

**Dossier zur Nutzenbewertung
gemäß § 35a SGB V**

Dapagliflozin (Forxiga[®] 5 mg/10 mg Filmtabletten)

Bristol-Myers Squibb GmbH & Co KGaA und
AstraZeneca GmbH

Modul 4B

*Zur Verbesserung der Blutzuckerkontrolle bei
erwachsenen Patienten mit Diabetes mellitus Typ 2 als
Ergänzung zu Diät und Bewegung,
in Add-on Kombination mit Metformin*

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	15
Abkürzungsverzeichnis	18
4 Modul 4 – allgemeine Informationen	21
4.1 Zusammenfassung der Inhalte von Modul 4.....	22
4.2 Methodik.....	36
4.2.1 Fragestellung.....	36
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	39
4.2.3 Informationsbeschaffung.....	41
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	42
4.2.3.2 Bibliografische Literaturrecherche.....	42
4.2.3.3 Suche in Studienregistern.....	44
4.2.3.4 Selektion relevanter Studien.....	45
4.2.4 Bewertung der Aussagekraft der Nachweise.....	45
4.2.5 Informationssynthese und -analyse.....	48
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	48
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	49
4.2.5.3 Meta-Analysen.....	53
4.2.5.4 Sensitivitätsanalysen.....	55
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	56
4.2.5.6 Indirekte Vergleiche.....	57
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	59
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	60
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	60
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	60
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	64
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	67
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	67
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	68
4.3.1.2.1 Studiendesign und Studienpopulationen.....	68
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	72
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	72
4.3.1.3.1 Hypoglykämien unter Berücksichtigung des HbA1c-Wertes – RCT.....	74
4.3.1.3.1.1 Ausmaß der Blutzuckerkontrolle (HbA1c) – RCT.....	74
4.3.1.3.1.2 Hypoglykämien – RCT.....	80
4.3.1.3.1.3 Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien.....	85
4.3.1.3.1.4 Gewichtsveränderung – RCT.....	87

4.3.1.3.1.5	Therapiezufriedenheit (DTSQ) – RCT	93
4.3.1.3.1.6	Blutdrucksenkung – RCT	96
4.3.1.3.1.7	Unerwünschte Ereignisse – RCT	101
4.3.1.3.1.8	Kardiale Ereignisse – RCT	105
4.3.1.3.2	Subgruppenanalysen – RCT	108
4.3.1.3.2.1	Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach dem Geschlecht – RCT	111
4.3.1.3.2.2	Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach dem Alter – RCT	114
4.3.1.3.2.3	Hypoglykämien stratifiziert nach dem Geschlecht – RCT	116
4.3.1.3.2.4	Hypoglykämien stratifiziert nach dem Alter – RCT	118
4.3.1.3.2.5	Gewichtsveränderung stratifiziert nach dem Geschlecht – RCT ..	119
4.3.1.3.2.6	Gewichtsveränderung stratifiziert nach dem Alter – RCT	122
4.3.1.3.2.7	Therapiezufriedenheit stratifiziert nach dem Geschlecht – RCT ..	124
4.3.1.3.2.8	Therapiezufriedenheit stratifiziert nach dem Alter – RCT	125
4.3.1.3.2.9	Blutdrucksenkung stratifiziert nach dem Geschlecht – RCT	125
4.3.1.3.2.10	Blutdrucksenkung stratifiziert nach dem Alter – RCT	128
4.3.1.3.2.11	Unerwünschte Ereignisse stratifiziert nach dem Geschlecht – RCT	131
4.3.1.3.2.12	Unerwünschte Ereignisse stratifiziert nach dem Alter – RCT	132
4.3.1.3.2.13	Kardiale Ereignisse stratifiziert nach dem Geschlecht – RCT	134
4.3.1.3.2.14	Kardiale Ereignisse stratifiziert nach dem Alter – RCT	135
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien	136
4.3.2	Weitere Unterlagen	140
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	140
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	141
4.3.2.1.1.1	Studien des pharmazeutischen Unternehmers	142
4.3.2.1.1.2	Studien aus der bibliografischen Literaturrecherche	146
4.3.2.1.1.3	Studien aus der Suche in Studienregistern	153
4.3.2.1.1.4	Resultierender Studienpool	156
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche	160
4.3.2.1.2.1	Studiendesign und Studienpopulationen	160
4.3.2.1.2.2	Verzerrungspotenzial auf Studienebene	193
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	198
4.3.2.1.3.1	Ausmaß der Blutzuckerkontrolle (HbA1c) – indirekte Vergleiche aus RCT	201
4.3.2.1.3.1.1	Hypoglykämien – indirekte Vergleiche aus RCT	213
4.3.2.1.3.1.2	Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien	222
4.3.2.1.3.1.3	Gewichtsveränderung – indirekte Vergleiche aus RCT	224
4.3.2.1.3.1.4	Blutdrucksenkung – indirekte Vergleiche aus RCT	235
4.3.2.1.3.1.5	Unerwünschte Ereignisse – indirekte Vergleiche aus RCT	241
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	261
4.3.2.2	Nicht randomisierte vergleichende Studien	261
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	261

4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	261
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	262
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien....	262
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	263
4.3.2.3	Weitere Untersuchungen.....	263
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	263
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	264
4.3.2.3.3	< Ergebnisse aus weiteren Untersuchungen.....	264
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	264
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	265
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen.....	265
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	266
4.4.1	Beurteilung der Aussagekraft der Nachweise	266
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	267
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	278
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	278
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	279
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	279
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	280
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	280
4.5.4	Verwendung von Surrogatendpunkten.....	280
4.6	Liste der eingeschlossenen Studien.....	282
4.7	Referenzliste.....	285
	Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche	298
	Anhang 4-B : Suchstrategien – Suche in Studienregistern.....	317
	Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund	321
	Anhang 4-D : Liste der abgebrochenen Studien	367
	Anhang 4-E : Liste der laufenden Studien.....	368
	Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....	371
	Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	538

Tabellenverzeichnis

	Seite
Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen	25
Tabelle 4-2: Zusammenfassende Darstellung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit: Metformin+Dapagliflozin gegen Metformin+Sulfonylharnstoff.....	34
Tabelle 4-3: Zusammenfassende Darstellung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit: Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitoren	36
Tabelle 4-4: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen	41
Tabelle 4-5: Methodik zur gemeinsame Betrachtung der Endpunkte Blutzuckerkontrolle (HbA1c) und Hypoglykämien.....	51
Tabelle 4-6: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	61
Tabelle 4-7: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	64
Tabelle 4-8: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	67
Tabelle 4-9: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	68
Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	69
Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	70
Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	70
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	71
Tabelle 4-14: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	72
Tabelle 4-15: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	73
Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	73
Tabelle 4-17: Operationalisierung von „Ausmaß der Blutzuckerkontrolle (HbA1c)“	75
Tabelle 4-18: Bewertung des Verzerrungspotenzials für „Ausmaß der Blutzuckerkontrolle (HbA1c)“ in RCT mit dem zu bewertenden Arzneimittel	75
Tabelle 4-19: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	77
Tabelle 4-20: Ergebnisse für „Ausmaß der Blutzuckerkontrolle (HbA1c)“ aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	78

Tabelle 4-21: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	79
Tabelle 4-22: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	80
Tabelle 4-23: Operationalisierung von Hypoglykämien	81
Tabelle 4-24: Bewertung des Verzerrungspotenzials für „Hypoglykämien“ in RCT mit dem zu bewertenden Arzneimittel	82
Tabelle 4-25: Ergebnisse für Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	83
Tabelle 4-26: Ergebnisse für Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	84
Tabelle 4-27: Operationalisierung von gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien	86
Tabelle 4-28: Bewertung des Verzerrungspotenzials für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien in RCT mit dem zu bewertenden Arzneimittel	86
Tabelle 4-29: Ergebnisse für „gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien“ aus RCT mit dem zu bewertenden Arzneimittel (Studien- und Zielpopulation)	87
Tabelle 4-30: Operationalisierung von Gewichtsveränderung	88
Tabelle 4-31: Bewertung des Verzerrungspotenzials für Gewichtsveränderung in RCT mit dem zu bewertenden Arzneimittel	89
Tabelle 4-32: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	90
Tabelle 4-33: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	91
Tabelle 4-34: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	92
Tabelle 4-35: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	92
Tabelle 4-36: Operationalisierung von Therapiezufriedenheit (DTSQ)	94
Tabelle 4-37: Bewertung des Verzerrungspotenzials für Therapiezufriedenheit (DTSQ) in RCT mit dem zu bewertenden Arzneimittel	95
Tabelle 4-38: Ergebnisse für Therapiezufriedenheit (DTSQ) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	95
Tabelle 4-39: Ergebnisse für Therapiezufriedenheit (DTSQ) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	96
Tabelle 4-40: Operationalisierung von Blutdrucksenkung	97
Tabelle 4-41: Bewertung des Verzerrungspotenzials für Blutdrucksenkung in RCT mit dem zu bewertenden Arzneimittel	98

Tabelle 4-42: Ergebnisse für Blutdrucksenkung aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation).....	99
Tabelle 4-43: Ergebnisse für Blutdrucksenkung aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	100
Tabelle 4-44: Operationalisierung von Unerwünschte Ereignisse.....	102
Tabelle 4-45: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel	102
Tabelle 4-46: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation).....	103
Tabelle 4-47: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	103
Tabelle 4-48: Ergebnisse für Unerwünschte Ereignisse – Anzeichen für Genital- und Harnwegsinfektionen aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	104
Tabelle 4-49: Ergebnisse für Unerwünschte Ereignisse – Anzeichen für Genital- und Harnwegsinfektionen aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	104
Tabelle 4-50: Operationalisierung von Kardiale Ereignisse	106
Tabelle 4-51: Bewertung des Verzerrungspotenzials für „Kardiale Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel.....	107
Tabelle 4-52: Ergebnisse für „Kardiale Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation).....	107
Tabelle 4-53: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	108
Tabelle 4-54: Interaktionstest der präspezifizierten Subgruppenanalysen für den Endpunkt Ausmaß der Blutzuckerkontrolle (HbA1c)	109
Tabelle 4-55: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	112
Tabelle 4-56: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation).....	113
Tabelle 4-57: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation).....	115
Tabelle 4-58: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	116
Tabelle 4-59: Ergebnisse für Hypoglykämien stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	117
Tabelle 4-60: Ergebnisse für Hypoglykämien stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	118
Tabelle 4-61: Ergebnisse für Gewichtsveränderung stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation).....	121
Tabelle 4-62: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation).....	122

Tabelle 4-63: Ergebnisse für Gewichtsveränderung stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	123
Tabelle 4-64: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) stratifiziert nach dem Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	124
Tabelle 4-65: Ergebnisse für Therapiezufriedenheit (DTSQ) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	124
Tabelle 4-66: Ergebnisse für Therapiezufriedenheit (DTSQ) stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	125
Tabelle 4-67: Ergebnisse für Blutdrucksenkung stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	126
Tabelle 4-68: Ergebnisse für Blutdrucksenkung stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	129
Tabelle 4-69: Ergebnisse für „Unerwünschte Ereignisse“ stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	131
Tabelle 4-70: Ergebnisse für Unerwünschte Ereignisse – Anzeichen für Genital- und Harnwegsinfektionen stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	132
Tabelle 4-71: Ergebnisse für Unerwünschte Ereignisse stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	133
Tabelle 4-72: Ergebnisse für Unerwünschte Ereignisse – Anzeichen für Genital- und Harnwegsinfektionen stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	134
Tabelle 4-73: Ergebnisse für Kardiale Ereignisse stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	135
Tabelle 4-74: Ergebnisse für Kardiale Ereignisse stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	136
Tabelle 4-75: Liste der Studien des pharmazeutischen Unternehmers – Studien für indirekte Vergleiche auf Basis von RCT	143
Tabelle 4-76: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel	146
Tabelle 4-77 : Relevante Studien aus der Suche in Studienregistern	153
Tabelle 4-78: Studienpool – Indirekte Vergleiche auf Basis von RCT	156
Tabelle 4-79: Charakterisierung der eingeschlossenen Studien – Studien für indirekte Vergleiche	162
Tabelle 4-80: Charakterisierung der Interventionen – Studien für indirekte Vergleiche	181
Tabelle 4-81: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche	186
Tabelle 4-82: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel	189
Tabelle 4-83: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel	194

Tabelle 4-84: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	198
Tabelle 4-85: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	200
Tabelle 4-86: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)	202
Tabelle 4-87: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)	203
Tabelle 4-88: Operationalisierung von Ausmaß der Blutzuckerkontrolle (HbA1c)	205
Tabelle 4-89: Bewertung des Verzerrungspotenzials für Ausmaß der Blutzuckerkontrolle (HbA1c) in RCT für indirekte Vergleiche	206
Tabelle 4-90: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT für indirekte Vergleiche (nach 24 Wochen).....	209
Tabelle 4-91: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT für indirekte Vergleiche (nach 52 Wochen).....	210
Tabelle 4-92: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT für indirekte Vergleiche – aktualisierte Literaturrecherche	211
Tabelle 4-93: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus indirekten Vergleiche nach 24 und 52 Wochen – Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel	212
Tabelle 4-94: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)	213
Tabelle 4-95: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)	214
Tabelle 4-96: Operationalisierung von Hypoglykämien	215
Tabelle 4-97: Bewertung des Verzerrungspotenzials für Hypoglykämien in RCT für indirekte Vergleiche	217
Tabelle 4-98: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche (nach 24 Wochen im Vergleich zu Metformin+Placebo)	219
Tabelle 4-99: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoff).....	220
Tabelle 4-100: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche – aktualisierte Literaturrecherche	221
Tabelle 4-101: Ergebnisse für Hypoglykämien aus indirekten Vergleiche nach 24 und 52 Wochen - Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel	221
Tabelle 4-102: Operationalisierung von gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien	222
Tabelle 4-103: Bewertung des Verzerrungspotenzials für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien in RCT mit dem zu bewertenden Arzneimittel.....	223

Tabelle 4-104: Ergebnisse für „gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien“ aus MTC mit dem zu bewertenden Arzneimittel.....	223
Tabelle 4-105: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)	224
Tabelle 4-106: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)	225
Tabelle 4-107: Operationalisierung von Gewichtsveränderung.....	227
Tabelle 4-108: Bewertung des Verzerrungspotenzials für „Gewichtsveränderung“ in RCT für indirekte Vergleiche	228
Tabelle 4-109: Ergebnisse für Gewichtsveränderung aus RCT für indirekte Vergleiche (nach 24 Wochen)	231
Tabelle 4-110: Ergebnisse für Gewichtsveränderung aus RCT für indirekte Vergleiche (nach 52 Wochen)	232
Tabelle 4-111: Ergebnisse für Gewichtsveränderung aus RCT für paarweise indirekte Vergleiche –aktualisierte Literaturrecherche	233
Tabelle 4-112: Ergebnisse für Gewichtsveränderung aus indirekten Vergleiche nach 24 und 52 Wochen - Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel	234
Tabelle 4-113: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)	235
Tabelle 4-114: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)	236
Tabelle 4-115: Operationalisierung von Blutdrucksenkung	237
Tabelle 4-116: Bewertung des Verzerrungspotenzials für Blutdrucksenkung in RCT für indirekte Vergleiche	238
Tabelle 4-117: Ergebnisse für systolische Blutdrucksenkung aus RCT für indirekte Vergleiche (nach 24 Wochen).....	239
Tabelle 4-118: Ergebnisse für systolische Blutdrucksenkung aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoff)	240
Tabelle 4-119: Ergebnisse für „Blutdrucksenkung“ aus RCT für indirekte Vergleiche – aktualisierte Literaturrecherche.....	240
Tabelle 4-120: Ergebnisse für systolische Blutdrucksenkung aus indirekten Vergleichen nach 24 Wochen- Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel....	241
Tabelle 4-121: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)	242
Tabelle 4-122: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)	243
Tabelle 4-123: Operationalisierung von unerwünschte Ereignisse.....	245
Tabelle 4-124: Bewertung des Verzerrungspotenzials für Gesamtrate unerwünschter Ereignisse in RCT für indirekte Vergleiche	248
Tabelle 4-125: Bewertung des Verzerrungspotenzials für Gesamtrate schwerwiegender unerwünschter Ereignisse in RCT für indirekte Vergleiche	249

Tabelle 4-126: Bewertung des Verzerrungspotenzials für Studienabbrüche aufgrund unerwünschter Ereignisse in RCT für indirekte Vergleiche	250
Tabelle 4-127: Ergebnisse für Gesamtrate unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 24 Wochen).....	253
Tabelle 4-128: Ergebnisse für Gesamtrate unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoffe) ...	254
Tabelle 4-129: Ergebnisse für Gesamtrate unerwünschter Ereignisse aus RCT für indirekte Vergleiche –aktualisierte Literaturrecherche	254
Tabelle 4-130: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 24 Wochen)	255
Tabelle 4-131: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoffe)	256
Tabelle 4-132: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT für indirekte Vergleiche –aktualisierte Literaturrecherche.....	256
Tabelle 4-133: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 24 Wochen)	257
Tabelle 4-134: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoffe)	258
Tabelle 4-135: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche –aktualisierte Literaturrecherche.....	259
Tabelle 4-136: Ergebnisse für schwerwiegende unerwünschte Ereignisse aus indirekten Vergleichen nach 24 und 52 Wochen - Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel	260
Tabelle 4-137: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse aus indirekten Vergleichen nach 24 und 52 Wochen - Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel	260
Tabelle 4-138: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	262
Tabelle 4-139: Operationalisierung von <Endpunkt xxx>.....	262
Tabelle 4-140: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	263
Tabelle 4-141: Operationalisierung von <Endpunkt xxx>.....	264
Tabelle 4-142: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	278
Tabelle 4-143: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	279
Tabelle 4-144: Ovid: Suchstrategie MEDLINE 1946 bis heute: am 11. Mai 2011	299
Tabelle 4-145: Ovid: Suchstrategie EMBASE 1980 bis heute: am 11. Mai 2011	301

Tabelle 4-146: Ovid: Suchstrategie Cochrane 1948 bis heute: am 11. Mai 2011.....	303
Tabelle 4-147: Ovid: Suchstrategie MEDLINE 1946 bis heute: am 04. Juni 2012.....	305
Tabelle 4-148: Ovid: Suchstrategie EMBASE 1980 bis heute: am 04. Juni 2012.....	307
Tabelle 4-149: Ovid: Suchstrategie Cochrane 1948 bis heute: am 04. Juni 2012	309
Tabelle 4-150: Ovid: Suchstrategie MEDLINE 1946 bis heute: am 09. Oktober 2012	310
Tabelle 4-151: Ovid: Suchstrategie EMBASE 1980 bis heute: am 09. Oktober 2012	312
Tabelle 4-152: Ovid: Suchstrategie Cochrane 1948 bis heute: am 09. Oktober 2012.....	315
Tabelle 4-153 (Anhang): Im Volltext ausgeschlossene nicht relevante Publikationen (nach Literaturrecherche vom 11. Mai 2011)	321
Tabelle 4-154 (Anhang): Im Volltext ausgeschlossene nicht relevante Publikationen (aktualisierte Literaturrecherche vom 04. Juni 2012)	363
Tabelle 4-155 (Anhang): Im Volltext ausgeschlossene nicht relevante Publikationen (aktualisierte Literaturrecherche vom 09. Oktober 2012).....	365
Tabelle 4-156 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel.....	367
Tabelle 4-157 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	369
Tabelle 4-158 (Anhang): Studiendesign und -methodik für Studie D1690C00004.....	372
Tabelle 4-159 (Anhang): Studiendesign und -methodik für Studie MB102014 (Bailey, 2010).....	381
Tabelle 4-160 (Anhang): Studiendesign und -methodik für die Studie D1690C00012 (Bolinder, 2012)	387
Tabelle 4-161 (Anhang): Studiendesign und -methodik für die Studie von Scott 2008	393
Tabelle 4-162 (Anhang): Studiendesign und -methodik für die Studie von Charbonnel 2006.....	397
Tabelle 4-163 (Anhang): Studiendesign und -methodik für die Studie von DeFronzo, 2009	402
Tabelle 4-164 (Anhang): Studiendesign und -methodik für die Studie von Bosi, 2007	407
Tabelle 4-165 (Anhang): Studiendesign und -methodik für die Studie von Taskinen 2011..	411
Tabelle 4-166 (Anhang): Studiendesign und -methodik für die Studie von Bergenstal, 2010a	416
Tabelle 4-167 (Anhang): Studiendesign und -methodik für die Studie von Raz, 2008	419
Tabelle 4-168 (Anhang): Studiendesign und -methodik für die Studie von Bolli, 2008	425
Tabelle 4-169 (Anhang): Studiendesign und -methodik für die Studie von Bergenstal, 2010b	430
Tabelle 4-170 (Anhang): Studiendesign und -methodik für die Studie von DeFronzo, 2005	437
Tabelle 4-171 (Anhang): Studiendesign und -methodik für die Studie von Nauck, 2009.....	442
Tabelle 4-172 (Anhang): Studiendesign und -methodik für die Studie von Pratley, 2010	448

Tabelle 4-173 (Anhang): Studiendesign und -methodik für die Studie von Kaku, 2009.....	453
Tabelle 4-174 (Anhang): Studiendesign und -methodik für die Studie von Charpentier 2001	458
Tabelle 4-175 (Anhang): Studiendesign und -methodik für die Studie von Papathanassiou, 2009	463
Tabelle 4-176 (Anhang): Studiendesign und -methodik für die Studie von Matthews, 2005	467
Tabelle 4-177 (Anhang): Studiendesign und -methodik für die Studie von Nauck, 2007.....	472
Tabelle 4-178 (Anhang): Studiendesign und -methodik für die Studie von Göke, 2010.....	477
Tabelle 4-179 (Anhang): Studiendesign und -methodik für die Studie von Filozof, 2010....	483
Tabelle 4-180 (Anhang): Studiendesign und -methodik für die Studie von Derosa 2011b...	488
Tabelle 4-181 (Anhang): Studiendesign und -methodik für die Studie von Derosa 2010.....	491
Tabelle 4-182 (Anhang): Studiendesign und -methodik für die Studie von Matthews, 2010	496
Tabelle 4-183 (Anhang): Studiendesign und -methodik für die Studie von Yang 2011.....	503
Tabelle 4-184 (Anhang): Studiendesign und -methodik für die Studie von Derosa 2011a ...	509
Tabelle 4-185 (Anhang): Studiendesign und -methodik für die Studie von Pan 2012	514
Tabelle 4-186 (Anhang): Studiendesign und -methodik für die Studie von DeFronzo 2012	519
Tabelle 4-187 (Anhang): Studiendesign und -methodik für die Studie von Gallwitz 2012 ..	525
Tabelle 4-188 (Anhang): Studiendesign und -methodik für die Studie von Yang 2012.....	532
Tabelle 4-189 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie D1690C00004.....	539
Tabelle 4-190 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie MB102014 / Bailey, 2010	570
Tabelle 4-191 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie D1690C00012.....	596
Tabelle 4-192 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Scott, 2008.....	622
Tabelle 4-193 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Charbonnel, 2006	645
Tabelle 4-194 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von DeFronzo, 2009	668
Tabelle 4-195 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Bosi, 2007.....	689
Tabelle 4-196 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Taskinen, 2011	713
Tabelle 4-197 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Bergenstal, 2010a	736
PanTabelle 4-198 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Bolli, 2008	746

Tabelle 4-199 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Bergenstal, 2010b	770
Tabelle 4-200 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von DeFronzo, 2005	796
Tabelle 4-201 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Nauck, 2009	819
Tabelle 4-202 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Pratley, 2010	845
Tabelle 4-203 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Kaku, 2009	871
Tabelle 4-204 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Charpentier, 2001	886
Tabelle 4-205 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Papathanassiou, 2009	896
Tabelle 4-206 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Matthews, 2005	906
Tabelle 4-207 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Nauck, 2007	929
Tabelle 4-208 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Göke, 2010	952
Tabelle 4-209 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Filozof, 2010	978
Tabelle 4-210 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Salvadeo, 2010	999
Tabelle 4-211 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Derosa, 2010	1009
Tabelle 4-212 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Matthews, 2010	1024
Tabelle 4-213 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Derosa, 2011a	1047
Tabelle 4-214 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Derosa 2011b	1063
Tabelle 4-215 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Yang et al., 2011	1078
Tabelle 4-216 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Pan et al., 2012	1104
Tabelle 4-217 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von DeFronzo., 2012	1127
Tabelle 4-218 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Gallwitz., 2012	1150

Tabelle 4-219 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten
für die Studie von Yang., 2012 1173

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	66
Abbildung 2: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	80
Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	85
Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	93
Abbildung 5: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	96
Abbildung 6: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	101
Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	105
Abbildung 8: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	108
Abbildung 9: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien zur Durchführung einer MTC (11. Mai 2011)	148
Abbildung 10: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien zur Durchführung einer MTC (04. Juni 2012)	150
Abbildung 11: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien zur Durchführung einer MTC (09. Oktober 2012)...	152
Abbildung 12: Netzwerk der Studien für Endpunkt „Ausmaß der Blutzuckerkontrolle“. A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk	204
Abbildung 13: Netzwerk der Studien für Endpunkt Hypoglykämien. A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk	215
Abbildung 14: Netzwerk der Studien für Endpunkt Gewicht. A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk	226
Abbildung 15: Netzwerk der Studien für Endpunkt Blutdrucksenkung. A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk	236
Abbildung 16: Netzwerk der Studien für Endpunkt Unerwünschte Ereignisse. A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk	244
Abbildung 17: Flussdiagramm Studiendesign und -methodik für Studie D1690C00004	380
Abbildung 18: Flussdiagramm Studiendesign und -methodik für Studie MB102014 (Bailey, 2010)	386
Abbildung 19: Flussdiagramm Studiendesign und -methodik für die Studie D1690C00012 (Bolinder, 2012)	392

Abbildung 20: Flussdiagramm Studiendesign und -methodik für die Studie von Scott 2008.....	396
Abbildung 21: Flussdiagramm Studiendesign und -methodik für die Studie von Charbonnel 2006.....	401
Abbildung 22: Flussdiagramm Studiendesign und -methodik für die Studie von DeFronzo, 2009.....	406
Abbildung 23: Flussdiagramm Studiendesign und -methodik für die Studie von Bosi 2007.....	410
Abbildung 24: Flussdiagramm Studiendesign und -methodik für die Studie von Taskinen, 2011.....	415
Abbildung 25: Flussdiagramm Studiendesign und -methodik für die Studie von Bergenstal 2010a.....	418
Abbildung 26: Flussdiagramm Studiendesign und -methodik für die Studie von Raz, 2008.....	424
Abbildung 27: Flussdiagramm Studiendesign und -methodik für die Studie von Bolli 2008.....	429
Abbildung 28: Flussdiagramm Studiendesign und -methodik für die Studie von Bergenstal 2010b.....	436
Abbildung 29: Flussdiagramm Studiendesign und -methodik für die Studie von DeFronzo, 2005.....	441
Abbildung 30: Flussdiagramm Studiendesign und -methodik für die Studie von Nauck, 2009.....	447
Abbildung 31: Flussdiagramm Studiendesign und -methodik für die Studie von Pratley, 2010.....	452
Abbildung 32: Flussdiagramm Studiendesign und -methodik für die Studie von Kaku, 2009.....	457
Abbildung 33: Flussdiagramm Studiendesign und -methodik für die Studie von Charpentier 2001.....	462
Abbildung 34: Flussdiagramm Studiendesign und -methodik für die Studie von Papathanassiou 2009.....	466
Abbildung 35: Flussdiagramm Studiendesign und -methodik für die Studie von Matthews, 2005.....	471
Abbildung 36: Flussdiagramm Studiendesign und -methodik für die Studie von Nauck, 2007.....	476
Abbildung 37: Flussdiagramm Studiendesign und -methodik für die Studie von Göke, 2010.....	482
Abbildung 38: Flussdiagramm Studiendesign und -methodik für die Studie von Filozof, 2010.....	487
Abbildung 39: Flussdiagramm Studiendesign und -methodik für die Studie von Derosa 2011b.....	490
Abbildung 40: Flussdiagramm Studiendesign und -methodik für die Studie von Derosa, 2010.....	495

Abbildung 41: Flussdiagramm Studiendesign und -methodik für die Studie von Matthews, 2010 und Ferrannini, 2009	502
Abbildung 42: Flussdiagramm Studiendesign und -methodik für die Studie von Yang 2011	508
Abbildung 43: Flussdiagramm Studiendesign und -methodik für die Studie von Derosa 2011a	513
Abbildung 44: Flussdiagramm Studiendesign und -methodik für die Studie von Pan, 2012	518
Abbildung 45: Flussdiagramm Studiendesign und -methodik für die Studie von DeFronzo 2012	524
Abbildung 46: Flussdiagramm Studiendesign und -methodik für Studie Gallwitz 2012	531
Abbildung 47: Flussdiagramm Studiendesign und -methodik für Studie Yang 2012	537

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ANCOVA	Analysis of covariance
ARR	Absolute Risiko Reduktion
AUC	Area under the curve
BMS/AZ	Bristol-Myers Squibb/AstraZeneca
BMI	Body-Mass-Index
BID	zweimal täglich
CONSORT	Consolidated Standards of Reporting Trials
DDG	Deutsche Diabetes-Gesellschaft
DMP	Disease-Management-Programm
DPP-4	Dipeptidyl-Peptidase-4
DTSQ	Diabetes Treatment Satisfaction Questionnaire
DTSQc	Diabetes Treatment Satisfaction Questionnaire, <i>Change-Version</i>
DTSQs	Diabetes Treatment Satisfaction Questionnaire, <i>Status-Version</i>
eGFR	Geschätzte glomeruläre Filtrationsrate
EKG	Elektrokardiogramm
EPAR	European Public Assessment Report
FEM	Fixed Effect Model
FFS	Freie Fettsäuren
FPG	Fasting plasma glucose
G-BA	Gemeinsamer Bundesausschuss
GLP-1	Glucagon-like peptide-1
HbA1c	Glykiertes Hämoglobin
HDL-C	High density lipoprotein-cholesterol
HOMA	Homeostasis Model Assessment
HIV	Human immunodeficiency virus
ICTRP	International Clinical Trials Registry Platform
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ISRCTN	International Standard Randomized Controlled Trial Number
ITT	Intention to treat
IWRS	Interactive web response system

Abkürzung	Bedeutung
KI	Konfidenzintervall
LDL-C	Low density lipoprotein-cholesterol
LOCF	Last observation carried forward
LT	Long term
MCMC	Markov-Chain-Monte-Carlo
MTC	Mixed Treatment Comparison (Netzwerk-Meta-Analyse)
MWD	Mittelwertdifferenz
NIDDM	Non-insulin-dependent diabetes mellitus
OAD	Orales Antidiabetikum
oGTT	Oraler Glukose-Toleranztest
OR	Odds ratio
PPG	Postprandial glucose
RCT	Randomized Controlled Trial
RSA-ÄndV	Verordnung zur Änderung der Risikostruktur-Ausgleichsverordnung
REM	Random Effect Model
RR	Relatives Risiko
SD	Standardabweichung
SE	Standardfehler
SGB	Sozialgesetzbuch
SGLT-2	Sodium Glucose Co-Transporter 2
SOC	System Organ Class
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TC	Totale Cholesterol
TG	Triglyzerid
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
TSH	Thyroid stimulierendes Hormon
TZD	Thiazolidinedione
UACR	Urin Albumin Kreatin Verhältnis
UE	Unerwünschte Ereignisse
ULN	Upper limit of normal

Abkürzung	Bedeutung
UKPDS	United Kingdom Prospective Diabetes Study
WHO	World Health Organization
XR	Extended release

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Für Arzneimittel, die zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt. Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, müssen jedoch vorgelegt werden. Zu diesem Zweck enthält die Dokumentvorlage den Abschnitt 4.4.4, der ausschließlich in Dossiers für die oben genannten Arzneimittel zur Behandlung eines seltenen Leidens bearbeitet werden soll. In diesem Abschnitt sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich der begründenden Nachweise vorzulegen. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

4.1 Zusammenfassung der Inhalte von Modul 4

Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.

Fragestellung

Kodierung B stellt den medizinischen Zusatznutzen von Dapagliflozin (Forxiga®) in der Zweifachkombination mit dem oralen Antidiabetikum Metformin (Anwendungsgebiet *Add-on* zu Metformin) gegenüber der zweckmäßigen Vergleichstherapie, der Kombination aus einem Sulfonylharnstoff und Metformin, sowie gegenüber der alternativen Vergleichstherapie Metformin und einem Dipeptidyl-Peptidase-4 (DPP-4)-Inhibitor, dar.

Es ergeben sich zwei Fragestellungen:

Primäre Fragestellung:

Ist für Metformin+Dapagliflozin in der Indikation Typ-2-Diabetes mellitus bei erwachsenen Patienten im Alter von 18 bis 74 Jahren, mit einer geschätzten glomerulären Filtrationsrate (eGFR) ≥ 60 mL/min/1,73 m² und ohne begleitender Therapie mit Schleifendiuretika, deren Blutzucker mit Metformin allein zusammen mit Diät und Bewegung nicht ausreichend kontrolliert ist, basierend auf den Daten aus RCT, ein medizinischer Zusatznutzen im Sinne des § 35a SGB V gegenüber der zweckmäßigen Vergleichstherapie nachweisbar?

Sekundäre Fragestellung:

Ist für Metformin+Dapagliflozin in der Indikation Typ-2-Diabetes mellitus bei erwachsenen Patienten im Alter von 18 bis 74 Jahren, mit einer eGFR ≥ 60 mL/min/1,73 m² und ohne begleitender Therapie mit Schleifendiuretika, deren Blutzucker mit Metformin allein zusammen mit Diät und Bewegung nicht ausreichend kontrolliert ist, und für die Sulfonylharnstoffe nicht geeignet sind und die Anwendung von Insulin noch nicht angezeigt ist, basierend auf den Daten aus RCT, ein medizinischer Zusatznutzen gegenüber der alternativen Vergleichstherapien nachweisbar?

Für die Bewertung des medizinischen Nutzens in der beschriebenen Indikation sind folgende patientenrelevante Endpunkte etabliert und wurden daher für die Zusatznutzenbewertung von Dapagliflozin, soweit Daten vorhanden, berücksichtigt:

- Hypoglykämien
- Gewichtsveränderungen
- Unerwünschte Ereignisse
- Folgekomplikationen und Mortalität
 - Kardiovaskuläre Morbidität und Mortalität
 - Zerebrovaskuläre Morbidität und Mortalität

- Gefäßbedingte nichtkardiale und nichtzerebrale Morbidität und Mortalität
 - Gesamtmortalität
 - Erblindung / Visusverschlechterung (Veränderungen des Augenhintergrundes)
 - Terminale Niereninsuffizienz mit Dialysenotwendigkeit
 - Amputationen
- Studienabbrüche aufgrund unzureichender glykämischer Kontrolle
 - Gesundheitsbezogene Lebensqualität und Therapiezufriedenheit
 - Hyperosmolare Komata
 - Stationäre Behandlungen

Darüber hinaus werden Angaben zum HbA1c- (glykiertes Hämoglobin) Wert als Maß für die langfristige Blutzuckerkontrolle extrahiert und dargestellt. Diese Angaben dienen zur Interpretation der Ergebnisse hinsichtlich patientenrelevanter Endpunkte, insbesondere bezüglich des Auftretens von Hypoglykämien. Aufgrund des Wirkmechanismus von Dapagliflozin und der Tatsache, dass Hypertonie einen zusätzlichen Risikofaktor für kardiovaskuläre Erkrankungen darstellt, wurden die Änderungen im systolischen und diastolischen Blutdruck als Surrogatparameter für das kardiovaskuläre Risiko berücksichtigt.

Datenquellen

Alle verfügbaren RCT zum zu bewertenden Arzneimittel und der zweckmäßigen Vergleichstherapie in der zu bewertenden Indikation wurden anhand einer umfassenden Informationsbeschaffung identifiziert (Abschnitt 4.3.1.1 sowie 4.3.2.1.1).

Für die Bewertung des medizinischen Zusatznutzens wurden die Daten aus dem Originalbericht der direkt vergleichenden randomisierten Phase III-Studie mit der Kombination aus Metformin+Dapagliflozin gegen die Kombination aus Metformin+Sulfonylharnstoff herangezogen. Die Originaldaten dieser Studie wurden für diese Nutzenbewertung zusätzlich für die Patientenpopulation, welche sich gemäß der gültigen Zulassung von Dapagliflozin für eine Behandlung mit Dapagliflozin eignen, ausgewertet. Dies sind die Patienten mit einer eGFR ≥ 60 mL/min/1,73 m² bzw. einer Kreatininclearance ≥ 60 mL/min, die <75 Jahre alt sind und nicht mit Schleifendiuretika vorbehandelt waren (Zielpopulation). Da eine direkt vergleichende Studie vorliegt, war die Durchführung eines indirekten Vergleiches nicht von Nöten und damit auch keine systematische Literaturrecherche für die zweckmäßige Vergleichstherapie.

Zur Identifizierung von relevanten Studien zu Dapagliflozin war keine bibliografische Literaturrecherche gefordert, da weltweit keine Zulassung vorliegt, die mehr als 12 Monate

vor Einreichung des Dossiers erteilt wurde. Die Erstzulassung wurde für Forxiga® am 12.11.2012 erteilt. Eine Literaturrecherche für das zu bewertende Arzneimittel entfiel damit.

Sowohl zum zu bewertenden Arzneimittel als auch zur Vergleichstherapie (Metformin+Sulfonylharnstoff) wurde eine Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien durchgeführt. Die Suche erfolgte in den öffentlich verfügbaren Studienregistern *clinicaltrials.gov* und dem *International Clinical Trials Registry Platform Search Portal* (ICTRP Search Portal).

Wie bereits im Modul 3, Kodierung B sowie Abschnitt 4.2.2 erläutert, sieht Bristol-Myers Squibb / AstraZeneca für die Patientengruppen, für die die Behandlung mit Sulfonylharnstoffen nicht in Frage kommt, die Kombinationsbehandlung von Metformin+DPP-4-Inhibitor als die zweckmäßige Vergleichstherapie an. Für diesen Zweck wurden die Ergebnisse einer Netzwerk-Meta-Analyse (*Mixed Treatment Comparison* [MTC]) dargestellt.

Dazu wurde eine bibliografische Literaturrecherche in den Datenbanken MEDLINE, EMBASE und Cochrane (*Cochrane Central Register of Controlled Trials*) durchgeführt. Für jede der Datenbanken wurde eine sequentielle und individuell adaptierte Suchstrategie entwickelt und separat angewandt.

Darüber hinaus wurde im Rahmen der MTC eine weitere Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien durchgeführt. Die Studienregistersuche für die MTC wurde in denselben Datenbanken wie für Dapagliflozin sowie im *International Standard Randomized Controlled Trial Number* (ISRCTN) Register durchgeführt.

Darüber hinaus wurden im Rahmen der MTC die Referenzlisten von geeigneten Studien und Meta-Analysen, welche in den letzten drei Jahren veröffentlicht wurden, nach relevanten Studien durchsucht.

Ergänzend zu der Literaturrecherche wurden für die MTC Tagungs- bzw. Konferenzberichte von verschiedenen Einrichtungen (*American College of Cardiology, American Diabetes Association, European Association for the Study of Diabetes, The Obesity Society und International Diabetes Federation*) für das Jahr 2010 durchsucht.

Ein-/Ausschlusskriterien für Studien

Die Ein- und Ausschlusskriterien der Studien zum Beleg des Zusatznutzens der Behandlung mit Dapagliflozin in Kombination mit Metformin sind in Tabelle 4-1 dargestellt.

Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen

	Einschlusskriterien	Ausschlusskriterien
Patientenpopulation	Erwachsene Patienten mit manifestem Typ-2-Diabetes, die unter einer Metformin-Therapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker aufweisen. ¹	Patienten mit anderen Diabetes Typen und behandlungsnaive Patienten. Tierexperimentelle Studien
Intervention	Behandlung mit Dapagliflozin (5 mg oder 10 mg einmal täglich) + Metformin	
Vergleichstherapie	Kombinationsbehandlung mit Metformin+Sulfonylharnstoff ²	Andere Kombinationstherapien oder Monotherapien
Endpunkte	Mindestens einer der oben genannten patientenrelevanten Endpunkte muss berücksichtigt werden (HbA1c-Wert nur in Kombination mit Hypoglykämien)	
Studientypen	RCT	Nicht randomisierte Studien, Review-Artikel <i>Case Reports</i>
Studiendauer	≥24 Wochen ³	<24 Wochen
Publikationstyp	Vollpublikation oder Bericht verfügbar, der den Kriterien des CONSORT Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht	Keine Vollpublikation oder Bericht verfügbar

RCT: Randomisierte kontrollierte Studie; HbA1c: Glykiertes Hämoglobin

¹ Unter besonderer Berücksichtigung von Patienten mit einer geschätzten glomerulären Filtrationsrate (eGFR) ≥60 mL/min/1,73 m² oder Kreatinin-clearance ≥60 mL/min und Patienten, die <75 Jahre alt sind und nicht mit Schleifendiuretika vorbehandelt waren (Zielpopulation).

² Für die *Mixed-Treatment-Comparison* (MTC) wurden auch andere Wirkstoffgruppen (inklusive Placebo) in Kombination mit Metformin zugelassen

³ In der MTC wurden Studien mit einer Mindestdauer von 18 Wochen eingeschlossen.

Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen

Die zur Quantifizierung des Zusatznutzens eingeschlossenen Studien wurden hinsichtlich ihrer Aussagekraft bewertet. Die Bewertung der eingeschlossenen Studien im Rahmen des vorliegenden Dossiers erfolgte anhand der zur Verfügung stehenden Informationen aus dem Studienbericht und aus Publikationen.

Die Bewertung des Verzerrungspotenzials erfolgte gemäß den Vorgaben der Dossievorlage. Zur Bewertung des Verzerrungspotenzials wurden sowohl allgemeine

(endpunktübergreifende) als auch endpunktspezifische Gesichtspunkte untersucht. Bei Verzerrungen auf Endpunktebene wurde jeder Endpunkt und gegebenenfalls jede Operationalisierung getrennt betrachtet. Auf Studienebene wurden folgende Verzerrungsaspekte betrachtet: Erzeugung der Randomisierungssequenz, Verdeckung der Gruppenzuteilung, Verblindung des Patienten sowie des behandelnden Arztes und ergebnisgesteuerte Berichterstattung. Auf Endpunktebene wurden, wie vorgegeben, die Aspekte Verblindung der Endpunkterheber, Umsetzung des *Intention-to-treat*-(ITT) Prinzips, ergebnisgesteuerte Berichterstattung berücksichtigt sowie die Behandlung fehlender Werte und *Drop-outs*.

Gemäß Verfahrensordnung wurde das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. Ein niedriges Verzerrungspotenzial lag dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führte nicht zum Ausschluss aus der Bewertung, sondern wurde in Sensitivitätsanalysen adressiert, diente der Diskussion heterogener Studienergebnisse und beeinflusste die Sicherheit der Aussage.

Mit Ausnahmen des HbA1c-Wertes sowie der Änderung im systolischen und diastolischen Blutdruck (letzteres ist ein Surrogatparameter für kardiovaskuläres Risiko) wurde jeder der in der Nutzen- bzw. Zusatznutzenbeurteilung betrachteten Endpunkte als patientenrelevant eingestuft (Abschnitt 4.2.5.2).

Für dichotome Endpunkte wurde der Therapieeffekt in der Regel anhand des *odds ratio* (OR) zusammengefasst. Im Fall, dass in einem Behandlungsarm kein Ereignis beobachtet wurde (Nullzelle), wurde für die Berechnung des OR eine Korrektur vorgenommen und 0,5 als fester Wert zu allen Zellen der Studienresultate hinzuaddiert. Gegebenenfalls, z. B. im Falle nicht hinreichend homogener Studienresultate oder im Falle hoher Ereignisraten, wurden auch andere Effektmaße berechnet (wie z. B. das relative Risiko [RR]) und für Sensitivitätsanalysen verwendet. Diese Effektmaße wurden jedoch nur dann im Dossier dargestellt, wenn es relevante Änderungen/Abweichungen im Vergleich zum OR gab. Für kontinuierliche Variablen wurden Mittelwertdifferenzen oder standardisierte Mittelwertdifferenzen (Hedges' *g*) berechnet.

Meta-Analysen

Für den Vergleich von Metformin+Dapagliflozin gegen Metformin+Sulfonylharnstoff wurde keine Meta-Analyse durchgeführt, da hier nur eine für die Fragestellung relevante Studie (eine direkt vergleichende Studie) vorliegt.

Für die Studien, die im Rahmen der MTC eingeschlossen wurden, konnte eine Meta-Analyse nur dann durchgeführt werden, falls die Daten hinreichend homogen waren. Die Meta-Analyse wurde dann mit der Software Review Manager (RevMan, Version 5) durchgeführt. Sofern berichtet, wurden primär die Resultate aus ITT-Analysen kombiniert, adjustierte Werte werden gegenüber ‚rohen‘ Werten bevorzugt, falls vorhanden. Alle Meta-Analysen basieren

auf einem Modell mit zufälligen Effekten (REM); Modelle mit festen Effekten (FEM) wurden im Rahmen einer Sensitivitätsanalyse verwendet. Zur Interpretation der Resultate basierend auf dem REM und dem FEM wurden Maße zur Bewertung und Quantifizierung von Heterogenität herangezogen. Falls Effekt oder Streuungsmaße der Einzelstudien nicht verfügbar waren, konnten diese, wenn möglich, aus Angaben aus den Studienberichten berechnet oder approximiert werden.

Zur Abschätzung der klinischen Relevanz wurden, falls keine validierten Irrelevanzgrenzen für den Gruppenunterschied vorlagen, Responderanalysen herangezogen. Bei Vorliegen eines validierten oder etablierten Responsekriteriums erfolgte eine Dichotomisierung im Sinne von Respondern/Non-Respondern anhand dieses Kriteriums. Danach erfolgte die Bewertung des Gruppenunterschieds anhand des OR (bzw. RR) für Response (bzw. Non-Response). Da bei Responderanalysen davon ausgegangen wird, dass die Responsedefinition eine Bewertung der klinischen Relevanz bereits auf individueller Ebene erlaubt, wurde ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen auch gleichzeitig als klinisch relevant gewertet.

Sensitivitätsanalysen und Subgruppenanalysen

Zur Überprüfung der Robustheit der verwendeten Methodik wurden neben den primären Analysen (*Last observation carried forward*, LOCF-Methode) auch *Repeated-Measurements*-Analysen als Sensitivitätsanalysen dargestellt.

Für die Zielpopulation gemäß Zulassung wurden die Ergebnisse hinsichtlich potenzieller Effektmodifikatoren, d. h. klinischer Faktoren, die die Effekte beeinflussen, untersucht. Möglicherweise variierende Effekte zwischen verschiedenen Subgruppen (Geschlecht und Alter [<65 Jahre und ≥ 65 bis <75 Jahre]) wurden mittels Interaktionstests bewertet. Diese wurden entweder aus vorliegenden Ergebnissen aus Regressionsanalysen, die Interaktionsterme beinhalten, oder aus den Ergebnissen von Subgruppenanalysen berechnet.

Ein p-Wert $<0,2$ lieferte einen Hinweis auf unterschiedliche Effekte zwischen den Subgruppen. Lag der p-Wert unter 0,05, so konnte von unterschiedlichen Effekten zwischen den Subgruppen ausgegangen werden (Beleg für eine Interaktion). Die Ergebnisse von Subgruppenanalysen wurden stets hinsichtlich ihrer Validität (Power, Präspezifizierung, qualitativ versus quantitativ) sowie hinsichtlich möglicher Auswirkungen auf das aus der Gesamtpopulation abgeleitete Fazit diskutiert.

Indirekte Vergleiche

Zur Bewertung des Zusatznutzens von Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitor wurden im Dossier auch indirekte Vergleiche in Form einer MTC dargestellt. Bei einer MTC können zwei oder mehr Behandlungsoptionen simultan miteinander verglichen und in eine Reihenfolge (*Ranking*) gebracht werden. Dabei wird die Evidenz aus direkten (*Head-to-Head*-Studien) und indirekten (Studien mit unterschiedlichen Komparatoren) Vergleichen miteinander kombiniert. Die Komparatoren zu Metformin+Dapagliflozin waren Glitazone (Thiazolidindione; TZD), *Glucagon-Like-Peptide-1* (GLP-1)-Agonisten, DPP-4-Inhibitoren oder Placebo jeweils als *Add-on* zu Metformin. Im Rahmen der Bewertung des

Zusatznutzens wurde nur der Vergleich Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitoren betrachtet.

Die indirekten Vergleiche im Rahmen der MTC wurden mit der Software WinBugs mithilfe der Bayesianischen Statistik auf Basis von Modellen mit zufälligen Effekten berechnet. Indirekte Vergleiche konnten nur dann durchgeführt werden, falls die Studien hinreichend homogen waren.

Da Bayesianische Modelle verwendet wurden, müssen noch Annahmen bezüglich A-priori-Verteilungen vorgenommen werden, sowie für die *Markov-Chain-Monte-Carlo*-(MCMC)-Simulationen Anzahl, Startwerte und Länge der gewählten Markov-Ketten. Im Folgenden wurde die nicht informative A-priori-Verteilung (Unwissenheit) verwendet, da diese keinen Einfluss auf die A-posteriori-Verteilung hat. Sowohl für dichotome und stetige Parameter (hier: Effektschätzer für eine Behandlung relativ zur Referenzbehandlung) ist die nicht-informative A-priori-Verteilung eine Normalverteilung mit Mittelwert 0 und Varianz 10.000. Auch für die *nuisance*-Parameter (hier studienspezifischer Mittelwert / *log-Odds* für die Referenzbehandlung) wurde diese A-priori-Verteilung angenommen. Bei Random-Effekt-Modellen wurden noch die Varianzen der beobachteten Effekte zwischen den Studien benötigt, für diese Varianz wurde ebenfalls eine A-priori-Verteilung benötigt. Hier wurde eine Gleichverteilung angenommen, deren Bereich für stetige Parameter klinisch plausibel ist. Um die A-posteriori-Verteilung zu schätzen, wurden die MCMC-Simulationen verwendet. Es wurden 2 Markov-Ketten mit unterschiedlichen Startwerten jedes zu schätzenden Parameters initiiert. Jede Kette hat 20.000 *burn-in*-Iterationen und mindestens 100.000 Iterationen. Jeder zu schätzende Parameter soll einen Monte Carlo-*Error* von weniger als 5% des Stichprobenstandardfehlers haben.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Zu einigen der vorab definierten Zielgrößen liegen keine relevanten Studien vor. Dies gilt insbesondere für die folgenden Endpunkte: Erblindung/Visusverschlechterung, terminale Niereninsuffizienz, Amputation, zerebrovaskuläre Morbidität und Mortalität, gefäßbedingte nichtkardiale und nichtzerebrale Morbidität und Mortalität, stationäre Behandlungen, hyperosmolare Komata und gesundheitsbezogene Lebensqualität.

Die in der folgenden Beschreibung des Zusatznutzens betrachteten Ergebnisse beziehen sich auf die Zielpopulation gemäß Zulassung (Typ-2-Diabetiker, die unter einer Metformin-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen, unter Ausschluss von Patienten mit einer eGFR <60 mL/min/1,73 m² bzw. Kreatinin-clearance <60 mL/min, Patienten, die ≥ 75 Jahre alt sind bzw. mit Schleifendiuretika vorbehandelt waren).

Die Ergebnisse zum **medizinischen Zusatznutzen** basieren auf einer direkt vergleichenden RCT, welche die Behandlung von Metformin+Dapagliflozin mit der Behandlung von Metformin+Sulfonylharnstoff vergleicht. Im Hinblick auf die Evidenzstufe von RCT (Evidenzstufe Ib), der hohen Studienqualität und Validität der Endpunkte besitzen die erbrachten Nachweise zum Zusatznutzen von Dapagliflozin im Vergleich zu einem

Sulfonylharnstoff eine hohe Validität und Aussagekraft. Damit können für alle betrachteten Endpunkte Aussagen gemacht werden, die als Hinweis gewertet werden müssen.

Darüber hinaus wurde eine Netzwerk-Meta-Analyse (*Mixed Treatment Comparison, MTC*) herangezogen, um einen Vergleich von Metformin+Dapagliflozin mit Metformin+DPP-4-Inhibitoren durchzuführen. In diese flossen ausschließlich RCT ein. Einschränkungen ergaben sich hinsichtlich der Operationalisierung einzelner Endpunkte. Damit können für alle betrachteten Endpunkte Aussagen gemacht werden, die als Anhaltspunkt gewertet werden müssen.

Metformin+Dapagliflozin gegen Metformin+Sulfonylharnstoff

Eine optimierte Blutzuckerkontrolle (und damit die HbA1c-Einstellung) wird von der Deutschen Diabetes-Gesellschaft (DDG) empfohlen (Zielvorgabe HbA1c <6,5%, sofern dies ohne das vermehrte Auftreten von Hypoglykämien und Gewichtszunahmen erreicht werden kann). Kann dieser Zielwert unter der OAD-Kombinationstherapie nicht erreicht werden, wird die Kombinationstherapie mit einem anderen oralen Antidiabetikum (OAD) bzw. einem GLP-1-Agonisten empfohlen. Der aktuelle Entwurf der Nationalen VersorgungsLeitlinie (NVL) zur Therapieplanung bei Typ-2-Diabetes empfiehlt einen HbA1c-Korridor von 6,5% bis 7,5% unter Berücksichtigung der individualisierten Therapieziele. Ähnliche Ausführungen zur Individualisierung der Therapieziele finden sich in einem aktuellen Positionspapier der Europäischen und Amerikanischen Diabetesgesellschaften (Consensus statement).

In der großen klinischen Studie „United Kingdom Prospective Diabetes Study“ (UKPDS) wurde gezeigt, dass sich bereits bei einer Senkung des HbA1c um 1% das Risiko mikrovaskulärer Erkrankungen um 37% und das Risiko makrovaskulärer Erkrankungen um 12-16% verringert. In der *Action in Diabetes and Vascular Disease: Preterax and Diamicon Modified Release Controlled Evaluation (ADVANCE)*-Studie bewirkte die intensiviertere antihyperglykämische Therapie eine signifikante Reduktion des Neuauftretens bzw. der Progression der Nephropathie und des kombinierten Endpunktes „mikro- und makrovaskuläre Erkrankungen“. Eine optimale Blutzuckereinstellung stellt somit ein wichtiges Therapieziel dar, insbesondere wenn sie unter Vermeidung von Nebenwirkungen wie z. B. Hypoglykämien erreicht werden kann.

In aktuellen Meta-Analysen prospektiv-randomisierter Studien bei Patienten mit Typ-2-Diabetes fand sich ferner eine signifikante Reduktion nicht-tödlicher Myokardinfarkte durch eine intensive Blutzuckersenkung. Schließlich legt die *Follow-up* Untersuchung der Patienten der UKPDS die Schlussfolgerung nahe, dass eine frühe normnahe Blutzucker-Einstellung auf lange Sicht zu einer Reduktion mikro- und makrovaskulärer Endpunkte sowie einer Reduktion der Sterblichkeit führen kann.

Hypoglykämien stellen vor allem bei Patienten mit lange bestehendem Diabetes und makrovaskulären Veränderungen einen erheblichen Risikofaktor für kardiovaskuläre Episoden dar. So kann eine Hypoglykämie zu Störungen des myokardialen Blutflusses und zu myokardialen Ischämien führen und damit myokardiale Schäden verursachen. Besonders Patienten mit kardiovaskulären Vorschädigungen haben ein hohes Risiko für kardiovaskuläre

Episoden nach einer Hypoglykämie, die auch zum Tod führen können. Auch bestimmte Berufsgruppen wie z. B. LKW-Fahrer oder Personen, die in großer Höhe arbeiten, sowie allein lebende Senioren sind speziell hypoglykämiegefährdet. Die Gefahr von Hypoglykämien beeinträchtigt die Fahrtüchtigkeit und kann sogar zum Entzug der Fahrerlaubnis führen. Für die Erlaubnis zum Führen eines Fahrzeuges wird aus juristischer Perspektive eine stabile Diabeteseinstellung ohne Hypoglykämien vorausgesetzt. Aus diesem Grunde kann die Therapie mit einem Sulfonylharnstoff regelhaft zur Notwendigkeit von Fahrverboten mit entsprechenden Konsequenzen für Mobilität und Berufstätigkeit der Patienten führen. Aufgrund des gesteigerten Hypoglykämierisikos kann ferner eine regelmäßige Blutzucker-Selbstkontrolle angezeigt sein, was mit einer entsprechenden Belastung für den Patienten und auch einem Anstieg der Therapie assoziierten Kosten verbunden sein kann.

Erschwerend kommt hinzu, dass Todesfälle durch Hypoglykämien vermutlich nicht in vollem Umfang in der Literatur berichtet werden, so dass deren tatsächliche Zahl systematisch unterschätzt wird und dieses Problem in der Realität noch größer ist als allgemein bekannt. Eine blutzuckersenkende Therapie, die das Auftreten von Hypoglykämien weitgehend vermeidet, stellt somit ein wichtiges Ziel der Diabetestherapie dar. Nicht zuletzt führen Hypoglykämien regelhaft zu Hospitalisierung und zusätzlicher Beanspruchung von Ressourcen des Gesundheitssystems.

Der Einfluss von Hypoglykämien auf die gesundheitsbezogene Lebensqualität von Typ-2-Diabetikern wurde in mehreren Studien untersucht. Es konnte gezeigt werden, dass das Auftreten sowie die Schwere einer Hypoglykämie die gesundheitsbezogene Lebensqualität mindert. Durch die ständige Angst dieser Patienten, eine Hypoglykämie zu erleiden, kann es zu Beeinträchtigungen des täglichen Soziallebens kommen. Die von den Patienten berichtete Angst vor einer hypoglykämischen Episode kann so groß sein, dass sie zu verminderter Therapieadhärenz führen kann (siehe Modul 3, Abschnitt 3.2.2). Dazu kommt, dass die Patienten durch häufigeres Essen versuchen, den Hypoglykämien entgegenzuwirken, was wiederum zur Gewichtszunahme führt. Eine bedeutsame Vermeidung der Hypoglykämien kann daher die Lebensqualität der Patienten voraussichtlich deutlich verbessern.

In Hinblick auf die Interpretation der direkten Vergleichsstudie zwischen Dapagliflozin und Glipizid muss somit berücksichtigt werden, dass Probleme der mangelnden Adhärenz in kontrollierten klinischen Studien in der Regel nicht zum Tragen kommen, was möglicherweise zu einer systematischen Nutzen-Überbewertung von Sulfonylharnstoffen in klinischen Studien im Vergleich zur praktischen Therapie führt.

Die **Blutzuckerkontrolle**, gemessen am HbA1c, unterschied sich nach 52 Wochen zwischen den beiden Behandlungsgruppen Metformin+Dapagliflozin versus Metformin+Sulfonylharnstoff nicht. Beide Behandlungen konnten den HbA1c-Wert effektiv und klinisch relevant (Reduktion des HbA1c >0,5%) senken. Die HbA1c-Senkung unter Metformin+Dapagliflozin blieb über die Zeit stabil, während die glykämische Kontrolle unter Metformin+Sulfonylharnstoff nachließ. Nach 104 Wochen war die HbA1c-Senkung in der Metformin+Dapagliflozin-Gruppe verglichen mit der Metformin+Sulfonylharnstoff-Gruppe größer (Mittelwertdifferenz [MWD] [95%-KI]: -0,19% [-0,34;-0,03]). Studienabbrüche

aufgrund einer unzureichenden glykämischen Kontrolle traten sowohl nach 52 Wochen als auch nach 104 Wochen statistisch signifikant häufiger in der Metformin+Sulfonylharnstoff-Gruppe auf (OR [95%-KI]: 0,08 [0,01;0,59] / 0,65 [0,44;0,97]).

Sowohl nach 52 als auch nach 104 Wochen konnte gezeigt werden, dass bei Patienten, die mit Metformin+Dapagliflozin behandelt wurden, signifikant weniger **Hypoglykämien** auftraten als bei den mit Metformin+Sulfonylharnstoff behandelten Patienten (OR [95%-KI] 0,05 [0,03;0,09] nach 52 Wochen / OR [95%-KI] 0,05 [0,03;0,09]; absolute Risiko Reduktion [ARR]: 44,2% nach 104 Wochen). Schwere Hypoglykämien waren insgesamt selten und wurden bei 0,8% der Patienten, die mit Metformin+Sulfonylharnstoff behandelt wurden, beobachtet. Bei den mit Metformin+Dapagliflozin behandelten Patienten trat dieses Ereignis nicht auf. Auch Studienabbrüche aufgrund einer Hypoglykämie traten bei Patienten, die mit Metformin+Dapagliflozin behandelt wurden nicht auf, wohingegen 6 (1,7%) Patienten, die mit der zweckmäßigen Vergleichstherapie behandelt wurden, die Studie aufgrund einer Hypoglykämie abbrachen.

Die unter Metformin+Dapagliflozin erreichte adäquate Blutzuckerkontrolle bei gleichzeitiger Hypoglykämievermeidung ist eine bedeutsame Vermeidung von Nebenwirkungen, sodass sich ein Hinweis auf einen beträchtlichen Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie (Metformin+Sulfonylharnstoff) ergibt.

Bezogen auf die **Gewichtsveränderung** zeigte sich in der direkten Vergleichsstudie ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen Metformin+Dapagliflozin gegen Metformin+Sulfonylharnstoff (MWD [95%-KI]: -5,02 kg [-5,57;-4,48] nach 52 Wochen und -5,38 kg [-6,07;-4,68] nach 104 Wochen). Die mit Metformin+Dapagliflozin behandelten Patienten nahmen im Mittel um 3,49 kg / 3,82 kg ab, während die mit Metformin+Sulfonylharnstoff behandelten Patienten an Gewicht zunahmen (1,54 kg / 1,56 kg). Der Anteil der Patienten, die mehr als 5% an Gewicht verloren hatten (präspezifizierte Responder-Analyse), war in der Metformin+Dapagliflozin-Gruppe signifikant größer. Während nach 52 Wochen nur 2,6% der Patienten aus der Metformin+Sulfonylharnstoff-Gruppe mindestens 5% abgenommen hatten, lag der Anteil an Patienten in der Metformin+Dapagliflozin-Gruppe bei 36,6% (OR [95%-KI]: 22,01 [10,93;44,33]). Nach 104 Wochen hatten 2,6% der Patienten aus der Metformin+Sulfonylharnstoff-Gruppe und 25,5% in der Metformin+Dapagliflozin-Gruppe mindestens 5% abgenommen (OR [95%-KI]:13,10 [6,45;26,60]).

In einer placebokontrollierten Studie, die auch in diesem Anwendungsgebiet durchgeführt wurde (D1690C00012) und speziell darauf ausgerichtet war, den Einfluss auf das Gewicht zu untersuchen, konnte gezeigt werden, dass der Gewichtsverlust hauptsächlich auf eine Reduktion der Körperfettmasse, des viszeralen Fettgewebes und des subkutanen Fettgewebes beruht.

Die signifikant größere und anhaltende Gewichtsreduktion unter Metformin+Dapagliflozin stellt eine bedeutsame Vermeidung von Nebenwirkungen dar und trägt zur erfolgreichen Diabetestherapie bei. Durch das bedeutsam höhere Auftreten relevanter Gewichtsreduktionen

(>5%) ergibt sich für Metformin+Dapagliflozin ein Hinweis für einen beträchtlichen Zusatznutzen.

Die Änderung in der **Therapiezufriedenheit** wurde mit dem *Diabetes Treatment Satisfaction Questionnaire (Change-Version, DTSQc)* nach 52 Wochen gemessen. Dieser kann Werte von -18 bis +18 annehmen, wobei positive Werte eine Verbesserung bedeuten. Ausgehend von ähnlich hohen Ausgangswerten im DTSQs (Status-Version) gaben Patienten unter Metformin+Dapagliflozin im Mittel einen DTSQc Skalenwert von 14,1 (Standardfehler [SE]: 0,317) an, während Patienten unter Metformin+Sulfonylharnstoff durchschnittlich einen Skalenwert von 13,6 Punkten (SE: 0,312) berichteten. Zwischen den Gruppen war kein statistisch signifikanter Unterschied vorhanden.

Damit ergibt sich kein Hinweis auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen in Bezug auf die Therapiezufriedenheit, gemessen mit dem DTSQc Fragebogen.

Patienten mit einer Metformin+Dapagliflozin-Behandlung hatten nach 52 Wochen eine höhere **Blutdrucksenkung** als Patienten mit einer Metformin+Sulfonylharnstoff-Behandlung. Der Unterschied zwischen den beiden Behandlungsgruppen war für den systolischen Blutdruck statistisch signifikant (MWD [95%-KI]: -5,3 mmHg [-7,1;-3,6]) und hielt auch nach 104 Wochen an (MWD [95%-KI]: -4,46 mmHg; [-6,67;-2,26]). Der Unterschied in der Reduktion des diastolischen Blutdrucks mit Metformin+Dapagliflozin gegenüber Metformin+Sulfonylharnstoff wies einen vergleichbaren Trend auf, war aber im Ausmaß geringer und erreichte keine statistische Signifikanz.

Aufgrund der unter Metformin+Dapagliflozin erzielten stärkeren Senkung des systolischen Blutdrucks, wird das Auftreten von nicht schwerwiegenden Symptomen verringert, sodass sich ein Hinweis auf einen geringen Zusatznutzen ergibt.

Die Gesamtrate der Patienten, die ein **unerwünschtes Ereignis**, ein schwerwiegendes unerwünschtes Ereignis oder einen Studienabbruch aufgrund eines unerwünschten Ereignisses hatten, war in beiden Gruppen nach 52 sowie nach 104 Wochen ähnlich. Es gab keine statistisch signifikanten Unterschiede. Auch der Anteil der Patienten, die ein kardiales Ereignis hatten, ließ keinen Unterschied zwischen den beiden Gruppen erkennen.

Bei den Anzeichen für **Genital- oder Harnwegsinfektionen** ergaben sich signifikante Unterschiede. Sie traten häufiger bei den Patienten auf, die mit Metformin+Dapagliflozin behandelt wurden, und zwar sowohl nach 52 Wochen (OR [95%-KI]: Anzeichen für Genitalinfektion: 5,92 [2,83;12,37], Anzeichen für Harnwegsinfektion: 2,02 [1,16;3,50]) als auch nach 104 Wochen (OR [95%-KI]: Anzeichen für Genitalinfektion: 6,67 [3,32;13,37], Anzeichen für Harnwegsinfektion: 1,76 [1,10;2,81]).

Die aufgetretenen Anzeichen für Genital- und Harnwegsinfektionen waren meist nur mild oder moderat in ihrer Intensität, mit den gängigen Behandlungsmethoden einfach zu behandeln und führten nur sehr selten zum Studienabbruch. Die Unterschiede bei den

Anzeichen für Harnwegsinfektionen betrafen den unteren Harntrakt. Das bestätigen auch die Ergebnisse einer Untersuchung zur Häufigkeit von Urogenitalinfektionen bei Typ-2-Diabetikern. Die Auswertung zeigte, dass Urogenitalinfektionen eine häufige Komplikation des Diabetes mellitus Typ 2 sind und die Mehrzahl der Infektionen in der Regel einfach zu behandeln sind und nicht zu schweren Verläufen führt.

Da sich die Gesamtrate der unerwünschten Ereignisse in den beiden Behandlungsgruppen nicht voneinander unterschied, ist nicht von einem Schaden auszugehen, denn Patienten unter einer Behandlung mit Metformin+Sulfonylharnstoff erleiden andere unerwünschte Ereignisse häufiger als unter Metformin+Dapagliflozin. Hier sind vor allem Schwindel (9,3% der mit Metformin+Sulfonylharnstoff behandelten Patienten im Vergleich zu 5,2% der mit Metformin+Dapagliflozin behandelten Patienten), vermehrtes Schwitzen (7,8% der Patienten in der Metformin+Sulfonylharnstoff-Gruppe im Vergleich zu 1,5% derer in der Metformin+Dapagliflozin-Gruppe) sowie Zittern (7,8% der mit Metformin+Sulfonylharnstoff behandelten Patienten gegenüber 0,5% der mit Metformin+Dapagliflozin behandelten Patienten) zu nennen.

Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitoren

Wie bereits im Rahmen des G-BA (Gemeinsamer Bundesausschuss)-Gespräches angemerkt, gibt es Patienten, bei denen die zweckmäßige Vergleichstherapie nicht angewendet werden kann. Das ist darin begründet, dass es unter der Behandlung mit Sulfonylharnstoffen vermehrt zum Auftreten von Hypoglykämien und zu Gewichtszunahmen kommt (siehe Abschnitt 4.4.2).

Wie in Abschnitt 4.2.2 erläutert sieht Bristol-Myers Squibb / AstraZeneca bei Patienten, für die Sulfonylharnstoffe nicht in Frage kommen, die Behandlung mit Metformin+DPP-4-Inhibitor als zweckmäßige Vergleichstherapie an. Die Ergebnisse der MTC zeigten gegenüber Metformin+DPP-4-Inhibitoren, dass Metformin+Dapagliflozin sowohl nach 24 als auch nach 52 Wochen eine signifikant stärkere Gewichtsreduktion bewirkt (MWD [95%-Kredibilitätsintervall]: -2,37 kg [-2,98;-1,77] nach 24 Wochen / -2,74 kg [-5,35;-0,10] nach 52 Wochen). Bezüglich den anderen betrachteten Endpunkten (Blutzuckerkontrolle (HbA1c), Hypoglykämien, Blutdrucksenkung, unerwünschte Ereignisse) gab es keinen Anhaltspunkt auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen. Durch die signifikant stärkere Gewichtsreduktion ergibt sich ein Anhaltspunkt auf einen geringen Zusatznutzen von Metformin+Dapagliflozin gegenüber Metformin+DPP-4-Inhibitor.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Anhand der direkt vergleichenden Studie D1690C00004 wurde der Zusatznutzen des zu bewertenden Arzneimittels Dapagliflozin in Kombination mit Metformin gegenüber der Kombination aus Metformin+Sulfonylharnstoff für erwachsene Typ-2 Diabetiker, deren Blutzucker mit der maximal verträglichen Dosis von Metformin alleine nicht ausreichend kontrolliert werden kann, beurteilt. Die folgende Tabelle stellt die Quantifizierung des Zusatznutzens zusammenfassend für alle betrachteten Endpunkte dar.

Tabelle 4-2: Zusammenfassende Darstellung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit: Metformin+Dapagliflozin gegen Metformin+Sulfonylharnstoff

Endpunkt	Quantifizierung des Zusatznutzens	Wahrscheinlichkeit
Hypoglykämien unter Berücksichtigung des HbA1c	Beträchtlich	Hinweis
Gewichtsveränderung	Beträchtlich	Hinweis
Therapiezufriedenheit	Kein Zusatznutzen	Hinweis
Blutdrucksenkung	Gering	Hinweis
Unerwünschte Ereignisse*	Kein Schaden	Hinweis
Kardiale Ereignisse	Kein Schaden	Hinweis

* Bewertung erfolgte für die Gesamtrate der unerwünschten Ereignisse, schwerwiegenden unerwünschten Ereignisse und Studienabbrüche aufgrund unerwünschter Ereignisse

Das metabolische Syndrom, dessen wichtigster Einzelfaktor die Adipositas darstellt, ist unter anderem wesentlich für das hohe kardiovaskuläre bzw. Makroangiopathierisiko verantwortlich, welches wiederum die wichtigste Ursache für die Morbidität und Sterblichkeit bei Typ-2-Diabetikern darstellt. Eine Verminderung des Körpergewichts stellt somit zusammen mit einer Ernährungsumstellung und vermehrter körperlicher Aktivität die Grundlage einer Diabetestherapie dar und spielt für die Patienten damit eine ähnlich wichtige Rolle wie die Blutzuckereinstellung. Insbesondere der Anteil der Körperfettmasse korreliert positiv mit den wichtigsten kardiovaskulären Risikofaktoren und es scheint, dass eine Verringerung der Fettmasse diese Risikofaktoren positiv beeinflussen kann.

Wie die Adipositas, so ist auch die essentielle Hypertonie mit einer Insulinresistenz assoziiert und begünstigt die Entwicklung der Mikro- und Makroangiopathie. Vor allem ein hoher systolischer Blutdruck ist mit der Entwicklung einer Mikroalbuminurie und einer Proteinurie assoziiert. Die Senkung eines erhöhten Blutdrucks kann die kardio- und zerebrovaskuläre Morbidität und Mortalität innerhalb weniger Jahre reduzieren.

Des Weiteren ist die Vermeidung von Hypoglykämien ein wichtiges Therapieziel bei Typ-2-Diabetikern. Hypoglykämien stellen einen erheblichen Risikofaktor für kardiovaskuläre Ereignisse dar, insbesondere bei Patienten, die schon lange an Diabetes mellitus Typ-2 erkrankt sind und makrovaskuläre Veränderungen aufweisen. Auch bestimmte Berufsgruppen, wie z. B. LKW-Fahrer oder Personen, die in großer Höhe arbeiten, sowie allein lebende Senioren sind speziell durch Hypoglykämien gefährdet. Erschwerend kommt hinzu, dass vermutlich die tatsächliche Zahl der Todesfälle durch Hypoglykämien systematisch unterschätzt wird und dieses Problem in der Realität noch größer ist als allgemein bekannt. Weiterhin konnte in mehreren Studien gezeigt werden, dass die Häufigkeit des Auftretens sowie die Schwere einer Hypoglykämie die gesundheitsbezogene Lebensqualität mindert. Durch die ständige Angst, eine Hypoglykämie zu erleiden, kann es zu Beeinträchtigungen z. B. beim Autofahren, bei der Freizeitgestaltung und damit im

Sozialleben kommen. Die von den Patienten berichtete Angst vor einer hypoglykämischen Episode kann so groß sein, dass sie zu verminderter Therapieadhärenz führen kann.

Mit diesem Hintergrund ergibt sich aus der Zusammenschau der Ergebnisse ein Hinweis auf einen beträchtlichen Zusatznutzen von Dapagliflozin bei Patienten mit Diabetes Mellitus Typ 2, die mit einer Metformin-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen.

Dapagliflozin weist ein Wirkprofil auf, welches das kardiovaskuläre Risiko senken sollte. Wie im *European Public Assessment Report* (EPAR) beschrieben, zeigt sich in einer Meta-Analyse kardiovaskulärer Ereignisse (kombinierter Endpunkt – kardiovaskulärer Tod, Herzinfarkt oder Schlaganfall) aller Phase IIb und III Studien, welche verblindet unabhängig adjudiziert wurden, ein positiver Trend für Dapagliflozin. Eine große Phase III Outcome-Studie ist geplant und sollte in Kürze beginnen.

Die Behandlung mit Metformin+Dapagliflozin stellt damit eine wirksame, langanhaltende (über 2 Jahre) und im Allgemeinen gut verträgliche blutzuckersenkende Therapie mit einem positiven Nutzen-Risiko-Verhältnis und geringem Risiko für die überaus belastenden und potenziell lebensgefährlichen Hypoglykämien dar. Die im Gegensatz zu Metformin+Sulfonylharnstoff auch nach 2 Jahren noch anhaltende effektive Blutzuckerkontrolle lässt vermuten, dass Metformin+Dapagliflozin unabhängig von der Diabetesdauer und damit unabhängig vom fortschreitenden Funktionsverlust der pankreatischen Inselzellen wirken kann.

Für Dapagliflozin gilt es zu berücksichtigen, dass seine insulinunabhängige Wirkungsweise und seine einmal tägliche Gabe (unabhängig von den Mahlzeiten) zu einer geringeren Komplexität der Therapie beitragen. Entsprechend sind einfach zu handhabende Therapieregimes wie z. B. Dapagliflozin -bei vergleichbarer Wirksamkeit und Sicherheit unter Studienbedingungen- unter Alltagsbedingungen vorteilhaft, weil damit auch die Therapie-Adhärenz verbessert werden kann.

Die Anwenderfreundlichkeit von Dapagliflozin ergibt sich zudem daraus, dass die Dosis im Gegensatz zu den Sulfonylharnstoffen normalerweise nicht fortlaufend während der Behandlung angepasst werden muss. Für die Dosistitrierung der Sulfonylharnstoffe sind häufigere Blut- und Harnzuckermessungen sowie Arztbesuche notwendig, die bei einer Therapie mit Dapagliflozin nicht notwendig sind. Darüber hinaus leiden die meisten Typ-2-Diabetiker unter vielen Komorbiditäten und müssen deshalb viele Begleitmedikamente einnehmen. Für die meisten in der Diabetestherapie gängigen Begleitmedikamente konnten keine klinisch relevanten Wechselwirkungen mit Dapagliflozin nachgewiesen werden, wodurch Dapagliflozin mit einer Vielfalt dieser Begleitmedikamente kombiniert werden kann.

Für Patienten, die nicht mit Sulfonylharnstoffen behandelt werden können, wurde der Zusatznutzen des zu bewertenden Arzneimittels Metformin+Dapagliflozin gegenüber der Kombination aus Metformin+DPP-4-Inhibitoren anhand einer MTC beurteilt. Die folgende

Tabelle stellt die Quantifizierung des Zusatznutzens zusammenfassend für alle betrachteten Endpunkte dar.

Tabelle 4-3: Zusammenfassende Darstellung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit: Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitoren

Endpunkt	Quantifizierung des Zusatznutzens	Wahrscheinlichkeit
Hypoglykämien unter Berücksichtigung des HbA1c	Kein Zusatznutzen	Anhaltspunkt
Gewichtsveränderung	Gering	Anhaltspunkt
Blutdrucksenkung	Kein Zusatznutzen	Anhaltspunkt
Unerwünschte Ereignisse	Kein Schaden	Anhaltspunkt

Bei vergleichbarer Sicherheit der beiden Interventionen ergibt sich aufgrund der Tatsache, dass Patienten, die mit Metformin+Dapagliflozin behandelt werden, signifikant mehr Gewicht verlieren als Patienten, die mit Metformin+DPP-4-Inhibitoren behandelt werden, insgesamt ein Anhaltspunkt auf einen geringen Zusatznutzen.

4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Diabetes mellitus Typ 2 (früher nicht-insulinabhängiger Diabetes mellitus, NIDDM) ist eine verbreitete chronisch progrediente Stoffwechselerkrankung mit dem Leitsymptom der Hyperglykämie, die zu makrovaskulären und mikrovaskulären Komplikationen führen kann. Diese sind mit Abstand die wichtigste Ursache für Morbidität und Sterblichkeit von Typ-2-Diabetikern. Dem Diabetes mellitus Typ 2 liegt eine Insulinresistenz in Verbindung mit einer gestörten Insulinsekretion zugrunde.

Intervention

Bei der untersuchten Intervention handelt es sich um Forxiga[®], eine Filmtablette mit dem Wirkstoff Dapagliflozin Propandiol Monohydrat. Dapagliflozin ist der erste zugelassene Vertreter der sogenannten Natrium-Glukose-Cotransporter 2 (*Sodium Glucose Co-Transporter 2*, SGLT-2)-Inhibitoren und ist als Monopräparat zugelassen. Es kann als restriktive Monotherapie oder in *Add-on*-Kombination mit anderen oralen Antidiabetika und/oder Insulin verabreicht werden. Dapagliflozin wird einmal täglich in der zugelassenen Wirkstärke (5 mg oder 10 mg Dapagliflozin) verabreicht. Die empfohlene Dosierung liegt bei 10 mg Dapagliflozin einmal täglich. Die Dosierung von 5 mg Dapagliflozin dient als Einstiegs-Dosis bei Patienten mit schwerer Lebererkrankung und kann bei guter Verträglichkeit auf 10 mg erhöht werden.

Kodierung B stellt den medizinischen Zusatznutzen von Dapagliflozin in der Zweifachkombination mit dem oralen Antidiabetikum Metformin (*Add-on* zu Metformin) dar.

Fragestellung

Die primäre Fragestellung dieser Nutzenbewertung adressiert den Zusatznutzen von Dapagliflozin gegenüber der vom G-BA festgelegten zweckmäßigen Vergleichstherapie, der Kombinationsbehandlung mit Sulfonylharnstoff+Metformin. Zusätzlich wird der medizinische Zusatznutzen von Dapagliflozin im Vergleich zu einer alternativen Vergleichstherapie (DPP-4-Inhibitoren+Metformin) beurteilt. Somit ergeben sich zwei Fragestellungen.

Primäre Fragestellung:

Ist für Metformin+Dapagliflozin in der Indikation Typ-2-Diabetes mellitus bei erwachsenen Patienten im Alter von 18 bis 74 Jahren, mit einer eGFR ≥ 60 ml/min/1,73 m² und ohne begleitender Therapie mit Schleifendiuretika, deren Blutzucker mit Metformin allein zusammen mit Diät und Bewegung nicht ausreichend kontrolliert ist, basierend auf den Daten aus RCT, ein medizinischer Zusatznutzen im Sinne des § 35a SGB V gegenüber der zweckmäßigen Vergleichstherapie nachweisbar?

Sekundäre Fragestellung:

Ist für Metformin+Dapagliflozin in der Indikation Typ-2-Diabetes mellitus bei erwachsenen Patienten im Alter von 18 bis 74 Jahren, mit einer eGFR ≥ 60 ml/min/1,73 m² und ohne begleitender Therapie mit Schleifendiuretika, deren Blutzucker mit Metformin allein

zusammen mit Diät und Bewegung nicht ausreichend kontrolliert ist, und für die Sulfonylharnstoffe nicht geeignet sind und die Anwendung von Insulin noch nicht angezeigt ist, basierend auf den Daten aus RCT, ein medizinischer Zusatznutzen gegenüber der alternativen Vergleichstherapien nachweisbar?

Vergleichstherapie

Die definierte zweckmäßige Vergleichstherapie für dieses Anwendungsgebiet ist die Kombination aus einem Sulfonylharnstoff und Metformin.

Für Patientengruppen, für die die Behandlung mit Sulfonylharnstoffen nicht in Frage kommt, stellt die Kombinationsbehandlung aus Metformin und einem DPP-4-Inhibitor aus Sicht von BMS / AZ die geeignete Vergleichstherapie dar (Begründung siehe Abschnitt 4.2.2).

Patientenpopulation

Die primäre Fragestellung dieses Nutzendossiers betrifft die Patientenpopulation gemäß Zulassung. Dapagliflozin ist in *Add-on*-Kombination mit Metformin indiziert für erwachsene Patienten mit manifestem Diabetes mellitus Typ 2, die trotz Diät und Bewegung sowie einer Metformin-Monotherapie einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen. Im Rahmen des Zulassungsprozesses ergab sich eine Zielpopulation von Dapagliflozin, die Patienten mit einer eGFR ≥ 60 mL/min/1,73 m² bzw. einer Kreatinin-clearance ≥ 60 mL/min, die < 75 Jahre alt sind und nicht mit Schleifendiuretika behandelt werden, einschließt. Um den Zulassungsstatus abzubilden, werden die Resultate dieser Teilpopulation (im Folgenden Zielpopulation genannt) präsentiert. Darüber hinaus werden zur Bestätigung die Ergebnisse der per Protokoll geplanten Gesamtpopulation der relevanten Studien betrachtet (im Folgenden Studienpopulation genannt).

Patientenrelevante Endpunkte in RCT zum Nachweis des Nutzens und Zusatznutzens

Für die Bewertung des medizinischen Nutzens und Zusatznutzens in der beschriebenen Indikation werden die in Abschnitt 4.2.2 genannten patientenrelevante Endpunkte berücksichtigt.

Darüber hinaus werden Angaben zum HbA1c-Wert als Maß für die langfristige Blutzuckerkontrolle extrahiert und dargestellt. Diese Angaben dienen zur Interpretation der Ergebnisse hinsichtlich patientenrelevanter Endpunkte, insbesondere bezüglich des Auftretens von Hypoglykämien. Aufgrund des Wirkmechanismus von Dapagliflozin und der Tatsache, dass Hypertonie ein zusätzlicher Risikofaktor für kardiovaskuläre Erkrankungen ist, werden die Änderungen im systolischen und diastolischen Blutdruck als Surrogatparameter für das kardiovaskuläre Risiko berücksichtigt.

Studientypen

Zu den oben aufgeführten patientenrelevanten Endpunkten ist eine Bewertung anhand von RCT möglich. RCT liefern für die Bewertung des Nutzens einer medizinischen Intervention die zuverlässigsten Ergebnisse und sind mit der geringsten Ergebnisunsicherheit behaftet. Aus

diesem Grund basiert die Bewertung des Nutzens und des Zusatznutzens von Dapagliflozin bevorzugt auf den Daten aus RCT.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.

Ziel der Ein- und Ausschlusskriterien ist es, diejenigen Studien zu charakterisieren, die grundsätzlich für die Beantwortung der Fragestellung gemäß 4.2.1 geeignet sind. Die Ein- und Ausschlusskriterien für Studien werden wie folgt definiert und begründet:

Population: Die Wahl der Selektionskriterien für die Studienpopulation ergibt sich aus der diesem Modul zugrunde liegenden Zielpopulation, für die Dapagliflozin zugelassen ist (Bristol-Myers Squibb und AstraZeneca GmbH, 2012a): erwachsene Patienten mit manifestem Diabetes mellitus Typ 2, die unter einer Metformin-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker aufweisen.

Intervention: Behandlung mit Dapagliflozin in der zugelassenen Wirkstärke (einmal täglich Dapagliflozin 5 mg oder 10 mg) in Kombination mit Metformin.

Vergleichstherapie: Die zweckmäßige Vergleichstherapie entspricht der vom G-BA im Beratungsgespräch am 20.12.2011 (Hein, 2012) festgelegten Therapie: eine Kombinationsbehandlung mit Metformin und einem Sulfonylharnstoff. Für die Begründung der Wahl der zweckmäßigen Vergleichstherapie siehe Modul 3B, Abschnitt 3.1.2.

Wie bereits im G-BA Beratungsgespräch angemerkt wurde, gibt es Patientengruppen, für die die Behandlung mit Sulfonylharnstoffen nicht in Frage kommt. Das sind zum einen Patienten, die eine Unverträglichkeit / Kontraindikation gegenüber Sulfonylharnstoffen haben und auch Patienten, für die das Auftreten von Hypoglykämien, welche unter der Behandlung mit Sulfonylharnstoffen vermehrt auftreten, problematisch sein kann (z. B. Bus- oder LKW-Fahrer oder Patienten, die in großer Höhe oder mit gefährlichen Maschinen arbeiten).

Für diese Patienten stellt damit die Kombinationsbehandlung von einem DPP-4-Inhibitor mit Metformin die geeignete orale Vergleichstherapie dar. Diese sind weder mit dem Auftreten von Hypoglykämien assoziiert, noch haben sie einen negativen Einfluss auf das Gewicht.

Endpunkte: Mindestens einer der folgenden patientenrelevanten Endpunkte muss berichtet sein:

- Hypoglykämien
- Gewichtsveränderungen
- Unerwünschte Ereignisse
- Folgekomplikationen und Mortalität
 - Kardiovaskuläre Morbidität und Mortalität
 - Zerebrovaskuläre Morbidität und Mortalität
 - Gefäßbedingte nichtkardiale und nichtzerebrale Morbidität und Mortalität
 - Gesamtmortalität
 - Erblindung / Visusverschlechterung (Veränderungen des Augenhintergrundes)
 - Terminale Niereninsuffizienz mit Dialysenotwendigkeit
 - Amputationen
- Studienabbrüche aufgrund unzureichender glykämischer Kontrolle
- Gesundheitsbezogene Lebensqualität und Therapiezufriedenheit
- Hyperosmolare Komata
- Stationäre Behandlungen

Diese Endpunkte stellen in der Indikation etablierte patientenrelevante Endpunkte dar und ermöglichen eine Beurteilung patientenrelevanter Therapieziele.

Darüber hinaus dienen Angaben zum HbA1c-Wert als Maß für die langfristige Blutzuckerkontrolle zur Interpretation der Ergebnisse hinsichtlich der patientenrelevanten Therapieziele, insbesondere bezüglich des Auftretens von Hypoglykämien. Die Änderungen im systolischen und diastolischen Blutdruck werden als Surrogatparameter für das kardiovaskuläre Risiko berücksichtigt. Weitere Details sind unter 4.2.5.2 näher beschrieben.

Studientyp: Da diese Endpunkte im Rahmen von RCT beurteilt werden können, schränkt das Einschlusskriterium zum Studientyp die Nutzenbewertung auf RCT ein, da RCT, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt werden, mit der geringsten Ergebnisunsicherheit behaftet sind.

Studiendauer: Entsprechend bisheriger Nutzenbewertungen von Arzneimitteln im Indikationsgebiet werden ausschließlich Studien mit einer Behandlungsdauer von mindestens 24 Wochen eingeschlossen, da Sulfonylharnstoffe in der Regel über mehrere Wochen auf-

titriert werden müssen und um einen Effekt der Intervention auf patientenrelevante Therapieziele (z. B. Gewichtsveränderungen) nachweisen zu können.

Tabelle 4-4: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen

	Einschlusskriterien	Ausschlusskriterien
Patientenpopulation	Erwachsene Patienten mit manifestem Typ-2-Diabetes, die unter einer Metformin-Therapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker aufweisen. ¹	Patienten mit anderen Diabetes Typen und behandlungsnaive Patienten.
Intervention	Behandlung mit Dapagliflozin (5 mg oder 10 mg einmal täglich) + Metformin	
Vergleichstherapie	Kombinationsbehandlung mit Metformin+Sulfonylharnstoff ²	Andere Kombinationstherapien oder Monotherapien
Endpunkte	Mindestens einer der oben genannten patientenrelevanten Endpunkte muss berücksichtigt werden (HbA1c-Wert nur in Kombination mit Hypoglykämien)	
Studientypen	RCT	Nicht randomisierte Studien, Review-Artikel <i>Case Reports</i>
Studiendauer	≥24 Wochen ³	<24 Wochen ³
Publikationstyp	Vollpublikation oder Bericht verfügbar, der den Kriterien des CONSORT Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht ⁴	Keine Vollpublikation oder Bericht verfügbar

RCT: Randomisierte kontrollierte Studie; HbA1c: Glykiertes Hämoglobin

¹ Unter besonderer Berücksichtigung von Patienten mit einer glomerulären Filtrationsrate ≥ 60 mL/min/1,73 m² oder Kreatinin-clearance ≥ 60 mL/min und Patienten, die <75 Jahre alt sind und nicht mit Schleifendiuretika vorbehandelt waren (Zielpopulation). Sofern nur eine Teilpopulation der Patienten der RCT die Ein- und Ausschlusskriterien erfüllt, so muss es sich um eine zahlenmäßig relevante Teilpopulation handeln (mindestens N=50 randomisierte Patienten pro Behandlungsarm).

² Glimperid, Glibenclamid oder Glipizid. Für die *Mixed-Treatment-Comparison* (MTC) wurden auch andere Wirkstoffgruppen (inklusive Placebo) in Kombination mit Metformin zugelassen.

³ In der MTC wurden Studien mit einer Mindestdauer von 18 Wochen eingeschlossen.

⁴ Studien, deren Ergebnisse auf der Webseite offizieller Studienregister publiziert wurden (z.B. clinicaltrials.gov) können ebenfalls berücksichtigt werden.

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen

Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist in folgenden Fällen erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt:

- Für das zu bewertende Arzneimittel wurde eine Zulassung in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wurde, mindestens 12 Monate vor dem Zeitpunkt der Dossiereinreichung erteilt. Dabei sind auch Zulassungen außerhalb Deutschlands bzw. Europas relevant.
- Es werden indirekte Vergleiche zur Bewertung des medizinischen Zusatznutzens herangezogen.

Sofern eine bibliografische Literaturrecherche erforderlich ist, soll diese mindestens in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Geben Sie dabei an, ob Sie eine gemeinsame Suche für

mehrere Fragestellungen durchgeführt haben (z. B. gemeinsame Suche nach Studien mit dem zu bewertenden Arzneimittel und Studien zur zweckmäßigen Vergleichstherapie für indirekte Vergleiche). Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, ob bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Zur Identifizierung von relevanten Studien zu Dapagliflozin ist keine bibliografische Literaturrecherche gefordert, da weltweit keine Zulassung vorliegt, die mehr als 12 Monate vor Einreichung des Dossiers erteilt wurde. Die Erstzulassung wurde für Forxiga® am 12.11.2012 erteilt. Eine Literaturrecherche für das zu bewertende Arzneimittel entfällt damit.

Da für die vom G-BA festgelegte Vergleichstherapie (Metformin+Sulfonylharnstoff) eine direkt vergleichende Studie vorliegt, ist die Durchführung eines indirekten Vergleiches nicht von Nöten und damit auch keine systematische Literaturrecherche für die zweckmäßige Vergleichstherapie.

Wie bereits erläutert, sieht BMS / AZ für die Patientengruppen, für die die Behandlung mit Sulfonylharnstoffen nicht in Frage kommt, die Kombinationsbehandlung von Metformin+DPP-4-Inhibitor als die zweckmäßige Vergleichstherapie an. Für diesen Zweck und um den Vergleich zu weiteren Antidiabetika (z. B. Glitazone, GLP-1-Analoga) herzustellen, wurde eine Netzwerk-Meta-Analyse (*Mixed Treatment Comparison*; MTC) herangezogen. Die für die MTC durchgeführte bibliografische Literaturrecherche wird nachfolgend beschrieben.

Beschreibung der Suche zur Durchführung einer MTC

Unter der Fragestellung „Identifikation von Studien (RCT) um die Sicherheit und Wirksamkeit von Dapagliflozin mit anderen Antidiabetika zu vergleichen“ wurde eine gemeinsame Suche nach Studien mit Dapagliflozin und Studien mit anderen Arzneimitteln zur Behandlung von Patienten mit Typ-2-Diabetes durchgeführt, mit dem Ziel eine MTC durchführen zu können.

Die bibliografische Literaturrecherche wurde in den Datenbanken MEDLINE, EMBASE und Cochrane (*Cochrane Central Register of Controlled Trials*) durchgeführt. Für jede der Datenbanken wurde eine sequentielle und individuell adaptierte Suchstrategie entwickelt und separat angewandt.

Bei allen Recherchen wurden nur Humanstudien berücksichtigt. Mithilfe von validierten Filtern wurde eine Einschränkung auf RCT vorgenommen. Die Publikationssprache wurde nicht eingeschränkt. Die Suchstrategien für diese Datenbanken sind in einheitlichen Blöcken, getrennt nach Intervention, Indikation und Studientypen aufgebaut. Zur Identifikation von Suchbegriffen für Diabetes und die einzelnen Komparatoren wurden Cochrane Reviews zur Hilfe genommen und mit *MeSH-terms* gesucht. Eine Einschränkung auf den Suchzeitraum wurde nicht vorgenommen. Die Suche wurde mithilfe der Ovid® Plattform durchgeführt.

Darüber hinaus wurden die Referenzlisten von geeigneten Studien und Meta-Analysen, welche in den letzten drei Jahren veröffentlicht wurden, nach relevanten Studien durchsucht.

Wurden ergänzend zur Publikation weitere oder detailliertere Informationen benötigt (z. B. wenn nur qualitative Werte (z. B. Differenz und p-Wert) aber keine quantitativen Werte angegeben waren oder die Daten nur grafisch aber nicht in Zahlen präsentiert wurden), wurde versucht den Prüfarzt oder den Sponsor zu kontaktieren.

Ergänzend zu der Literaturrecherche wurden Tagungs- bzw. Konferenzberichte von verschiedenen Einrichtungen (*American College of Cardiology, American Diabetes Association, European Association for the Study of Diabetes, The Obesity Society und International Diabetes Federation*) für das Jahr 2010 durchsucht.

Die detaillierte Vorgehensweise und die Ergebnisse der systematischen bibliografischen Literaturrecherche sowie die Ergebnisse aus den Referenzlisten und Tagungsberichten sind in einem separaten Bericht (Oxford Outcomes Ltd., 2011b), welcher 2 Mal aktualisiert wurde (Oxford Outcomes Ltd., 2012a; Oxford Outcomes Ltd., 2012c) und in Abschnitt 4.3.2.1.1.2 beschrieben.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien von Dritten vollständig identifiziert werden.

Die Suche soll mindestens in den Studienregistern clinicaltrials.gov, clinicalstudyresults.org sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden.

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Beschreiben Sie nachfolgend, in welchen Studienregistern die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, ob bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.

Um sicherzustellen, dass alle verfügbaren Studiendaten über Dapagliflozin für die Bewertung berücksichtigt werden, wird eine Studienregistersuche nach abgeschlossenen, abgebrochenen

und laufenden Studien mit dem zu bewertenden Arzneimittel Dapagliflozin in den gängigen medizinischen Studienregistern anhand der nachfolgend dargestellten Methodik durchgeführt.

Die detaillierte Suchstrategie und deren Ergebnisse sind in Anhang 4-B dokumentiert.

Die Suche erfolgt entsprechend den Vorgaben in den Studienregistern *clinicaltrials.gov* und über das *ICTRP Search Portal*. Da das Studienregister *clinicalstudyresults.org* nicht mehr zur Verfügung steht, wurde hier keine Suche vorgenommen.

Darüber hinaus wurde im Rahmen der MTC eine weitere Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien durchgeführt. Die Studienregistersuche für die MTC wurde in denselben Datenbanken wie für Dapagliflozin sowie im ISRCTN-Register durchgeführt.

Die detaillierte Suchstrategie und deren Ergebnisse sind in einem separaten Bericht (Oxford Outcomes Ltd., 2011b) dokumentiert und in Anhang 4-B sowie Abschnitt 4.3.2.1.1 dargestellt.

4.2.3.4 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2 und 4.2.3.3 beschriebenen Rechenschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die in der Studienregistersuche identifizierten Studien für das zu bewertende Arzneimittel wird unter Anwendung der in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien nach Titel und Abstracts unabhängig von zwei Personen selektiert. Diskrepanzen werden durch Diskussion aufgelöst. Die selektierten Studien werden mit dem schon bekannten Studienpool abgeglichen und neue, passende Studien werden in einen Studienpool zusammengeführt.

Die Literaturrecherche und Studienregistersuche für die MTC wurde genauso gehandhabt und auch nach den in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien definiert. Eine MTC dient dazu mehrere Therapieoptionen simultan miteinander zu vergleichen. Deshalb wurden neben Dapagliflozin auch andere Interventionen wie Sulfonylharnstoffe, DPP-4-Inhibitoren, Glitazone und GLP-1-Analoga eingeschlossen¹. Um einen möglichst großen Studienpool zu erhalten wurden Studien mit einer Mindestdauer von 18 Wochen berücksichtigt.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und

¹ Da die zweckmäßige Vergleichstherapie für Patienten, für die die Behandlung mit Sulfonylharnstoffen nicht geeignet ist, die Behandlung mit einem DPP-4-Inhibitor darstellt, wird der indirekte Vergleich im Rahmen der MTC nur für den Vergleich Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitor beschrieben.

zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-G):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie des Behandlers
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die zur Quantifizierung des Zusatznutzens eingeschlossenen Studien werden hinsichtlich ihrer Aussagekraft bewertet. Die Bewertung der eingeschlossenen Studien im Rahmen des vorliegenden Dossiers erfolgt anhand der zur Verfügung stehenden Informationen aus dem Studienbericht und aus Publikationen.

Bewertung des Verzerrungspotenzials

Sowohl allgemeine (endpunktübergreifende) als auch endpunktspezifische Gesichtspunkte werden untersucht. Bei Verzerrungen auf Endpunktebene wird jeder Endpunkt und gegebenenfalls jede Operationalisierung getrennt betrachtet. Entsprechend der Verfahrensordnung des G-BA zur Bewertung des Nutzens von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V vom 20. Januar 2011 werden dabei folgende Aspekte bewertet:

Verzerrungsaspekte der Ergebnisse auf Studienebene:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des behandelnden Arztes
- Ergebnisgesteuerte Berichterstattung

Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- Ergebnisgesteuerte Berichterstattung
- Sonstige Aspekte: Behandlung fehlender Werte und *Drop-outs*

Die Bewertungen des Verzerrungspotenzials der Einzelstudien werden tabellarisch zusammengefasst. Gemäß Verfahrensordnung wird das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern.

Abschließend wird unter Berücksichtigung der oben genannten Aspekte das Verzerrungspotenzial auf Ebene der Endpunkte als „hoch“ oder „niedrig“ eingestuft. Von einem ‚niedrigen‘ Verzerrungspotenzial wird ausgegangen, wenn mit großer

Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Eine relevante Verzerrung liegt vor, falls sich die Aussagen und Schlussfolgerungen bezüglich der Ergebnisse bei Behebung des verzerrenden Aspektes verändern würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ wird in der Bewertung berücksichtigt. In diesem Fall werden dann Sensitivitätsanalysen durchgeführt. Die Klassifizierung dient der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

Die Vorgehensweise bei der Bewertung des Verzerrungspotenzials für jede Studie und jeden Endpunkt (Operationalisierung) richtete sich nach den Hinweisen der Bewertungsbögen in Anhang 4-G.

Eine zusammenfassende Bewertung der Verzerrungsaspekte erfolgte ausschließlich für RCT.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)². Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-³ bzw. STROBE-Statements⁴ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Für alle für die gegebene Fragestellung relevanten Studien werden Detailinformationen zu Design und Methodik extrahiert und im Abschnitt 4.3.1.2.1 beschrieben. Die Beschreibung erfolgt gemäß den Punkten 2b bis 14 des CONSORT-Statements (siehe Anhang 4-F). Der Patientenfluss wird gemäß CONSORT-Flow-Chart dargestellt.

² Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c332.

³ Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Publ Health* 2004; 94(3): 361-366.

⁴ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 4.3.2 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

Benennen Sie die Patientencharakteristika und Endpunkte, zu denen Ergebnisse dargestellt werden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Zur Charakterisierung der in die Nutzenbewertung eingeschlossenen Studien hinsichtlich der untersuchten Patientenpopulation werden sowohl demografische als auch krankheitsspezifische Daten für die Einzelstudien zusammengefasst. Folgende Faktoren werden berücksichtigt: Alter, Geschlecht, Gewicht, *Body-Mass-Index* (BMI), HbA1c vor Randomisierung, Krankheitsdauer und Anzahl der Studienabbrecher. Stetige Variablen werden anhand des Mittelwertes und der Standardabweichung (SD) beschrieben, kategorielle Merkmale anhand der beobachteten Häufigkeiten bzw. Anteile.

Die Ergebnisse der in den Studien berichteten patientenrelevanten Endpunkte werden entsprechend der vorliegenden Operationalisierung gegenübergestellt und vergleichend beschrieben. Hierbei werden die in den Studienberichten dargestellten Maßzahlen, wie z. B. adjustierte Mittelwerte mit Standardfehler bzw. Anteile und entsprechende Mittelwertdifferenzen berichtet. Zusätzlich dazu wird für dichotome Zielvariablen das OR als das vom G-BA präferierte Effektmaß berechnet. Im Fall, dass in einem Behandlungsarm kein Ereignis beobachtet wird (Nullzelle), wird für die Berechnung des OR eine Korrektur vorgenommen und 0,5 als fester Wert zu allen Zellen der Studienresultate hinzuaddiert. Alle Effektmaße werden mit den jeweiligen Intervallschätzern angegeben. Wenn nicht in den Publikationen angegeben, wurde der Standardfehler für die Differenz von Mittelwerten (A minus B) mittels der Formel $SE_{Dif} = \sqrt{SE_A^2 + SE_B^2}$ berechnet. Für die Umrechnung von Streuungsmaßen normalverteilter, kontinuierlicher Variablen wurden übliche Zusammenhänge verwendet, z. B. Standardfehler=Standardabweichung/ \sqrt{n} .

Die folgenden patientenrelevanten Zielgrößen sollen berücksichtigt werden:

- Hypoglykämien
- Gewichtsveränderungen

- Unerwünschte Ereignisse
- Folgekomplikationen und Mortalität
 - Kardiovaskuläre Morbidität und Mortalität
 - Zerebrovaskuläre Morbidität und Mortalität
 - Gefäßbedingte nichtkardiale und nichtzerebrale Morbidität und Mortalität
 - Gesamtmortalität
 - Erblindung / Visusverschlechterung (Veränderungen des Augenhintergrundes)
 - Terminale Niereninsuffizienz mit Dialysenotwendigkeit
 - Amputationen
- Studienabbrüche aufgrund unzureichender glykämischer Kontrolle
- Gesundheitsbezogene Lebensqualität und Therapiezufriedenheit
- Hyperosmolare Komata
- Stationäre Behandlungen

In Bezug auf die Verbesserung des Gesundheitszustandes werden Hypoglykämien nur unter der Berücksichtigung der Blutzuckerkontrolle (HbA1c) betrachtet. Das liegt darin begründet, dass bei der blutzuckersenkenden Behandlung von Patienten mit Diabetes mellitus Typ 2 die Vermeidung von Hypoglykämien ein wichtiges Therapieziel ist (IQWiG, 2009a), dabei aber berücksichtigt werden muss, dass eine intensiviertere Therapie zur Blutzuckersenkung ein höheres Risiko für das Erleiden schwerwiegender Hypoglykämien zur Folge haben kann (UK Prospective Diabetes Study Group, 1998b; Patel et al., 2008; Ismail-Beigi et al., 2010). Somit könnte eine geringere Hypoglykämierate in einer der Behandlungsgruppen möglicherweise allein durch eine geringere Intensität der Blutzuckersenkung zu erklären sein und wäre nicht zwingend auf einen Effekt der Substanz zurückzuführen.

Aus diesem Grund werden die Ergebnisse für den Endpunkt „Hypoglykämien“ ausschließlich zusammen mit dem Surrogatparameter „Ausmaß der Blutzuckerkontrolle (HbA1c)“ beurteilt. In der nachfolgenden Tabelle 4-5 ist die für diesen Zweck herangezogene Methodik dargestellt.

Tabelle 4-5: Methodik zur gemeinsame Betrachtung der Endpunkte Blutzuckerkontrolle (HbA1c) und Hypoglykämien

Ergebnis für die Einzelendpunkte		Ergebnis der gemeinsamen Betrachtung -Bewertung des Zusatznutzens
Blutzuckerkontrolle (HbA1c)	Hypoglykämie-Rate	
Vorteil	Vorteil	Zusatznutzen
Vorteil	Kein Unterschied	Kein Zusatznutzen ²
Vorteil	Nachteil	Kein Zusatznutzen ¹
Kein Unterschied	Vorteil	Zusatznutzen
Kein Unterschied	Kein Unterschied	Kein Zusatznutzen
Kein Unterschied	Nachteil	Schaden
Nachteil	Vorteil	Kein Zusatznutzen
Nachteil	Kein Unterschied	Kein Zusatznutzen ²
Nachteil	Nachteil	Schaden

Kein Unterschied: kein Vor- bzw. Nachteil für eine der beiden Behandlungsoptionen.

Nachteil: Statistisch signifikanter und klinisch relevanter Unterschied zu Ungunsten des zu bewertenden Arzneimittels; Als klinisch relevant wird ein HbA1c Unterschied von >0,35% angesehen. Dies ist in Übereinstimmung mit den Vorgaben von der FDA (FDA, 2008), auf die sich auch das IQWiG im Abschlussbericht Blutglukoseselbstmessung bezieht (IQWiG, 2009b). Eine patientenrelevante Hypoglykämie ist eine Hypoglykämie, welche symptomatisch ist und als unerwünschtes Ereignis berichtet wird. Ein niedriger Blutzuckerwert ohne Symptome wird nicht als relevant angesehen.

Vorteil: Statistisch signifikanter und klinisch relevanter Unterschied zu Gunsten des zu bewertenden Arzneimittels. Als klinisch relevant wird in Analogie zur Definition des Nachteils ein HbA1c Unterschied von >0,35% angesehen. Im IQWiG-Abschlussbericht des Projektes „Urin- und Blutzuckerselbstmessung bei Diabetes mellitus Typ 2“ wurden beim Vertauschen der Null- und Alternativhypothese beim Test auf Nichtunterlegenheit auch andere Schwellen, z.B. eine Halbierung der vorgegebenen Nichtunterlegenheitsschwelle (>0,175%) diskutiert (IQWiG, 2009b). Eine patientenrelevante Hypoglykämie ist eine Hypoglykämie, welche symptomatisch ist und als unerwünschtes Ereignis berichtet wird. Ein niedriger Blutzuckerwert ohne Symptome wird nicht als relevant angesehen.

¹Für diese Konstellationen der Einzelendpunkte muss eine einzelfallbezogene Gesamtbeurteilung durchgeführt werden, da das Ausmaß des Vorteils für Blutzuckersenkung gegenüber dem Ausmaß des Nachteils für Hypoglykämie-Rate abgewogen werden muss.

²Auch beim Vorliegen eines Vorteils oder Nachteils beim Endpunkt HbA1c, wird der Endpunkt Blutzuckerkontrolle (HbA1c) lediglich zur Beurteilung der Hypoglykämie-Rate herangezogen; ein klinisch relevanter Unterschied von HbA1c allein rechtfertigt daher noch nicht die Gesamtbeurteilung Zusatznutzen bzw. „weniger Nutzen“, da die Relevanz dieses Endpunkts als Surrogatparameter durchaus in Frage gestellt wird.

Im Vergleich zu einem aktiven Komparator (Zusatznutzen) lässt sich der patientenrelevante Zusatznutzen vor allem durch eine Verbesserung der Therapieerträglichkeit in Bezug auf die Hypoglykämie-Rate nachweisen. Der Effekt auf den Parameter HbA1c wird daher im Rahmen der Bewertung des Zusatznutzens als nachrangig gegenüber dem Effekt auf die Hypoglykämie-Rate angesehen.

Bezüglich der Verbesserung der Lebensqualität wird die Therapiezufriedenheit anhand des *Diabetes Treatment Satisfaction Questionnaire Status* (DTSQs) berücksichtigt (Bradley C., 1994). Der DTSQ ist ein Messinstrument, das die Therapiezufriedenheit anhand von drei

Dimensionen widerspiegelt: Behandlungszufriedenheit, Hyperglykämiewahrnehmung und Hypoglykämiewahrnehmung. Der DTSQ ist ein validierter Fragebogen, welcher häufig zur Erfassung der Therapiezufriedenheit bei Diabetes mellitus Typ 2 angewendet wird und kann zur Nutzenbewertung herangezogen werden. Des Weiteren wird der DTSQ von der *World Health Organization* (WHO) und der internationalen Diabetesgesellschaft zu Messung der Therapiezufriedenheit empfohlen (Bradley C., 1994).

Zur Analyse der Verträglichkeit und Sicherheit werden die Gesamtrate unerwünschter Ereignisse, Gesamtrate schwerwiegender unerwünschter Ereignisse, Therapieabbrüche aufgrund unerwünschter Ereignisse und kardiale Ereignisse (im Rahmen der Erhebung der unerwünschten Ereignisse in der Studie) betrachtet. Da im Verlauf einer Dapagliflozin-Behandlung ein erhöhtes Auftreten von Anzeichen für Genital- bzw. Harnwegsinfektionen beobachtet wurde, werden auch die Gesamtrate der Anzeichen für Genitalinfektionen und die Gesamtrate der Anzeichen für Harnwegsinfektionen als unerwünschte Ereignisse von speziellem Interesse dargestellt.

Gewichtsveränderungen werden im Rahmen der Betrachtung der unerwünschten Ereignisse in den bisherigen Nutzenbewertungen des IQWiG im Indikationsgebiet Diabetes mellitus Typ 2 explizit dargestellt (IQWiG, 2009a). Da Diabetes häufig mit den anderen Erkrankungen des sogenannten metabolischen Syndroms, d. h. Übergewicht, Hypertonie und Fettstoffwechselstörung, assoziiert ist und viszerales Übergewicht einen gemeinsamen Risikofaktor für die Grunderkrankung (Diabetes mellitus) und kardiovaskuläre Erkrankungen darstellt, gilt ein gewichtsneutraler bzw. ein gewichtsreduzierender Effekt einer antidiabetischen Therapie in den Leitlinien als vorteilhaft und stellt somit ein patientenrelevantes Therapieziel dar. Daher wurden Gewichtsveränderungen als patientenrelevant betrachtet.

Alle in der Zusatznutzenbeurteilung betrachteten Endpunkte werden als patientenrelevant eingestuft, da sie einerseits die krankheitsbedingte Morbidität widerspiegeln und andererseits patientenrelevante Therapieziele abbilden sowie die Behandlungsverträglichkeit und Akzeptanz beschreiben und damit umfassend den Gesundheitszustand des Patienten widerspiegeln. Zu diesem Ergebnis kommt auch das IQWiG, das alle oben genannten Endpunkte als patientenrelevant einstuft (IQWiG, 2009a). Insbesondere die Vermeidung von Hypoglykämien stellt ein relevantes Therapieziel einer antihyperglykämischen Therapie des Diabetes mellitus Typ 2 dar.

Darüber hinaus wird die Änderung im systolischen und diastolischen Blutdruck als Surrogatparameter für das kardiovaskuläre Risiko berücksichtigt. Eine Begründung der Verwendung dieses Surrogatendpunktes und eine Darlegung, warum dieser Endpunkt Aussagen zu dem patientenrelevanten Endpunkt kardiovaskuläre Morbidität und Mortalität zulässt findet sich in Abschnitt 4.5.4.

Es werden für alle Endpunkte sowohl die Studienpopulation als auch die *post hoc* analysierte zulassungsrelevante Population (entspricht der Zielpopulation) dargestellt.

4.2.5.3 Meta-Analysen

Sofern die vorliegenden Studien dazu geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst werden. Dabei sollen verschiedene Studientypen nicht in einer Meta-Analyse zusammengefasst werden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten⁵ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g , als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär anhand des Odds Ratio durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁶ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität⁷ erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Meta-Analyse ist ein statistisches Verfahren um die Ergebnisse mehrerer Studien, die die gleiche Frage bearbeiten, quantitativ zu einem Ergebnis zusammenzufassen und dadurch die Aussagekraft (Genauigkeit der Effektschätzer) gegenüber Einzelstudien zu erhöhen.

⁵ DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-188.

⁶ Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

⁷ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.

Sind die Einzelstudien hinreichend ähnlich in ihrer Fragestellung, in ihren Populations- und Design-Charakteristika und die zu kombinierenden Studienresultate für die einzelnen Endpunkte hinreichend homogen (Bewertung auf der Heterogenität auf Basis des Heterogenitäts Test oder I^2 -Maß), erfolgt die quantitative Zusammenfassung der Studienresultate mittels einer Meta-Analyse.

Die Meta-Analyse wird mit geeigneter Software, wie etwa Review Manager (RevMan, Version 5), durchgeführt. Sofern berichtet, werden primär die Resultate aus ITT-Analysen kombiniert. Falls vorhanden werden adjustierte Werte gegenüber ‚rohen‘ Werten bevorzugt. Alle Meta-Analysen basieren auf einem Modell mit zufälligen Effekten (REM); Modelle mit festen Effekten (FEM) werden im Rahmen einer Sensitivitätsanalyse verwendet. Zur Interpretation der Resultate basierend auf dem REM und dem FEM werden Maße zur Bewertung und Quantifizierung von Heterogenität (siehe unten) herangezogen. Falls Effekt- oder Streuungsmaße der Einzelstudien nicht verfügbar sind, können diese, wenn möglich, durch Angaben aus den Studienberichten berechnet oder approximiert werden.

Für dichotome Endpunkte wird der Therapieeffekt in der Regel anhand des OR zusammengefasst. Gegebenenfalls, z. B. im Falle nicht hinreichend homogener Studienresultate oder im Falle hoher Ereignisraten, werden auch andere Effektmaße berechnet (wie z. B. das RR) und für Sensitivitätsanalysen verwendet (siehe Kapitel 4.2.5.4). Diese Effektmaße werden jedoch nur dann im Dossier dargestellt, wenn es relevante Änderungen/Abweichungen im Vergleich zum OR gibt. Für kontinuierliche Variablen werden Mittelwertdifferenzen berechnet.

Zur Abschätzung der klinischen Relevanz werden, falls keine validierten Irrelevanzgrenzen für den Gruppenunterschied vorliegen, Responderanalysen herangezogen. Bei Vorliegen eines validierten oder etablierten Responsekriteriums erfolgt eine Dichotomisierung im Sinne von Respondern/Non-Respondern anhand dieses Kriteriums. Danach erfolgt die Bewertung des Gruppenunterschieds anhand des OR (bzw. RR) für Response (bzw. Non-Response). Da bei Responderanalysen davon ausgegangen wird, dass die Responsedefinition eine Bewertung der klinischen Relevanz bereits auf individueller Ebene erlaubt, wird ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen auch gleichzeitig als klinisch relevant gewertet.

Die zusammenfassende Darstellung der Resultate der Einzelstudien sowie eine visuelle Einschätzung möglicher Heterogenität der Studienergebnisse erfolgt mit Hilfe eines *Forest-Plots*, in dem die Therapieeffekte der Einzelstudien (sowie – im Fall hinreichender Homogenität – ein zusammenfassender Effektschätzer) mit dem entsprechenden Konfidenzintervall abgetragen werden. Falls nur eine einzelne Studie vorhanden ist, oder die Einzelstudien klinische Heterogenität aufweisen, werden deren Ergebnisse nur tabellarisch dargestellt.

Die quantitative Untersuchung der statistischen Heterogenität erfolgt mit Hilfe eines statistischen Heterogenitätstests (Q-Statistik, p-Wert) und der Berechnung der I^2 -Statistik (Higgins et al., 2003). Um statistische Heterogenität festzustellen wird ein Signifikanzlevel

von 0,2 verwendet. Für die I^2 Statistik gelten die im Cochrane Handbuch (Higgins und Green, 2008) beschriebenen Grenzwerte für Heterogenität (>50% indiziert erhebliche Heterogenität). Ist die statistische Heterogenität nicht bedeutsam, d. h. $p > 0,2$ (Heterogenitätstest), so wird ein zusammenfassender Schätzer berechnet. Auch im Fall erheblicher Heterogenität werden zusammenfassende Schätzer berechnet (Skipka und Bender, 2010); die Interpretierbarkeit der Schätzer wird allerdings kritisch diskutiert und mögliche Ursachen der Heterogenität werden mittels Sensitivitäts- und Subgruppenanalysen untersucht.

In folgenden Fällen werden keine Meta-Analysen durchgeführt:

1. Falls nur eine für die Fragestellung relevante RCT zur Verfügung steht.
2. Falls die Studien aufgrund klinischer Heterogenität einen zusammenfassenden Schätzer nicht sinnvoll und nicht interpretierbar erscheinen lassen. Klinische Heterogenität versteht sich dabei als die Variabilität zwischen bzw. die Vielfältigkeit (*diversity*) bezüglich Studiencharakteristika, Populationscharakteristika, Begleiterkrankungen, Begleitbehandlungen, Endpunktoptimalisierungen, Beobachtungsdauer, etc.

In diesen Fällen werden die Resultate der patientenrelevanten Endpunkte, wie in der/den Einzelstudie(n) analysiert, dargestellt. Insbesondere für quantitative Parameter werden, falls notwendig, die Therapieeffekte zusätzlich hinsichtlich ihrer klinischen Relevanz bewertet.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden, wenn sich die Studien bezüglich dieser Faktoren unterscheiden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Sensitivitätsanalysen dienen zur Überprüfung der Robustheit der Ergebnisse und Schlussfolgerungen gegenüber der Variation methodischer Faktoren. Sensitivitätsanalysen können sich zum einen auf die Variation meta-analytischer Methodik beziehen (wie z. B. die Verwendung eines Modells mit festen Effekten anstelle eines Modells mit zufälligen Effekten oder die Nicht-Berücksichtigung von Studien mit geringerer Qualität). Zum anderen können sie die Robustheit der Resultate der Einzelstudie im Hinblick auf methodische Variation, wie

z. B. das Ersetzen fehlender Werte untersuchen. Dabei ist man auf die Verfügbarkeit solcher Analysen aus der Einzelstudie angewiesen.

Folgende Sensitivitätsanalysen sind im Ergebnisteil adressiert: Zur Überprüfung der Robustheit der verwendeten Methodik und zum Ersetzen fehlender Werte bei stetigen Zielvariablen wurden neben den primären Analysen (*Last observation carried forward*, LOCF Methode) auch *Repeated-Measurements* Analysen dargestellt.

Im Rahmen der MTC wurde der Effekt der Berücksichtigung bzw. Nicht-Berücksichtigung qualitativ schwächerer Studien beurteilt.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollten, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. -stadium

Sollten sich aus den verfügbaren Informationen Anhaltspunkte für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen.

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach

Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, d. h. klinischer Faktoren, die die Effekte beeinflussen, untersucht. Möglicherweise variierende Effekte zwischen verschiedenen Subgruppen werden mittels Interaktionstests bewertet. Diese werden entweder aus vorliegenden Ergebnissen aus Regressionsanalysen, die Interaktionsterme beinhalten oder aus den Ergebnissen von Subgruppenanalysen berechnet.

Ein p-Wert $<0,2$ liefert einen Hinweis auf unterschiedliche Effekte zwischen den Subgruppen. Liegt der p-Wert unter $0,05$, so kann von unterschiedlichen Effekten zwischen den Subgruppen ausgegangen werden (Beleg für eine Interaktion). Die Ergebnisse von Subgruppenanalysen werden stets hinsichtlich ihrer Validität (*Power*, Präspezifizierung, qualitativ versus quantitativ) sowie hinsichtlich möglicher Auswirkungen auf das aus der Gesamtpopulation abgeleitete Fazit diskutiert. Falls kein Hinweis auf eine signifikante Interaktion zwischen Behandlung und dem subgruppenbildenden Faktor besteht, wird keine nach Subgruppen getrennte Meta-Analyse durchgeführt. Meta-Regressionen werden bei geringer Studienzahl (<5) nicht durchgeführt.

Für folgende Subgruppen werden die Ergebnisse für die Subgruppenkategorien separat dargestellt:

- Alter (<65 Jahre und ≥ 65 bis <75 Jahre)
- Geschlecht (männlich; weiblich)

Für die im Studienprotokoll geplanten Subgruppenanalysen werden Interaktionstests berichtet. Diese beziehen sich ausschließlich auf den Endpunkt HbA1c und auf zusammenfassende sicherheitsrelevante Endpunkte.

Für alle Subgruppenanalysen werden ausschließlich solche Trennpunkte verwendet, die im Studienprotokoll präspezifiziert waren. Der Trennpunkt für die Subgruppenkategorien bzgl. Alter der Patienten teilt die Patienten in Alterskategorien ein, die den gebräuchlichen Alterseinteilungen in der Indikation entspricht.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich⁸. Komplexe Verfahren für den simultanen Vergleich von

⁸ Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-148.

mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“⁹, „Multiple-Treatment-Meta-Analysen“¹⁰ oder auch „Netzwerk-Meta-Analysen“¹¹, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹². Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹³.

Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Um auch den Zusatznutzen von Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitoren aufzuzeigen, werden in Abschnitt 4.3.2.1 die Ergebnisse einer MTC dargestellt.

⁹ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

¹⁰ Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

¹¹ Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17(3): 279-301.

¹² B. Schöttker, D. Lühmann, D. Boukhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

¹³ Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

Bei einer MTC können zwei oder mehr Behandlungsoptionen simultan miteinander verglichen und in eine Reihenfolge (*Ranking*) gebracht werden. Dabei wird die Evidenz aus direkten (*Head-to-head-Studien*) und indirekten (Studien mit unterschiedlichen Komparatoren) Vergleichen miteinander kombiniert.

Die hier in dem Dossier dargestellte MTC wurde von der Oxford Group durchgeführt. Es wurden zwei Netzwerke beschrieben, die sich hinsichtlich ihrer Studiendauer unterscheiden (Netzwerk A: 24±6 Wochen und Netzwerk B: 52 Wochen). Es wurde ein Vergleich von Metformin+Dapagliflozin mit Glitazonen, GLP-1-Analoga, DPP-4-Inhibitoren, Sulfonylharnstoffen oder Placebo jeweils als *Add-on* zu Metformin durchgeführt. Im Rahmen der Ergebnisdarstellung werden zunächst die Ergebnisse der direkten Vergleiche dargestellt. Falls die Studien hinreichend homogen waren, wurde die MTC mit der Software WinBugs mithilfe der Bayesianischen Statistik auf Basis von Modellen mit zufälligen Effekten berechnet.

Da Bayesianische Modelle verwendet wurden, müssen noch Annahmen bezüglich A-priori-Verteilungen vorgenommen werden, sowie für die MCMC-Simulationen Anzahl, Startwerte und Länge der gewählten Markov-Ketten. Im Folgenden werden nicht informative A-priori-Verteilungen (Unwissenheit) verwendet, da diese keinen Einfluss auf die A-posteriori-Verteilung hat.

Für dichotome und stetige Parameter (hier: Effektschätzer für eine Behandlung relativ zur Referenzbehandlung) ist die nicht-informative A-priori-Verteilung eine Normalverteilung mit Mittelwert 0 und Varianz 10.000 (Cooper et al., 2009; Dias et al., 2010). Auch für die *nuisance* Parameter (hier: studienspezifischer Mittelwert / *log-Odds* für die Referenzbehandlung) wird diese A-priori Verteilung angenommen. Dadurch wird gewährleistet, dass das beobachtete Ergebnis nicht von der A-priori-Verteilung beeinflusst ist (Ergebnisse der Bayesianischen Analyse und der frequentistischen Analyse sind bei Wahl nicht-informativer A-prioris nahezu identisch). Bei Random-Effekt-Modellen werden noch die Varianzen der beobachteten Effekte zwischen den Studien benötigt, für diese Varianz wird ebenfalls eine A-priori-Verteilung benötigt. Hier wird eine Gleichverteilung angenommen, deren Bereich für die stetigen Parameter klinisch plausibel ist. Um die A-Posterior-Verteilung zu schätzen, werden die MCMC-Simulationen verwendet. Es werden 2 Markov-Ketten mit unterschiedlichen Startwerten jedes zu schätzenden Parameters initiiert. Jede Kette hat 20.000 *burn-in* Iterationen und mindestens 100.000 Iterationen. Jeder zu schätzende Parameter soll einen *Monte Carlo Error* von weniger als 5% des Stichprobenstandardfehlers haben.

Im Rahmen der Bewertung des Zusatznutzens wurde nur der Vergleich Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitoren betrachtet, da Metformin+DPP-4-Inhibitoren als zweckmäßig in der Population angesehen wird, welche eine Kontraindikation bzw. Unverträglichkeit gegen Metformin+Sulfonylharnstoff hat.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus

randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2)
- Weitere Untersuchungen (Abschnitt 4.3.2.3)

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Benennen Sie in der nachfolgenden Tabelle alle Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war.

Von den randomisierten kontrollierten Studien, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden, sollen die Studien, deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle randomisierten kontrollierten Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.

Benennen Sie nur randomisierte, kontrollierte Studien, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-6: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
D1690C00004	ja	abgeschlossen	Behandlungsphase: 52 Wochen	Metformin (1500–2500 mg/Tag)+ Dapagliflozin (2,5-10 mg/Tag) Metformin (1500–2500 mg/Tag)+Glipizid (5 mg-20 mg/Tag)
D1690C00004 (LT I)	nein	abgeschlossen	Verlängerungsphase I: bis 104 Wochen	Metformin (1500–2500 mg/Tag)+ Dapagliflozin (2,5-10 mg/Tag) Metformin (1500–2500 mg/Tag)+Glipizid (5 mg-20 mg/Tag)
D1690C00004 (LT II)	nein	laufend	Verlängerungsphase II: bis 208 Wochen	Metformin (1500–2500 mg/Tag)+ Dapagliflozin (2,5-10 mg/Tag) Metformin (1500–2500 mg/Tag)+Glipizid (5 mg-20 mg/Tag)
MB102014	ja	abgeschlossen	Behandlungsphase: 24 Wochen Verlängerungsphase: bis 104 Wochen	Metformin (≥ 1500 mg/Tag)+ Dapagliflozin (2,5, 5 und 10 mg/Tag) Placebo+Metformin
D1690C00012	ja	abgeschlossen	Behandlungsphase: 24 Wochen Verlängerungsphase: bis 104 Wochen	Metformin (≥ 1500 mg/Tag)+ Dapagliflozin (10 mg) Placebo+Metformin
MB102003	nein	abgeschlossen	2 Wochen	Dapagliflozin (5, 25 und 100 mg/Tag)+/- Metformin Placebo+/-Metformin
MB102020	nein	abgebrochen	4 Wochen	Dapagliflozin (5 mg) +Vorbehandlung Dapagliflozin (20 mg) +Vorbehandlung

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
MB102045	nein	abgeschlossen	12 Wochen	Dapagliflozin (5 mg/Tag)+Vorbehandlung Placebo+Vorbehandlung
MB102035	nein	abgeschlossen	12 Wochen	Dapagliflozin (10 mg/Tag)+Vorbehandlung (Metformin und/oder Sulfonylharnstoffe) Placebo+Vorbehandlung (Metformin und/oder Sulfonylharnstoffe)
MB102054	nein	abgeschlossen*	24 Wochen	Dapagliflozin (5 und 10 mg/Tag)+Vorbehandlung Placebo+Vorbehandlung
MB102055	nein	laufend	24 Wochen	Dapagliflozin (5 und 10 mg/Tag)+Metformin (1500-3000 mg/Tag)+Vorbehandlung Placebo+Metformin+Vorbehandlung
MB102073	nein	laufend	12 Wochen	Dapagliflozin (10 mg/Tag)+Vorbehandlung Placebo+Vorbehandlung
MB102077	nein	laufend	12 Wochen	Dapagliflozin (10 mg/Tag)+Vorbehandlung Placebo+Vorbehandlung
D1691C00003	nein	abgeschlossen	16 Wochen	Metformin (≥ 1500 mg/Tag)+ Dapagliflozin (2,5, 5 und 10 mg/Tag) Placebo+Metformin
D1690C00018	nein	abgeschlossen	Behandlungsphase: 24 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00018 (LT I)	nein	abgeschlossen	Verlängerungsphase I: bis 52 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00018 (LT II)	nein	laufend	Verlängerungsphase II: bis 104 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
D1690C00019	nein	abgeschlossen	Behandlungsphase: 24 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00019 (LT I)	nein	abgeschlossen	Verlängerungsphase I: bis 52 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00019 (LT II)	nein	laufend	Verlängerungsphase II: bis 104 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00007	nein	Abgebrochen vor Einschluss des ersten Patienten	24 Wochen	Dapagliflozin (10 mg/Tag)+Metformin Sitagliptin (100 mg/Tag) +Metformin
CV181169	nein	laufend	24 Wochen	Dapagliflozin (10 mg)+Metformin XR (1500-2000 mg) +Saxagliptin (5 mg) Dapagliflozin (10 mg)+Metformin XR (1500-2000 mg) +Placebo Metformin XR (1500-2000 mg) +Saxagliptin (5 mg) +Placebo
CV181168	nein	laufend	52 Wochen	Dapagliflozin (10 mg)+Metformin (≥ 1500 mg) +Saxagliptin (5 mg) Dapagliflozin (10 mg)+Metformin (≥ 1500 mg) +Placebo

LT: Long term; XR: Extended release

* Es liegt noch kein Studienbericht vor.

Geben Sie an, welchen Stand die Information in Tabelle 4-6 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus sollte nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Der Informationsstand für den Status der angegebenen Studien bezieht sich auf den 29.10.2012.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-6 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-7: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
MB102014	Die Studie ist placebokontrolliert und enthält die falsche bzw. keine Vergleichstherapie
D1690C00012	Die Studie ist placebokontrolliert und enthält die falsche bzw. keine Vergleichstherapie
D1690C00004 (LT II)	Die Verlängerungsphase II (bis 208 Wochen) läuft noch. Der Studienbericht wird für Mitte 2013 erwartet.
MB102003	Studiendauer zu gering.
MB102020	Phase-1-Studie. Die Studiendauer beträgt nur 4 Wochen.
MB102045	Studiendauer zu gering.
MB102035	Studiendauer zu gering.
MB102054	Die Studie wurde kürzlich abgeschlossen. Es liegen noch keine Ergebnisse vor. Die Studie wurde daher in Anhang 4-E aufgenommen.
MB102055	Die Studie läuft noch. Es liegen noch keine Ergebnisse vor.
MB102073	Studiendauer zu gering. Studie läuft noch.
MB102077	Studiendauer zu gering. Studie läuft noch.
D1691C00003	Studiendauer zu gering und falsche Vergleichstherapie.
D1690C00018 (inkl. LT I und II)	Die Studie ist placebokontrolliert und enthält die falsche bzw. keine Vergleichstherapie
D1690C00019 (inkl. LT I und II)	Die Studie ist placebokontrolliert und enthält die falsche bzw. keine Vergleichstherapie
D1690C00007	Die Studie enthält die falsche Vergleichstherapie (Sitagliptin). Außerdem wurde die Studie vor Einschluss des ersten Patienten abgebrochen.
CV181169	Die Studie enthält die falsche Vergleichstherapie.
CV181168	Die Studie läuft noch. Es liegen keine Ergebnisse vor.

Stellen Sie die abgebrochenen und laufenden Studien, die nicht in die Bewertung eingehen, in Anhang 4-D und Anhang 4-E dar.

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Sofern eine bibliografische Literaturrecherche durchgeführt wurde, beschreiben Sie nachfolgend das Ergebnis dieser Recherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und

wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-C.

Falls die Recherche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel gemeinsam mit der Recherche zu einer anderen Fragestellung (z. B. zu indirekten Vergleichen) durchgeführt wurde, unterteilen Sie die Angaben zu relevanten Treffern und Studien entsprechend diesen Fragestellungen.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.

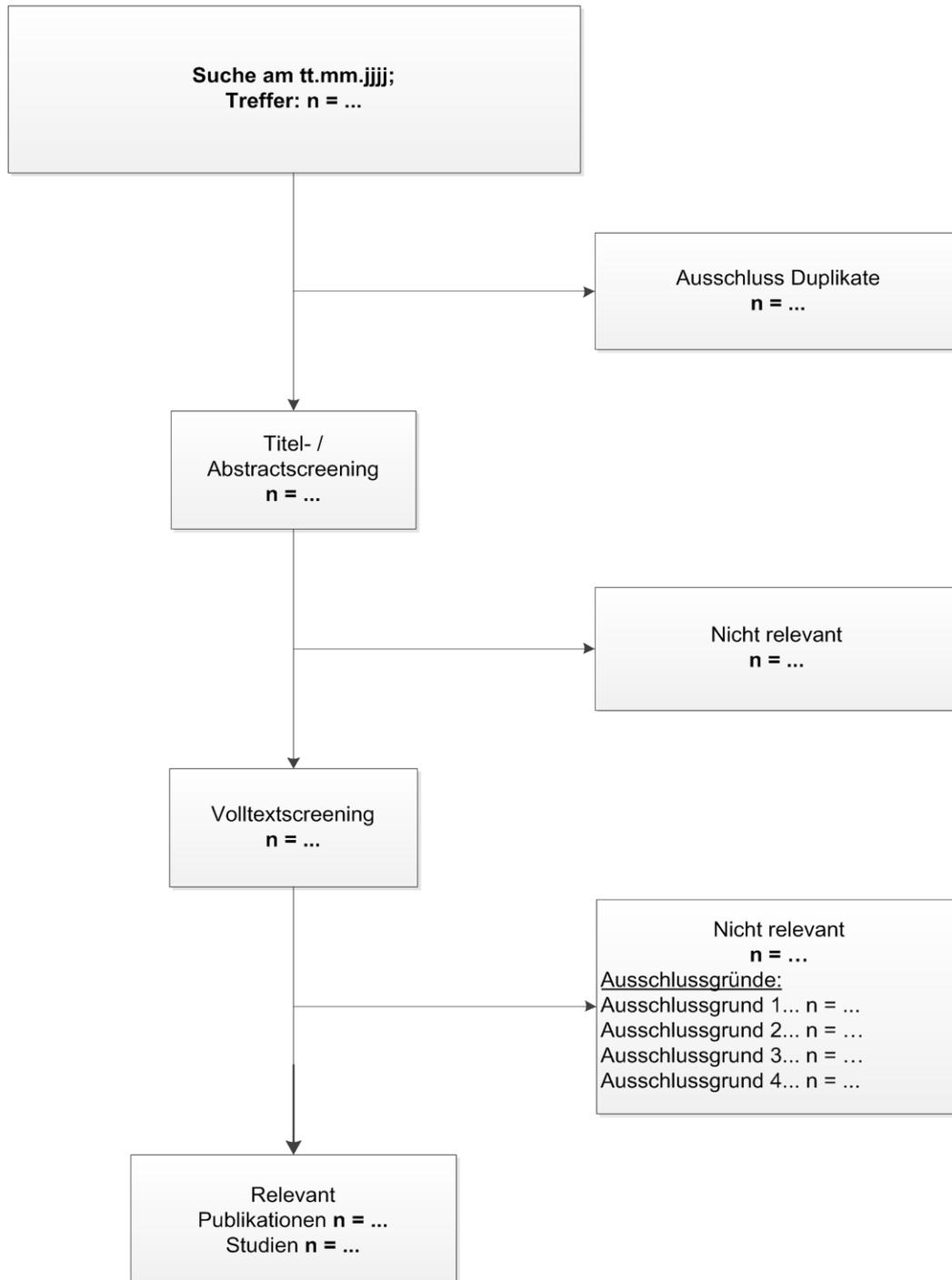


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Wie bereits in Abschnitt 4.2.3.2 beschrieben, wurde keine bibliografische Literaturrecherche durchgeführt.

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-6) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde (sofern eine solche durchgeführt wurde). Fügen Sie für jede Studie eine neue Zeile ein.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-8: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. ^b)
D1690C00004	Clinicaltrials.gov [NCT00660907] (Clinicaltrials.gov, 2012a) WHO ICTRP [EUCTR2007-005220-33-GB] (ICTRP Search Portal, 2012a)	Ja	n.d.
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.			

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle für die Nutzenbewertung berücksichtigten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige

Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-9: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
--						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
D1690C 00004	Ja	Ja	nein	Ja (AstraZeneca, 2010b) (AstraZeneca, 2011b)	Ja (Clinicaltrials.gov, 2012a) (ICTRP Search Portal, 2012a)	Ja (Nauck et al., 2011)
a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.						
b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.						
c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-F zu hinterlegen.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
D1690C00004	RCT, Phase III, doppelblind, parallel, aktivkontrolliert, multizentrisch (95 Zentren in 10 Ländern, davon 16 Zentren in Deutschland)	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit Metformin alleine nicht ausreichend kontrolliert werden kann (HbA1c >6,5% und ≤10%)	Metformin+Dapa- gliflozin (n= 406) Metformin+ Glipizid (n= 408)	<i>Enrollment-Phase:</i> 1 Woche <i>Dose Stabilization</i> Phase: 8 Wochen <i>Placebo-Lead-in</i> Phase: 2 Wochen Behandlung: 52 Wochen Verlängerungsphase I: bis 104 Wochen Verlängerungsphase II: bis 208 Wochen Follow-up-Phase:3 Wochen nach der letzten Studienmedikation	International 3/2008 – 12/2010 Weiterbeobach- tung bis 12/2012 läuft noch	Primärer Endpunkt: HbA1c-Änderung in Woche 52 im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: HbA1c-Änderung in Woche 104 im Vergleich zum Ausgangswert Gewichtveränderung in Woche 52 / 104 im Vergleich zum Ausgangswert Anteil der Patienten, deren Gewichtsabnahme mehr als 5% in Woche 52 / 104 im Vergleich zum Ausgangswert betrug Anzahl Hypoglykämien (innerhalb 52 / 104 Wochen) Kardiale Ereignisse (innerhalb 52 / 104 Wochen) Unerwünschte Ereignisse (innerhalb 52 / 104 Wochen)

Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Metformin + Dapagliflozin	Metformin + Glipizid	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
D1690C00004	Metformin 1500, 2000, oder 2500 mg/Tag (gleichbleibend) + Dapagliflozin	Metformin 1500, 2000, oder 2500 mg/Tag (gleichbleibend) + Glipizid	<i>Dose Stabilization</i> Phase: 8 Wochen: stabile Dosis von Metformin <i>Placebo-Lead-in</i> Phase: Einfachblinde Placebo-Gabe, offene Metformin-Gabe (1500, 2000, 2500 mg pro Tag) Behandlung 52 Wochen doppelblind: - davon waren die ersten 18 Wochen eine Titrationsphase: Patienten wurden titriert bis sie einen optimalen Effekt (FPG < 100 mg/dL, <6,1 mmol/L) oder die höchste verträgliche Dosis erreicht hatten. Patienten haben mit Dosis Level 1 (Dapagliflozin 2,5 mg / Glipizid 5 mg) begonnen und konnten dann in 3- Wochen-Intervallen zu Dosis Level 2 (Dapagliflozin 5 mg / Glipizid 10 mg) und 3 (Dapagliflozin 10 mg / Glipizid 20 mg) hochtitriert werden - die letzten 34 Wochen bestanden aus einer Erhaltungssphase (die Dosis, die die Patienten am Ende der Titrationsphase erreichten, wurde beibehalten). Verlängerungsphase I: bis 104 Wochen Verlängerungsphase II: bis 208 Wochen (laufend) Follow-up-Phase: 3 Wochen nach der letzten Studienmedikation

Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD)	Geschlecht w / m (%)	BMI (kg/m²) Mittelwert (SD)	Gewicht (kg) Mittelwert (SD)
D1690C00004					
Metformin+Dapagliflozin	400	58,1 (9,37)	44,8 / 55,3	31,71 (5,104)	88,44 (16,323)
Metformin+Glipizid	401	58,6 (9,80)	45,1 / 54,9	31,23 (5,053)	87,60 (16,970)

BMI: *Body-Mass-Index*, SD: Standardabweichung

Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Baseline HbA1c (%)	Diabetesdauer (Jahre)	Anzahl der Studienabbrecher n (%)
		Mittelwert (SD)	Mittelwert (SD)	
D1690C00004				Nach 52 Wochen:
Metformin+Dapagliflozin	400	7,69 (0,855)	6,08 (4,611)	84 (20,7)
Metformin+Glipizid	401	7,74 (0,886)	6,55 (5,902)	94 (23,0)
				Nach 104 Wochen*:
				202 (49,8)
				220 (53,9)

*die häufigste Ursache für den Abbruch der Studie war, dass die Patienten nicht mehr den studienspezifischen Einschlusskriterien entsprachen und gemäß Protokoll abbrechen mussten. Der häufigste Ausschlussgrund war, dass Patienten aufgrund der prädefinierten glykämischen Werte ausgeschlossen werden mussten, da es zu diesem Zeitpunkt keine Notfallmedikation gab.

SD: Standardabweichung

Beschreiben Sie die Studien zusammenfassend. Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

Metformin+Dapagliflozin versus Metformin+Sulfonylharnstoff:

Studie D1690C00004 ist eine randomisierte, doppelblinde, aktivkontrollierte, 52-wöchige Phase III Studie mit einer Verlängerungsphase I von 52 Wochen und einer Verlängerungsphase II von 104 Wochen. In zwei Gruppen wurde die Wirksamkeit und Sicherheit von Metformin+Dapagliflozin und Metformin+Glipizid verglichen. Eingeschlossen wurden Typ-2-Diabetiker über 18 Jahre, bei denen eine Therapie, die Metformin enthielt, zur Blutzuckerkontrolle nicht mehr ausreichend war. Mögliche andere orale Antidiabetika (OAD) wurden vor der Dosis-Stabilisierungsperiode abgesetzt. Die Patienten erhielten während der Behandlungsphase eine gleichbleibende Metformin-Dosis von 1500-2500 mg/Tag und die doppelblinde Studienmedikation (Dapagliflozin oder Glipizid). Die 52-wöchige Behandlungsphase bestand aus einer 18-wöchigen Titrationsphase und einer anschließenden Erhaltungsphase. Während der Titrationsphase wurde in jeweils dreiwöchigem Abstand eine Dosisanpassung (bei Dapagliflozin: 2,5, 5 und 10 mg und bei Glipizid: 5, 10 und 20 mg) bis zu einem optimalen Effekt (definiert als FPG <110 mg/dL) oder der höchsten verträglichen Dosis durchgeführt. Die demografischen Charakteristika zwischen den beiden Gruppen unterschieden sich nicht. Das durchschnittliche Alter lag bei 58,1 / 58,6 Jahren, das mittlere Gewicht bei 88,44 / 87,60 kg und der HbA1c-Ausgangswert bei 7,69 / 7,74%. Der Anteil der Frauen (44,8 und 45,1%) sowie die Krankheitsdauer (6,08 und 6,55 Jahre) waren in beiden Behandlungsgruppen ebenfalls vergleichbar. Der Anteil der Studienabbrecher nach 52 Wochen lag in der Metformin+Dapagliflozin-Gruppe bei 20,7% und in der Metformin+Glipizid-Gruppe bei 23,0%. Nach 104 Wochen war die Anzahl der Studienabbrecher dann 49,8% in der Metformin+Dapagliflozin-Gruppe bzw. 53,9% in der Metformin+Sulfonylharnstoff-Gruppe.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-14: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
D1690C00004	ja	ja	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

D1690C00004: Es handelt sich hierbei um eine doppelblinde, randomisierte aktivkontrollierte Studie. Die Randomisierungssequenz wurde adäquat durchgeführt. Die Verdeckung der Gruppenzuteilung (mit Zuteilungsverhältnis 1:1) erfolgte zentral und unabhängig durch ein *interactive web response system* (IWRS). Die Patienten und Behandler waren verblindet. Die Erhebung und Auswertung der Ergebnisse zu den einzelnen definierten Endpunkten erfolgte planmäßig und wie vorab definiert. Auch nach den ersten Auswertungen der Studienergebnisse, verblieben die Ärzte sowie die Patienten weiterhin verblindet. Es ergeben sich keine Anhaltspunkte für ein Verzerrungspotenzial der Studie.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-15: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Hypo- glykämien	Gewichts- veränderung	Un- erwünschte Ereignisse	Folge- komplika- tionen	Gesamt- mortalität	Systolischer und diastolischer Blutdruck
D1690C00004 (inkl. LT I)	ja	ja	ja	ja*	nein**	ja

*Kardiale Ereignisse werden zur Beurteilung der patientenrelevanten Endpunkte „kardiovaskuläre Morbidität und Mortalität“ verwendet. ** Mortalität war nicht als Endpunkt vorgesehen und wurde daher nicht adäquat im Rahmen von einer Überlebenszeitanalyse untersucht. Insgesamt gab es keinen Todesfall in der Metformin+Dapagliflozin-Gruppe und 4 Todesfälle in der Metformin+Sulfonylharnstoff-Gruppe nach 104 Wochen. Die Auswertung der Anzahl der Todesfälle erfolgte im Rahmen der schwerwiegenden unerwünschten Ereignisse.

Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Hyperosmolare bzw. ketoazidotische Komata	Stationäre Behand- lungen	Gesundheits- bezogene Lebensqualität	Therapie- zufriedenheit	Ausmaß der Blutzucker- senkung (HbA1c)
D1690C00004 (inkl. LT I)	nein	nein	nein	ja	ja*

* Inklusive Studienabbrüche aufgrund unzureichender glykämischer Kontrolle

Nicht berichtete Zielgrößen

Zu einigen der vorab definierten Zielgrößen liegen keine relevanten Studien vor. Dies gilt für die folgenden Endpunkte: Erblindung/Visusverschlechterung, terminale Niereninsuffizienz, Amputation, zerebrovaskuläre Morbidität und Mortalität, gefäßbedingte nichtkardiale und nichtzerebrale Morbidität und Mortalität, stationäre Behandlungen, hyperosmolare Komata und gesundheitsbezogene Lebensqualität.

Verwendete Daten

Alle im folgenden beschriebenen Ergebnisse werden jeweils für die **Studienpopulation** (Patienten mit manifestem Diabetes mellitus Typ 2, die unter einer Metformin-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker aufweisen) sowie für die **Zielpopulation**, für die Dapagliflozin zugelassen ist (Ausschluss von Patienten mit einer eGFR <60 mL/min/1,73 m² bzw. Kreatininclearance <60 mL/min, Patienten, die ≥75 Jahre alt sind bzw. mit Schleifendiuretika vorbehandelt waren), tabellarisch dargestellt. Die Daten für die Zielpopulation wurden *post hoc* analysiert (Bristol-Myers Squibb und AstraZeneca GmbH, 2012b)

Die Beschreibung der Ergebnisse bezieht sich im Folgenden auf die zugelassene Zielpopulation. Die Ergebnisse für die Studienpopulation werden nur dann betrachtet, wenn sie von den Ergebnissen für die Zielpopulation wesentlich abweichen.

4.3.1.3.1 Hypoglykämien unter Berücksichtigung des HbA1c-Wertes – RCT

Wie in Abschnitt 4.2.5.2 erläutert, muss für die Beurteilung der Hypoglykämierate das Ausmaß der Blutzuckerkontrolle (gemessen am HbA1c-Wert) bekannt sein und kann nur gemeinsam betrachtet werden.

In den folgenden Abschnitten 4.3.1.3.1.1 und 4.3.1.3.1.2 wurden daher zunächst das Ausmaß der Blutzuckerkontrolle (gemessen am HbA1c) und dann das Auftreten von Hypoglykämien getrennt voneinander operationalisiert und beschrieben. Im Anschluss wurden diese Endpunkte gemeinsam betrachtet (Abschnitt 4.3.1.3.1.3). Die Methodik anhand welcher die gemeinsame Betrachtung vorgenommen wurde kann Tabelle 4-5 entnommen werden.

4.3.1.3.1.1 Ausmaß der Blutzuckerkontrolle (HbA1c) – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-17: Operationalisierung von „Ausmaß der Blutzuckerkontrolle (HbA1c)“

Studie	Operationalisierung
D1690C00004	Anhand der absoluten Änderung des HbA1c-Wertes (in %) von Studienbeginn bis Woche 52 (falls der Wert nicht erhoben wurde, wurde der letztmögliche vor Woche 52 (LOCF) betrachtet) bzw. bis Woche 104 (<i>Repeated-Measurements</i> Modell). Der HbA1c-Wert wurde für alle Studienteilnehmer im Zentrallabor bestimmt. Anhand von Studienabbrüchen aufgrund unzureichender glykämischer Kontrolle

LOCF: *Last observation carried forward*

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-18: Bewertung des Verzerrungspotenzials für „Ausmaß der Blutzuckerkontrolle (HbA1c)“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D1690C00004	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Ausmaß der Blutzuckerkontrolle (gemessen anhand des HbA1c) war der primäre Endpunkt der Studie D1690C00004 und wurde anhand der Differenz der HbA1c-Werte zu Woche 52 relativ zum Wert zu Studienbeginn beurteilt. Der Vergleich der Behandlungsgruppen erfolgte auf der Basis der Differenz der adjustierten Mittelwerte. Die Adjustierung erfolgte bezüglich des Ausgangswertes. Als weitere Operationalisierung des Ausmaßes der Blutzuckerkontrolle wurden Studienabbrüche aufgrund unzureichender glykämischer Kontrolle in beiden Behandlungsgruppen dargestellt.

Die Endpunkterheber waren zu jedem Zeitpunkt der Erhebung verblindet. In beiden Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn vergleichbare Ausfallraten. Das ITT-Prinzip wurde adäquat umgesetzt, indem die randomisierten Patienten, die mindestens eine Studienmedikationseinnahme und einen Folgewert hatten, analysiert wurden (*Full Analysis Set*). Fehlende Werte wurden durch die LOCF-Methode ersetzt. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Somit kann das Verzerrungspotenzial für den Endpunkt „Ausmaß der Blutzuckerkontrolle“ unabhängig von der verwendeten Operationalisierung mit niedrig bewertet werden. Um die Langzeitwirkung

und auch die Robustheit der Ergebnisse zu untersuchen, wurde die Änderung des HbA1c-Wertes zusätzlich mit einem *Repeated-Measurements* Modell untersucht.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-19: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppenunterschied Mittelwert % (SE) [95%-KI]		
	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung# Mittelwert % (SE)	N	Ausgangswert Mittelwert % (SD)		Endwert Mittelwert % (SD)	Adjustierte Veränderung# Mittelwert % (SE)
	Metformin+Dapagliflozin			Metformin+Sulfonylharnstoff					
52 Wochen*									
	400	7,69 (0,855)	7,19 (0,760)	-0,52 (0,0403)	401	7,74 (0,886)	7,21 (1,090)	-0,52 (0,0402)	0,00 (0,0569)**
	Metformin+Dapagliflozin			Metformin+Sulfonylharnstoff					
52 Wochen°									
	321		7,14 (0,742)	-0,52 (0,0430)	315		7,13 (0,986)	-0,50 (0,0433)	-0,02 (0,0610) [-0,14;0,10]
104 Wochen°									
	233		7,05 (0,681)	-0,32 (0,0536)	208		7,06 (0,767)	-0,14 (0,0551)	-0,18 (0,0764) [-0,33;-0,03]

basierend auf einer Kovarianzanalyse mit Behandlung als Faktor und dem Ausgangswert als Kovariate

*LOCF

**KI nicht angegeben, nur Nichtunterlegenheitstest p-Wert <0,0001

° Repeated-Measurements Modell

SD: Standardabweichung; SE: Standardfehler; KI: Konfidenzintervall

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-20: Ergebnisse für „Ausmaß der Blutzuckerkontrolle (HbA1c)“ aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppenunterschied Mittelwert % (SE) [95%-KI]		
	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert % (SE)	N	Ausgangswert Mittelwert % (SD)		Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert % (SE)
	Metformin+Dapagliflozin			Metformin+Sulfonylharnstoff					
52 Wochen*									
	318	7,77 (0,870)	7,21 (0,778)	-0,55 (0,0458)	354	7,72 (0,877)	7,22 (1,066)	-0,51 (0,0434)	-0,03 (0,0632) [-0,16;0,09]
	Metformin+Dapagliflozin			Metformin+Sulfonylharnstoff					
52 Wochen°									
	272		7,16 (0,746)	-0,54 (0,0472)	284		7,14 (0,949)	-0,49 (0,0458)	-0,05 (0,0657) [-0,18;0,08]
104 Wochen°									
	196		7,06 (0,576)	-0,33 (0,0564)	192		7,08 (0,777)	-0,14 (0,0556)	-0,19 (0,0788) [-0,34;-0,03]

*LOCF

° Repeated-Measurements Modell

SD: Standardabweichung; SE: Standardfehler; KI: Konfidenzintervall

Für die Zielpopulation wurden folgende Resultate bezüglich der Blutzuckerkontrolle beobachtet: Nach 52 Wochen ergab sich kein statistisch signifikanter Unterschied zwischen der Metformin+Dapagliflozin Gruppe und der Metformin+Sulfonylharnstoff-Gruppe. Ausgehend von einem mittleren HbA1c vor Behandlung von 7,77% bzw. 7,72% in der Metformin+Dapagliflozin bzw. der Metformin+Sulfonylharnstoff-Gruppe, wurde der HbA1c in der Metformin+Dapagliflozin Gruppe um durchschnittlich 0,55% gesenkt, in der Metformin+Sulfonylharnstoff Gruppe betrug die durchschnittliche HbA1c-Reduktion 0,51%. Die a priori definierte Unterlegenheitsschranke von 0,35% wurde nicht unterschritten, d. h. Metformin+Dapagliflozin war in der Blutzuckerkontrolle der Behandlung mit Metformin+Sulfonylharnstoff nicht unterlegen. Das galt sowohl in der Studienpopulation als auch in der Zielpopulation. Die Ergebnisse auf der Basis eines *Repeated-Measurements*-Modells zeigten konsistente Ergebnisse. Die HbA1c-Senkung unter Metformin+Dapagliflozin blieb über die Zeit erhalten, während die glykämische Kontrolle unter Metformin+Sulfonylharnstoff nachließ, sodass nach 104 Wochen die HbA1c-Senkung bei der Metformin+Dapagliflozin-Gruppe größer war (*Repeated-Measurements*-Modell).

Als unterstützende Wirksamkeitsmessung wurden Studienabbrüche aufgrund unzureichender glykämischer Kontrolle erhoben. Eine unzureichende glykämische Kontrolle lag in der Studie vor, wenn folgende Messwerte zu bestimmten Visiten festgestellt wurden: Nüchternblutzucker (FPG) >270 mg/dL Visite 5-8; FPG >240 mg/dL Visite 9-11; FPG >220 mg/dL Visite 12-13; FPG >200 mg/dL oder HbA1c \geq 8.0% an Visite 15-18.

Tabelle 4-21: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	N	Studienabbrüche aufgrund unzureichender glykämischer Kontrolle, n (%; SE)
D1690C00004		
Nach 52 Wochen		
Metformin+Dapagliflozin	400	1 (0,2;0,250),
Metformin+Sulfonylharnstoff	401	15 (3,6;0,935)
OR [95%-KI]		0,06 [0,01;0,49]
Nach 104 Wochen		
Metformin+Dapagliflozin	400	56 (14,5;1,675)
Metformin+Sulfonylharnstoff	401	89 (21,6;1,973)
OR [95%-KI]		0,57 [0,40;0,82]

SE: Standardfehler, OR: Odds Ratio, KI: Konfidenzintervall

Tabelle 4-22: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Studienabbrüche aufgrund unzureichender glykämischer Kontrolle, n (%), SE)
D1690C00004		
Nach 52 Wochen		
Metformin+Dapagliflozin	318	1 (0,3;0,314),
Metformin+Sulfonylharnstoff	354	14 (4,0;1,028)
OR [95%-KI]		0,08 [0,01;0,59]
Nach 104 Wochen		
Metformin+Dapagliflozin	318	51 (15,9;1,977)
Metformin+Sulfonylharnstoff	354	80 (22,8;2,131)
OR [95%-KI]		0,65 [0,44;0,97]

SE: Standardfehler, OR: Odds Ratio, KI: Konfidenzintervall

Innerhalb der Beobachtungsdauer von 52 Wochen brach in der Metformin+Dapagliflozin-Gruppe ein Patient aufgrund unzureichender glykämischer Kontrolle die Studie ab, dagegen brachen in der Metformin+Sulfonylharnstoff-Gruppe 14 Patienten die Studie ab. Auch nach 104 Wochen brachen in der Metformin+Dapagliflozin-Gruppe statistisch signifikant weniger Patienten die Studie wegen unzureichender Blutzuckerkontrolle ab (15,9%; adjustiert bzgl. des HbA1c-Ausgangswertes) als in der Metformin+Sulfonylharnstoff-Gruppe (22,8%; adjustiert bzgl. des HbA1c-Ausgangswertes).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 2: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Da nur eine RCT vorliegt, wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.2 Hypoglykämien – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die

Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-23: Operationalisierung von Hypoglykämien

Studie	Operationalisierung
D1690 C00004	<p>Anzahl von Patienten mit mindestens einer Hypoglykämie innerhalb der 52 / 104 Wochen (mittels <i>diary review</i>).</p> <p>Hypoglykämien wurden unterteilt in:</p> <p>Schwere Hypoglykämien: Anzahl der Patienten mit mindestens einer symptomatischen Hypoglykämie, bei der externe Hilfe von Nöten war oder Blutzuckerwert <54 mg/dL</p> <p>Leichte Hypoglykämie: -Anzahl der Patienten mit mindestens einer symptomatischen Hypoglykämie mit einem Blutzuckerwert von <63 mg/dL, (unabhängig davon ob externe Hilfe von Nöten war) oder -Anzahl der Patienten mit einer asymptomatischen Hypoglykämie und einem bestätigten Blutzuckerwert von < 63 mg/dL</p> <p>Andere Hypoglykämien: Anzahl der Patienten mit einer symptomatischen Hypoglykämie (patientenberichtet), keine bestätigte Messung wurde durchgeführt</p> <p>Bestätigte Hypoglykämien: Anzahl von Patienten mit mindestens einer bestätigten Hypoglykämie innerhalb der 52/104 Wochen (Blutzuckerwert ≤50 mg/dL) <i>Post-hoc</i> Berechnung</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-24: Bewertung des Verzerrungspotenzials für „Hypoglykämien“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D1690C00004	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Auftreten einer Hypoglykämie war in dieser Studie neben der Gewichtsveränderung ein hierarchisch getesteter sekundärer *Key-Endpunkt*. Es wurde die Anzahl der Patienten, bei denen mindestens eine Hypoglykämie berichtet wurde (mittels eines von den Patienten geführten Tagebuchs und eines für die Studie ausgegebenen Glukometers), erhoben. Die Endpunkterheber waren zu jedem Zeitpunkt verblindet. In allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn vergleichbare Ausfallraten. Das ITT-Prinzip wurde adäquat umgesetzt, da das *Full Analysis Set* bestehend aus allen randomisierten Patienten, die mindestens eine Studienmedikationseinnahme und einen Folgewert hatten, analysiert wurden. Da dies ebenfalls ein *Safety-Endpunkt* war, wurde die Analyse zusätzlich mit dem *Safety Analysis Set* durchgeführt. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Somit kann das Verzerrungspotenzial für diesen Endpunkt mit niedrig bewertet werden.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-25: Ergebnisse für Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie D1690C00004	Zu bewertendes Arzneimittel N / n (%)		Vergleichstherapie N / n (%)		Odds Ratio [95%-KI]
	Metformin+Dapagliflozin		Metformin+Sulfonylharnstoff		
52 Wochen					
Patienten mit mindestens einer Hypoglykämie**	400	14 (3,5*) SE: 0,919	401	162 (40,8*) SE: 2,403	0,05 [0,03;0,09]
Hypoglykämie***	406	14 (3,4)	408	162 (39,7)	0,05 [0,03;0,10]
schwer		0		3 (0,7)	0,14 [0,01;2,77]
leicht		7 (1,7)		147 (36,0)	0,03 [0,01;0,07]
andere		7 (1,7)		40 (9,8)	0,16 [0,07;0,36]
Abbruch aufgrund von Hypoglykämien		0		6 (1,5)	0,08 [0,00;1,36]
104 Wochen					
Hypoglykämie***	406	17 (4,2)	408	187 (45,8)	0,05 [0,03;0,09]
schwer		0		3 (0,7)	0,14 [0,01;2,77]
leicht		10 (2,5)		173 (42,4)	0,03 [0,02;0,07]
andere		8 (2,0)		45 (11,0)	0,16 [0,08;0,35]
Abbruch aufgrund von Hypoglykämien		0		7 (1,7)	0,07 [0,00;1,15]

*adjustiert (bzgl. des HbA1c Ausgangswertes) mit Standardfehler

** sekundärer Key-Endpunkt mittels *Full Analysis Set* analysiert,***Sicherheitsendpunkt mittels *Safety Analysis Set* analysiert

SE: Standardfehler, KI: Konfidenzintervall

Für die Zielpopulation wurden zusätzlich die bestätigten Hypoglykämien (Anzahl von Patienten mit mindestens einer bestätigten Hypoglykämie [Blutzuckerwert ≤ 50 mg/dL]) als Sensitivitätsanalyse betrachtet. Der *Cut-off* von 50 mg/dL wurde in Anlehnung an den IQWiG-Bericht zu den Gliniden gewählt, da dieser eine hohe Messsicherheit aufweist (IQWiG, 2009a).

Tabelle 4-26: Ergebnisse für Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie D1690C00004	Zu bewertendes Arzneimittel N / n (%*)		Vergleichstherapie N / n (%*)		Odds Ratio [95%-KI]
	Metformin+Dapagliflozin		Metformin+Sulfonylharnstoff		
52 Wochen					
Patienten mit mindestens einer Hypoglykämie***	318	11 (3,4) SE:1,025	354	150 (42,1) SE: 2,579	0,05 [0,03;0,09]
Hypoglykämie**	318	11 (3,5)	359	150 (41,8)	0,05 [0,03;0,10]
schwer		0		3 (0,8)	0,16 [0,01;3,16]
leicht		5 (1,6)		137 (38,2)	0,03 [0,01;0,07]
andere		6 (1,9)		35 (9,7)	0,18 [0,08;0,44]
Abbruch aufgrund von Hypoglykämien	318	0	359	5 (1,4)	0,10 [0,01;1,84]
Bestätigte Hypoglykämien	318	3 (0,9) SE 0,542	354	48 (13,4) SE: 1,801	0,06 [0,02;0,20]
104 Wochen					
Hypoglykämien***	318	14 (4,4) SE: 1,150	354	173 (48,6) SE: 2,613	0,05 [0,03;0,09]
Hypoglykämie**	318	14 (4,4)	359	173 (48,2)	0,05 [0,03;0,09]
schwer		0		3 (0,8)	0,16 [0,01;3,11]
leicht		8 (2,5)		161 (44,8)	0,03 [0,02;0,07]
andere		6 (1,9)		39 (10,9)	0,16 [0,07;0,38]
Abbruch aufgrund von Hypoglykämien	318	0	359	6 (1,7)	0,09 [0,00;1,52]
Bestätigte Hypoglykämien	318	3 (0,9) SE 0,542	354	56 (15,7) SE: 1,923	0,05 [0,02;0,16]

* adjustiert (bzgl. des HbA1c-Ausgangswertes) mit Standardfehler

** Sicherheitsendpunkt mittels *Safety Analysis Set* analysiert*** sekundärer Key-Endpunkt (mittels *Full Analysis Set*) analysiert

SE: Standardfehler, KI: Konfidenzintervall

Sowohl nach 52 als auch nach 104 Wochen hatten statistisch signifikant mehr Patienten in der Metformin+Sulfonylharnstoff-Gruppe mindestens eine Hypoglykämie im Vergleich zur Metformin+Dapagliflozin-Gruppe (42,1% / 48,6% gegen 3,4% / 4,4%). Schwere Hypoglykämien traten sowohl nach 52 als auch nach 104 Wochen nur sehr selten auf (0% unter Metformin+Dapagliflozin versus 0,8% unter Metformin+Sulfonylharnstoff). In der Metformin+Sulfonylharnstoff-Gruppe gab es insgesamt 5 (1,4%) Studienabbrüche aufgrund von Hypoglykämien nach 52 Wochen bzw. 6 (1,7%) nach 104 Wochen, wohingegen es in der

Metformin+Dapagliflozin-Gruppe keine gab. Insgesamt ergaben sich zwischen den Ergebnissen der Studienpopulation und den Ergebnissen der Zielpopulation keine wesentlichen Unterschiede. Auch bezüglich der Sensitivitätsanalyse gab es in der Metformin+Dapagliflozin-Gruppe gegenüber der Metformin+Sulfonylharnstoff-Gruppe signifikant weniger bestätigte Hypoglykämien sowohl nach 52 als auch nach 104 Wochen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Da nur eine Studie vorliegt, wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.3 Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung

der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-27: Operationalisierung von gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

Studie	Operationalisierung
D1690C00004	Anhand der Häufigkeiten von Hypoglykämien unter Berücksichtigung der Blutzuckerkontrolle

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-28: Bewertung des Verzerrungspotenzials für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D1690C00004	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Da das Verzerrungspotenzial für die einzelnen Endpunkte „Ausmaß der Blutzuckerkontrolle“ und „Hypoglykämien“ als niedrig bewertet wurde, kann auch für die gemeinsame Betrachtung der Endpunkte das Verzerrungspotenzial niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Im Folgenden werden zunächst die in den vorherigen Kapiteln 4.3.1.3.1.1 und 4.3.1.3.1.2 dargestellten Endpunkte „Ausmaß der Blutzuckerkontrolle“ und „Hypoglykämien“

gemeinsam betrachtet, da das Auftreten von Hypoglykämien nur unter Berücksichtigung der Blutzuckerkontrolle sinnvoll beurteilt werden kann.

Tabelle 4-29: Ergebnisse für „gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien“ aus RCT mit dem zu bewertenden Arzneimittel (Studien- und Zielpopulation)

Zeitpunkt	Studie D1690C00004	Blutzuckerkontrolle	Gesamtrate Hypoglykämien	Beurteilung
52 Wochen	Metformin+Dapagliflozin vs. Metformin+Sulfonylharnstoff	kein Unterschied	Vorteil	Zusatznutzen
104 Wochen	Metformin+Dapagliflozin vs. Metformin+Sulfonylharnstoff	kein Unterschied	Vorteil	Zusatznutzen

Kein Unterschied: kein Vor- bzw. Nachteil für eine der beiden Behandlungsoptionen.
Vorteil: Statistisch signifikanter und klinisch relevanter Unterschied zu Gunsten des zu bewertenden Arzneimittels. Als klinisch relevant wird in Analogie zur Definition des Nachteils ein HbA1c Unterschied von >0,35% angesehen. Im IQWiG-Abschlussbericht des Projektes „Urin- und Blutzuckerselbstmessung bei Diabetes mellitus Typ 2“ wurden beim Vertauschen der Null- und Alternativhypothese beim Test auf Nichtunterlegenheit auch andere Schwellen, z.B. eine Halbierung der vorgegebenen Nichtunterlegenheitsschwelle (>0,175%) diskutiert. Eine patientenrelevante Hypoglykämie ist eine Hypoglykämie, welche symptomatisch ist und als unerwünschtes Ereignis berichtet wird. Ein niedriger Blutzuckerwert ohne Symptome wird nicht als relevant angesehen.

Bei dem Vergleich von Metformin+Dapagliflozin mit Metformin+Sulfonylharnstoff lag hinsichtlich der HbA1c-Senkung zum Zeitpunkt 52 Wochen keine Unterlegenheit in einer der Behandlungsgruppen vor. Da die glykämische Kontrolle unter Metformin+Sulfonylharnstoff im Verlauf stark abnahm und unter Metformin+Dapagliflozin relativ konstant blieb, war die HbA1c-Senkung unter Metformin+Dapagliflozin bei Woche 104 statistisch signifikant stärker. Darüber hinaus gab es in der Metformin+Sulfonylharnstoff-Gruppe eine statistisch signifikant höhere Anzahl von Patienten mit Hypoglykämien. Aus der gemeinsamen Betrachtung von Blutzuckerkontrolle und Anzahl der Hypoglykämien ergibt sich somit eine Überlegenheit für Metformin+Dapagliflozin gegenüber Metformin+Sulfonylharnstoff und somit ein Zusatznutzen für Metformin+Dapagliflozin.

4.3.1.3.1.4 Gewichtsveränderung – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-30: Operationalisierung von Gewichtsveränderung

Studie	Operationalisierung
D1690C00004	Anhand der absoluten Änderung des Körpergewichts in kg von Studienbeginn bis Woche 52 (falls der Wert nicht erhoben wurde, wurde der letztmögliche vor Woche 52 (LOCF) betrachtet) bzw. bis Woche 104 (<i>Repeated Measurement</i> Modell). Das Gewicht wurde ohne Schuhe, mit leichter Kleidung und nüchtern („on a fasting stomach“) gemessen Anhand einer Responder-Analyse: <ul style="list-style-type: none"> - Anzahl der Patienten mit einer Gewichtsabnahme von mindestens 5% nach Woche 52 und 104.

LOCF: *Last observation carried forward*

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-31: Bewertung des Verzerrungspotenzials für Gewichtsveränderung in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D1690C00004	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Gewichtsveränderung war in dieser Studie ein hierarchisch getesteter sekundärer Key-Endpunkt und wurde anhand der Differenz des Gewichts zu Woche 52 mit dem Gewicht zu Studienbeginn verglichen. Ergänzend wurde die Anzahl der Patienten mit einer Gewichtsabnahme von mehr als 5% in beiden Gruppen untersucht. Die Endpunkterheber waren zu jedem Zeitpunkt der Erhebung verblindet, in allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn vergleichbare Ausfallraten. Das ITT-Prinzip wurde adäquat umgesetzt, das *Full Analysis Set* bestehend aus allen randomisierten Patienten, die mindestens eine Studienmedikationseinnahme und einen Folgewert hatten, wurde analysiert. Fehlende Werte wurden durch die LOCF-Methode ersetzt. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Somit wird das Verzerrungspotenzial für diesen Endpunkt mit niedrig bewertet. Um die Langzeitwirkung und auch die Robustheit der Ergebnisse zu untersuchen, wurde der Endpunkt zusätzlich mit einem *Repeated-Measurement-Modell* über einen Zeitraum von 104 Wochen untersucht.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-32: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert in kg (SE) [95%-KI]
	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert kg (SE)	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert kg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen**									
	400	88,44 (16,323)	85,22 (16,191)	-3,22 (0,1756)	401	87,60 (16,970)	89,04 (17,560)	1,44 (0,1754)	-4,65 (0,2483)* [-5,14;-4,17]
52 Wochen°									
	323		85,90 (16,267)	-3,38 (0,1954)	315		89,46 (17,422)	1,55 (0,1969)	-4,93 (0,2776) [-5,47;-4,38]
104 Wochen°									
	234		85,61 (16,331)	-3,70 (0,2352)	211		90,90 (17,243)	1,36 (0,2427)	-5,06 (0,3379) [-5,73;-4,40]

SD: Standardabweichung; SE: Standardfehler, KI: Konfidenzintervall

*ANCOVA-Modell (*Analysis of covariance*) mit Behandlung als Faktor und Ausgangsgewicht als Kovariate. p <0,0001

** LOCF

° *Repeated-Measurements*-Modell

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-33: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert kg (SE) [95%-KI]
	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte** Veränderung Mittelwert % (SE)	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte** Veränderung Mittelwert kg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
	318	88,79 (16,809)	85,30 (16,667)	-3,49 (0,2000)	354	87,39 (17,308)	88,93 (17,879)	1,54 (0,1896)	-5,02 (0,2757) [-5,57;-4,48]
52 Wochen°									
	274		85,98 (16,657)	-3,58 (0,2150)	284		89,08 (17,562)	1,65 (0,2096)	-5,23 (0,3006) [-5,82;-4,64]
104 Wochen°									
	197		85,66 (16,558)	-3,82 (0,2510)	195		90,32 (17,335)	1,56 (0,2482)	-5,38 (0,3531) [-6,07;-4,68]

*LOCF

** adjustiert nach Ausgangsgewicht

SD: Standardabweichung; SE: Standardfehler, KI: Konfidenzintervall

° Repeated-Measurements Modell

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-34: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	Metformin+ Dapagliflozin N / n (%*; SE)		Metformin+ Sulfonylharnstoff N / n (%*; SE)		Odds Ratio [95%-KI]
D1690C00004					
52 Wochen	400	133 (33,3;2,354)	401	10 (2,5;0,779)	19,48 [10,05;37,73]
104 Wochen	400	95 (23,8;2,128)	401	11 (2,8;0,815)	11,04 [5,81;20,98]

*adjustiert nach Ausgangsgewicht
SE: Standardfehler, KI: Konfidenzintervall

Tabelle 4-35: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Metformin+ Dapagliflozin N / n (%*, SE)		Metformin+ Sulfonylharnstoff N / n (%*, SE)		Odds Ratio [95%-KI]
D1690C00004					
52 Wochen	318	116 (36,6;2,695)	354	9 (2,6;0,837)	22,01 [10,93;44,33]
104 Wochen	318	81 (25,5;2,443)	354	9 (2,6;0,836)	13,10 [6,45;26,60]

*adjustiert nach Ausgangsgewicht
SE: Standardfehler, KI: Konfidenzintervall

Patienten mit einer Metformin+Dapagliflozin-Behandlung hatten einen mittleren Gewichtsverlust von -3,49 kg, während Patienten mit einer Metformin+Sulfonylharnstoff-Behandlung innerhalb eines Jahres im Mittel um 1,54 kg Gewicht zunahmten. Es ergibt sich ein statistisch signifikanter Unterschied von -5,02 kg (95% KI: [-5,57;-4,48]) zugunsten von Metformin+Dapagliflozin. Das Resultat zeigte sich als robust in der *Repeated-Measurements*-Analyse: auch hier ergab sich nach 52 Wochen ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten der Metformin+Dapagliflozin-Gruppe (-5,23 kg, 95%-KI: [-5,82;-4,64]). Dieser Unterschied war über 104 Wochen stabil. Zu Woche 104 hatten die Patienten unter Metformin+Dapagliflozin durchschnittlich 5 kg mehr Gewicht verloren als Patienten unter Metformin+Sulfonylharnstoff (MWD: -5,38 kg 95%-KI: [-6,07;-4,68]). Ein signifikant größerer Anteil (36,6%) der Patienten mit Metformin+Dapagliflozin-Behandlung hatte nach 52 Wochen ein um mehr als 5% - relativ zum Ausgangsgewicht - reduziertes Körpergewicht, wohingegen nur 2,6% Patienten mit Metformin+Sulfonylharnstoff-Behandlung eine Gewichtsreduktion von mehr als 5% erreichten (OR [95%-KI]: 22,01 [10,93;44,33]). Nach 104 Wochen waren es noch 81 (25,5%) Patienten in der Metformin+Dapagliflozin-Gruppe und 9 (2,6%) in der Metformin+Sulfonylharnstoff-Gruppe (OR [95%-KI]: 13,10 [6,45;26,60]).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Da nur eine Studie vorliegt, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.1.5 Therapiezufriedenheit (DTSQ) – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind

- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-36: Operationalisierung von Therapiezufriedenheit (DTSQ)

Studie	Operationalisierung
D1690C00004	Zufriedenheit der Patienten mit der Behandlung (anhand des DTSQ-Fragebogens zur Therapiezufriedenheit von Diabetes-Patienten). Der DTSQ hat 2 Versionen (Status-Version DTSQs und die <i>Change</i> -Version DTSQc). Der DTSQ wurde von den Patienten selbst ausgefüllt. Der DTSQs enthält eine Skala von 0–36 und wurde nach Randomisierung (Woche 0) ausgefüllt. Je höher der <i>Score</i> , desto zufriedener war der Patient mit seiner Therapie. In Woche 52 füllten die Patienten den DTSQc aus, der die relative Veränderung in der Therapiezufriedenheit misst. Es konnten Werte von -18 zu +18 erreicht werden, wobei 0 keine Veränderung bedeutete.

DTSQ: *Diabetes Treatment Satisfaction Questionnaire*

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-37: Bewertung des Verzerrungspotenzials für Therapiezufriedenheit (DTSQ) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D1690C00004	niedrig	ja	unklar	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Erfassung der Therapiezufriedenheit der Patienten mittels eines Fragebogens war in der Studie ein sekundärer Endpunkt und der Status wurde in Woche 52 erhoben. Die Patienten waren zu jedem Zeitpunkt bezüglich der Behandlung verblindet, in allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn vergleichbare Ausfallraten. Fehlende Werte wurden nicht imputiert. Es gab daher unterschiedlich Abbrecherraten zwischen den Gruppen, so dass eine adäquate Umsetzung des ITT-Prinzips fraglich ist. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Es wird angenommen, dass das Verzerrungspotenzial für diesen Endpunkt niedrig ist.

Tabelle 4-38: Ergebnisse für Therapiezufriedenheit (DTSQ) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert DTSQs Mittelwert (SD)	Endwert DTSQc Adjustierter Mittelwert* (SE)	N	Ausgangswert DTSQs Mittelwert (SD)	Endwert DTSQc Adjustierter Mittelwert (SE)	
52 Wochen	Metformin+Dapagliflozin			Metformin+Sulfonylharnstoff			
	312	30,9 (5,41)	14,3 (0,288)	303	31,5 (5,01)	13,6 (0,292)	0,7 (0,411) [-0,1;1,5]*

SD: Standardabweichung; SE: Standardfehler, KI: Konfidenzintervall

*p: 0,0797 (ANCOVA),

DTSQ: *Diabetes Treatment Satisfaction Questionnaire*

Tabelle 4-39: Ergebnisse für Therapiezufriedenheit (DTSQ) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppen- unterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert DTSQs Mittelwert (SD)	Endwert DTSQc Adjustierter Mittelwert (SE)	N	Ausgangswert DTSQs Mittelwert (SD)	Endwert DTSQc Adjustierter Mittelwert (SE)	
52 Wochen	Metformin+Dapagliflozin			Metformin+Sulfonylharnstoff			
	266	31,0 (5,49)	14,1 (0,317)	274	31,5 (5,17)	13,6 (0,312)	0,5 (0,445) [-0,3;1,4]

SD: Standardabweichung; SE: Standardfehler, KI: Konfidenzintervall

DTSQ: *Diabetes Treatment Satisfaction Questionnaire*

Die mittleren Werte der Therapiezufriedenheit basierend auf den DTSQc waren in beiden Behandlungsgruppen hoch (Metformin+Dapagliflozin: 14,1 und Metformin+Sulfonylharnstoff: 13,6), was eine Verbesserung der Therapiezufriedenheit in beiden Behandlungsgruppen bedeutet. In der Metformin+Dapagliflozin-Gruppe war der Wert höher als in der Metformin+Sulfonylharnstoff-Gruppe, jedoch ergab sich kein signifikanter Unterschied. Zum Zeitpunkt 104 Wochen wurde der DTSQ nicht erhoben.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 5: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Da nur eine RCT vorliegt, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.1.6 Blutdrucksenkung – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die

Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-40: Operationalisierung von Blutdrucksenkung

Studie	Operationalisierung
D1690C00004	Anhand der Änderung des systolischen und diastolischen Blutdruckes (mmHg) von Studienbeginn bis Woche 52 (falls der Wert nicht erhoben wurde, wurde der letztmögliche vor Woche 52 (LOCF) betrachtet) bzw. bis Woche 104 (<i>Repeated-Measurements</i> Analyse). Es wurden an den Visiten 3 Blutdruckmessungen (es wurde der Durchschnitt von den 3 Messungen für die Analyse verwendet) vorgenommen mit einer Pause von einer Minute dazwischen. Davor sollte der Patient mindestens 5 Minuten in einer sitzenden Position ruhig verbringen. Es sollte immer das gleiche Blutdruckgerät für alle Visiten für den einzelnen Patienten verwendet werden. Eine Änderung der antihypertensiven Basistherapie wurde nicht berücksichtigt.

LOCF: *Last observation carried forward*

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-41: Bewertung des Verzerrungspotenzials für Blutdrucksenkung in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D1690C00004	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Blutdrucksenkung war in dieser Studie ein vordefinierter sekundärer Endpunkt und wurde als Differenz des systolischen bzw. diastolischen Blutdruckes zu Woche 52 zum jeweiligen Wert zu Studienbeginn berechnet. Die Endpunkterheber waren zu jedem Zeitpunkt der Erhebung verblindet. In allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn geringe Ausfallraten. Das ITT-Prinzip wurde adäquat umgesetzt, da alle randomisierten Patienten, die mindestens eine Studienmedikationseinnahme und einen Folgewert hatten, in die Analyse eingeschlossen wurden und entsprechend ihrer Randomisierung ausgewertet wurden (*Full Analysis Set*). Dieser Endpunkt war auch ein *Safety-Endpunkt* und wurde mittels *Safety Analysis Set* untersucht (wird nicht dargestellt). Fehlende Werte wurden durch die LOCF-Methode ersetzt. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Es gilt zu berücksichtigen, dass eine Änderung der antihypertensiv Basistherapie der Patienten nicht berücksichtigt wurde und es daher zu einer Verzerrung des Endpunkts kommen kann. Da dies jedoch für beide Behandlungsgruppen der Fall war, kann von einem konstanten Behandlungseffekt ausgegangen werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt mit niedrig bewertet. Um die Langzeitwirkung und auch die Robustheit der Ergebnisse zu untersuchen, wurde der Endpunkt zusätzlich mit einem *Repeated-Measurements-Modell* untersucht.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-42: Ergebnisse für Blutdrucksenkung aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppen- unterschied Mittelwert mmHg (SE) [95%-KI]
	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	N	Ausgangs- wert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
systolisch	399	132,8 (14,89)	128,8 (14,21)	-4,3 (0,585)	396	133,8 (14,69)	134,3 (14,17)	0,8 (0,587)	-5,0 (0,829) [-6,7;-3,4] **
diastolisch	399	80,6 (8,42)	79,0 (8,92)	-1,6 (0,367)	396	80,6 (8,46)	80,2 (8,64)	-0,4 (0,369)	-1,2 (0,521) [-2,3;-0,2]**
52 Wochen°									
systolisch	323		129,60 (13,861)	-3,79 (0,6396)	314		134,36 (14,148)	0,93 (0,6473)	-4,72 (0,9101) [-6,50;-2,93]
diastolisch	323		79,60 (8,840)	-1,30 (0,4015)	314		80,16 (8,645)	-0,43 (0,4061)	-0,87 (0,5712) [-1,99;0,25]
104 Wochen°									
systolisch	234		130,18 (14,244)	-2,69 (0,7725)	211		134,97 (13,775)	1,20 (0,8067)	-3,89 (1,1169) [-6,08;-1,69]
diastolisch	234		78,84 (8,714)	-1,95 (0,4619)	211		79,22 (8,488)	-1,51 (0,4817)	-0,45 (0,6678) [-1,76;0,86]

SD: Standardabweichung; SE: Standardfehler, KI: Konfidenzintervall

*LOCF

**ANCOVA-Modell p < 0,0001 / p = 0,0179

°Repeated-Measurements-Modell

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-43: Ergebnisse für Blutdrucksenkung aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
systolisch	318	132,6 (14,41)	128,3 (13,12)	-4,4 (0,649)	351	133,0 (13,87)	133,8 (13,70)	1,0 (0,618)	-5,3 (0,896) [-7,1;-3,6]
diastolisch	318	81,1 (8,54)	79,5 (8,82)	-1,6 (0,415)	351	80,9 (8,27)	80,4 (8,61)	-0,5 (0,395)	-1,1 (0,573) [-2,2;0,0]
52 Wochen°									
systolisch	274		129,19 (12,858)	-3,99 (0,6831)	283		133,94 (13,578)	1,06 (0,6687)	-5,05 (0,9560) [-6,93;-3,17]
diastolisch	274		79,90 (8,714)	-1,30 (0,4395)	283		80,36 (8,605)	-0,48 (0,4296)	-0,82 (0,6148) [-2,03;0,39]
104 Wochen°									
systolisch	197		129,60 (13,087)	-3,15 (0,7929)	195		134,86 (13,098)	1,31 (0,7917)	-4,46 (1,1205) [-6,67;-2,26]
diastolisch	197		79,01 (8,269)	-2,14 (0,4863)	195		79,62 (8,174)	-1,37 (0,4850)	-0,77 (0,6874) [-2,12;0,58]

*LOCF

SD: Standardabweichung; SE: Standardfehler, KI: Konfidenzintervall

° Repeated-Measurements-Modell

Patienten mit einer Metformin+Dapagliflozin-Behandlung hatten im Mittel nach 52 Wochen eine größere Blutdrucksenkung (systolisch: -4,4 mmHg / diastolisch: -1,6 mmHg) als Patienten mit einer Metformin+Sulfonylharnstoff-Behandlung (systolisch: +1,0 mmHg / diastolisch: -0,5 mmHg). Der Unterschied zwischen den Behandlungsoptionen war für den systolischen Blutdruck signifikant (-5,3 mmHg), für den diastolischen Blutdruck nicht (-1,1 mmHg). Mit dem *Repeated-Measurements*-Modell war der Unterschied im systolischen Blutdruck nach 52 und 104 Wochen ebenfalls statistisch signifikant zugunsten von Metformin+Dapagliflozin. Die Differenz im diastolischen Blutdruck war auch nach 104 Wochen nicht signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 6: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Da nur eine RCT vorliegt, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.1.7 Unerwünschte Ereignisse – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-

Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-44: Operationalisierung von Unerwünschte Ereignisse

Studie	Operationalisierung
D1690C00004	Gesamtrate unerwünschter Ereignisse Schwerwiegende unerwünschte Ereignisse Studienabbruch wegen unerwünschter Ereignisse Anzeichen einer Genital- bzw. Harnwegsinfektion

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-45: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D1690C00004	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt „Unerwünschte Ereignisse“ in der Studie ist niedrig. Die Erfassung der unerwünschten Ereignisse erfolgte fortlaufend von Randomisierung bis zum letzten definierten Zeitpunkt. Die Endpunkterheber waren zu jedem Zeitpunkt der Erhebung verblindet. In allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn vergleichbare Ausfallraten. Der Endpunkt wurde bei behandelten Patienten, die mindestens eine Dosis der Studienmedikation bekommen hatten, erhoben (*Safety Analysis Set*). Eine statistische Auswertung der Ergebnisse fand wie geplant statt.

Tabelle 4-46: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	N	Gesamtrate unerwünschter Ereignisse*, n (%)	Schwerwiegende unerwünschte Ereignisse*, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse*, n (%)
D1690C00004				
Nach 52 Wochen				
Metformin+Dapagliflozin	406	318 (78,3)	35 (8,6)	37 (9,1)
Metformin+Sulfonylharnstoff	408	318 (77,9)	46 (11,3)	24 (5,9)
OR [95%-KI]		1,02 [0,73;7,43]	0,74 [0,47;1,18]	1,60 [0,94;2,73]
Nach 104 Wochen				
Metformin+Dapagliflozin	406	337 (83,0)	51 (12,6)	40 (9,9)
Metformin+Sulfonylharnstoff	408	338 (82,8)	62 (15,2)	31 (7,6)
OR [95%-KI]		1,01 [0,70;1,46]	0,80 [0,54;1,19]	1,33 [0,81;2,17]

*Nur Hypoglykämien, die als schweres unerwünschtes Ereignis berichtet wurden, sind bei der Gesamtrate unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse und Studienabbrüche aufgrund unerwünschter Ereignisse eingeschlossen.

OR: Odds Ratio, KI: Konfidenzintervall

Tabelle 4-47: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Gesamtrate unerwünschter Ereignisse*, n (%)	Schwerwiegende unerwünschte Ereignisse*, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse*, n (%)
D1690C00004				
Nach 52 Wochen				
Metformin+Dapagliflozin	318	255 (80,2)	23 (7,2)	14 (4,4)
Metformin+Sulfonylharnstoff	359	283 (78,8)	39 (10,9)	17 (4,7)
OR [95%-KI]		1,09 [0,75;1,58]	0,64 [0,37;1,10]	0,93 [0,45;1,91]
Nach 104 Wochen				
Metformin+Dapagliflozin	318	271 (85,2)	36 (11,3)	16 (5,0)
Metformin+Sulfonylharnstoff	359	297 (82,7)	51 (14,2)	21 (5,8)
OR [95%-KI]		1,20 [0,80;1,82]	0,77 [0,49;1,22]	0,85 [0,44;1,66]

*Nur Hypoglykämien, die als schweres unerwünschtes Ereignis berichtet wurden, sind bei der Gesamtrate unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse und Studienabbrüche aufgrund unerwünschter Ereignisse eingeschlossen.

OR: Odds Ratio, KI: Konfidenzintervall

Die Gesamtrate der Patienten, die mindestens ein unerwünschtes Ereignis (80,2% versus 78,8%) hatten, die Gesamtrate der Patienten mit schwerwiegenden unerwünschten Ereignissen (7,2% versus 10,9%) sowie der Anteil der Patienten, die aufgrund unerwünschter Ereignisse die Studie abbrachen (4,4% versus 4,7%) war in beiden Gruppen ähnlich. Auch nach 104 Wochen ergaben sich keine Unterschiede zwischen den Behandlungsoptionen.

Ergänzend zu den Daten der unerwünschten Ereignisse werden hier auch die Daten zu unerwünschten Ereignissen von besonderem Interesse – Anzeichen für Genital- und Harnwegsinfektionen – dargestellt. Die Symptome zu diesen Ereignissen wurden proaktiv von den Prüfarzten bei jeder Visite erfragt und dokumentiert. Da die Symptome häufig nur im Nachhinein erfragt werden konnten und ein Kulturnachweis dann nicht mehr möglich war, wurden unter dem Endpunkt unspezifische Symptome wie häufiger Harndrang, Genitalrötung, –juckreiz oder Dysurie, sowie nachgewiesene Genital- und Harnwegsinfektionen aufsummiert.

Tabelle 4-48: Ergebnisse für Unerwünschte Ereignisse – Anzeichen für Genital- und Harnwegsinfektionen aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	N	Anzahl der Patienten mit Anzeichen für eine Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen für eine Harnwegsinfektion, n (%)
D1690C00004			
Nach 52 Wochen			
Metformin+Dapagliflozin	406	50 (12,3)	44 (10,8)
Metformin+Sulfonylharnstoff	408	11 (2,7)	26 (6,4)
OR [95%-KI]		5,07 [2,60;9,89]	1,79 [1,08;2,96]
Nach 104 Wochen			
Metformin+Dapagliflozin	406	60 (14,8)	55 (13,5)
Metformin+Sulfonylharnstoff	408	12 (2,9)	37 (9,1)
OR [95%-KI]		5,72 [3,03;10,81]	1,57 [1,01;2,44]

OR: Odds Ratio, KI: Konfidenzintervall

Tabelle 4-49: Ergebnisse für Unerwünschte Ereignisse – Anzeichen für Genital- und Harnwegsinfektionen aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Anzahl der Patienten mit Anzeichen für eine Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen für eine Harnwegsinfektion, n (%)
D1690C00004			
Nach 52 Wochen			
Metformin+Dapagliflozin	318	42 (13,2)	37 (11,6)
Metformin+Sulfonylharnstoff	359	9 (2,5)	22 (6,1)
OR [95%-KI]		5,92 [2,83;12,37]	2,02 [1,16;3,50]
Nach 104 Wochen			
Metformin+Dapagliflozin	318	51 (16,0)	48 (15,1)
Metformin+Sulfonylharnstoff	359	10 (2,8)	33 (9,2)
OR [95%-KI]		6,67 [3,32;13,37]	1,76 [1,10;2,81]

OR: Odds Ratio, KI: Konfidenzintervall

Unter Metformin+Dapagliflozin hatten die Patienten nach 52 Wochen sowie nach 104 Wochen signifikant häufiger Anzeichen für eine Genitalinfektion bzw. Harnwegsinfektion als unter Metformin+Sulfonylharnstoff. Die Mehrzahl der Anzeichen für Genitalinfektionen deutete auf Pilzinfektionen hin.

Die Symptome waren in der Regel leicht bis mittelschwer in der Ausprägung und konnten mit den gängigen Behandlungsmethoden (antimikrobielle Wirkstoffe, überwiegend Antibiotika) einfach behandelt werden.

Für die Studienpopulation werden die aufgetretenen Anzeichen für Genital- und Harnwegsinfektionen nach 104 Wochen im Folgenden detailliert beschrieben:

Lediglich drei Patienten (0,7%) in der mit Metformin+Dapagliflozin behandelten Gruppe brachen aufgrund Anzeichen für eine Genitalinfektion und ein (0,2%) Patient aufgrund eines Anzeichens für eine Harnwegsinfektion die Behandlung ab. In der Metformin+Sulfonylharnstoff-Gruppe wurde die Behandlung bei einem Patienten wegen eines Anzeichens für eine obere Harnwegsinfektion abgebrochen.

Da es sich bei der Anzahl der Anzeichen für Genital- und Harnwegsinfektionen um Patientenangaben aufgrund bestimmter Symptome handelt und nicht in jedem Fall um Laboruntersuchungen auf entsprechende Keime, ist die Inzidenz der nachgewiesenen Infekte geringer als die in den Tabellen angegebenen Zahlen.

Entzündungen der oberen Harnwege waren selten, sie traten nur bei zwei Patienten unter Metformin+Sulfonylharnstoff auf.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Da nur eine RCT vorliegt, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.1.8 Kardiale Ereignisse – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-50: Operationalisierung von Kardiale Ereignisse

Studie	Operationalisierung ^a
D1690C00004	Anzahl der Patienten mit mindestens einem kardialen Ereignis innerhalb von 52 bzw. 104 Wochen. Die Erhebung der kardialen Ereignisse erfolgte im Rahmen der Erhebung der unerwünschten Ereignisse (SOC für kardiale Ereignisse)

SOC: System Organ Class

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-51: Bewertung des Verzerrungspotenzials für „Kardiale Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D1690C00004	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt „Kardiale Ereignisse“ in der Studie war niedrig. Für diesen Endpunkt lag eine Auswertung der MedDRA-Systemorganklasse (SOC) „Kardiale Ereignisse“ vor. In der SOC Ebene sind kardiale Ereignisse angemessen zusammengefasst. Zudem stellt die Art der Auswertung sicher, dass Patienten, die mehrere Ereignisse aus dieser Klasse berichteten nur einmal in der Auswertung berücksichtigt werden. Die Endpunkterheber waren zu jedem Zeitpunkt der Erhebung verblindet, in allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn geringe Ausfallraten. Der Endpunkt wurde bei behandelten Patienten, die mindestens eine Dosis der Studienmedikation bekommen hatten, erhoben (*Safety Analysis Set*).

Tabelle 4-52: Ergebnisse für „Kardiale Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	N	Gesamtrate kardialer Ereignisse, n (%)
D1690C00004		
Nach 52 Wochen		
Metformin+Dapagliflozin	406	30 (7,4)
Metformin+Sulfonylharnstoff	408	32 (7,8)
OR [95%-KI]		0,94 [0,56;1,57]
Nach 104 Wochen		
Metformin+Dapagliflozin	406	38 (9,4)
Metformin+Sulfonylharnstoff	408	42 (10,3)
OR [95%-KI]		0,90 [0,57;1,43]

OR: Odds Ratio, KI: Konfidenzintervall

Tabelle 4-53: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
D1690C00004		
Nach 52 Wochen		
Metformin+Dapagliflozin	318	24 (7,5)
Metformin+Sulfonylharnstoff	359	26 (7,2)
OR [95%-KI]		1,05 [0,59;1,86]
Nach 104 Wochen		
Metformin+Dapagliflozin	318	31 (9,7)
Metformin+Sulfonylharnstoff	359	36 (10,0)
OR [95%-KI]		0,97 [0,58;1,61]

OR: Odds Ratio, KI: Konfidenzintervall

Kardiale Ereignisse traten über den Zeitraum von 52 Wochen in beiden Behandlungsgruppen in vergleichbarer Häufigkeit auf (7,5% versus 7,2%). Nach 104 Wochen hatten in der Metformin+Dapagliflozin-Gruppe 31 (9,7%) Patienten ein kardiales Ereignis, wohingegen in der Metformin+Sulfonylharnstoff-Gruppe 36 (10,0%) Patienten ein solches Ereignis hatten. Signifikante Unterschiede zwischen den Behandlungsoptionen fanden sich nicht.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 8: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Da nur eine RCT vorliegt, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.2 Subgruppenanalysen – RCT

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Für die im Studienprotokoll geplanten Subgruppenanalysen wurden Interaktionstests berichtet. Diese beziehen sich ausschließlich auf den Endpunkt HbA1c und auf zusammenfassende sicherheitsrelevante Endpunkte und basieren auf der Studienpopulation.

Im Folgenden werden zunächst die p-Werte der Interaktionstests der präspezifizierten Subgruppenanalysen gemäß Studienbericht für den Endpunkt „Ausmaß der Blutzuckerkontrolle (HbA1c)“ vom Ausgangswert bis Woche 52 angegeben:

Tabelle 4-54: Interaktionstest der präspezifizierten Subgruppenanalysen für den Endpunkt Ausmaß der Blutzuckerkontrolle (HbA1c)

Merkmal	Subgruppen	Interaktion p-Wert
Geschlecht	(männlich / weiblich)	p=0,1061
Alter	(<65 / ≥65 / ≥75 Jahre)	p=0,2178
Alter und Geschlecht	(Alter der Frau ≤50 / Alter der Frau >50 Jahre)	p=0,1307
Ethnische Zugehörigkeit	(weiß / schwarz / asiatisch / anders)	p=0,0362
Ethnizität	(Latino / kein Latino)	p=0,4041
Ausgangswert des Body Mass Index	(<25 kg/m ² , ≥25 kg/m ²)	p=0,8960
Geografische Herkunft	(Latein-Amerika / Europa)	p=0,5515
HbA1c-Ausgangswert	(<8% / ≥8-<9% / ≥9%)	p=0,2038
eGFR-Ausgangswert	(<30 / ≥30-<60 / ≥60-<90 / ≥90 mL/min/1,73m ²)	p=0,0055

HbA1c: Glykiertes Hämoglobin; eGFR: geschätzte glomeruläre Filtrationsrate

Signifikante Interaktionen wurden hinsichtlich folgender Faktoren beobachtet: Zwischen den Behandlungsgruppen und der ethnischen Zugehörigkeit sowie zwischen den Behandlungsgruppen und dem eGFR-Ausgangswert. Für die Interaktion hinsichtlich der ethnischen Zugehörigkeit ist jedoch keine Interpretation der Interaktion möglich, da die Subgruppen der Asiaten (n=27 / 36), der Schwarzen (n=26 / 24) und der Patienten anderer ethnischer Zugehörigkeit (n=20 / 20) sehr viel kleiner waren als die Subgruppe der Weißen (n=327 / 323). Auch ist der Behandlungseffekt in jeder dieser Subgruppen gleich (es gibt keinen signifikanten Unterschied zwischen den Behandlungsoptionen). Patienten mit eGFR ≥60 mL/min/1,73m² zu Beginn der Studie hatten mit der Metformin+Dapagliflozin-Behandlung eine größere HbA1c-Senkung als mit der Metformin+Sulfonylharnstoff-Behandlung, wohingegen Patienten mit einer Nierenfunktionsstörung (*Baseline*

eGFR <60 mL/min/1,73m²) eine geringere Wirksamkeit unter der Behandlung mit Metformin+Dapagliflozin aufwiesen. Für Patienten mit einer Nierenfunktionsstörung (*Baseline* eGFR <60 mL/min/1,73m² bzw. Kreatininclearance <60 mL/min) wurde im Rahmen des Zulassungsverfahrens die Anwendung von Dapagliflozin nicht empfohlen (Bristol-Myers Squibb und AstraZeneca GmbH, 2012a).

Zusätzlich werden für alle Endpunkte Subgruppenanalysen (*post hoc*) nach

- Geschlecht
- Alter (<65 Jahre und ≥65 bis <75 Jahre)

durchgeführt. Diese Daten beruhen auf der Zielpopulation gemäß Zulassung (Bristol-Myers Squibb und AstraZeneca GmbH, 2012b).

Der Trennpunkt für den Faktor Alter wurde analog zu der in der Studie geplanten Klassifizierung für Alter gewählt. Der Trennpunkt erscheint insofern angemessen, als die Deutsche Diabetes Gesellschaft (DDG) den älteren Patienten als in der Regel älter als 65 Jahre definiert (Hader C et al., 2004). Das entspricht annähernd den oben dargestellten a priori definierten Alters-Trennpunkten. Dadurch können Subgruppenanalysen für erwachsene (<65 Jahre) und ältere (≥65 Jahre) Patienten dargestellt werden.

Eine Subgruppenanalyse nach Schweregrad der Diabetes Erkrankung war nicht erforderlich. Die Einteilung der Schweregradstufen richtet sich beim Diabetes bei den intensiver zu behandelnden Stufen in erster Linie nach der Notwendigkeit der Insulingabe in unterschiedlichen Stufen. Diabetes mellitus kann gemäß der DDG in die folgenden Stadien eingeteilt werden (Kerner et al., 2004; AkdÄ, 2009):

- Normale Blutglukoseregulation,
- Gestörte Glukosetoleranz,
- Typ-2-Diabetiker, der ohne Insulin behandelt wird,
- Typ-2-Diabetiker, der Insulin zur Blutzuckerkontrolle benötigt
- Typ-1-Diabetiker, für den die äußere Insulinzufuhr lebensnotwendig ist.

Da die hier betrachtete Patientenpopulation vollständig der Gruppe „Typ-2 Diabetiker, der ohne Insulin behandelt wird“ zugeordnet werden kann, ist eine Einteilung nach dem Schweregrad nicht sinnvoll.

Abweichend davon wurde in der IQWiG Nutzenbewertung von Linagliptin zwischen der Vorbehandlung der Patienten (Metformin allein oder Metformin+andere OAD) unterschieden. Eine Zusatzanalyse mit der Teilpopulation, die mit Metformin allein vorbehandelt war, wurde dem Studienbericht beigelegt

(M4B_D1690C00004_Zusatzanalyse_Teilpopulation_Metforminvorbehandlung). Es zeigte sich keine unterschiedliche Wirksamkeit bzw. Verträglichkeit aufgrund der Vorbehandlung.

Die Bewertung des Zusatznutzens beruht nur auf einer Studie. Aus diesem Grund wurden keine Meta-Analysen durchgeführt.

4.3.1.3.2.1 Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach dem Geschlecht – RCT

Bezüglich des Ausmaßes der Blutzuckerkontrolle anhand des HbA1c ergab sich kein Hinweis auf eine signifikante Interaktion mit dem Geschlecht der Patienten (Interaktionstest: $I^2=0\%$; $p=0,52$). Auch für die Operationalisierung Studienabbrüche aufgrund unzureichender glykämischer Kontrolle ergab sich kein Hinweis auf eine Effektmodifikation durch das Geschlecht der Patienten ($I^2=0\%$, $p=0,58$).

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-55: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert % (SE)	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert % (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
männlich	178	7,76 (0,852)	7,12 (0,746)	-0,62 (0,0635)	203	7,71 (0,861)	7,10 (1,105)	-0,62 (0,0595)	0,00 (0,0870) [-0,17;0,17]
weiblich	140	7,77 (0,895)	7,32 (0,806)	-0,44 (0,0642)	151	7,75 (0,902)	7,39 (0,990)	-0,36 (0,0618)	-0,08 (0,0891) [-0,26;0,09]
Interaktionstest**									I ² =0%;p=0,52
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen°									
männlich	155		7,11 (0,746)	-0,62 (0,0669)	171		7,03 (0,995)	-0,59 (0,0635)	-0,03 (0,0923) [-0,21;0,15]
weiblich	117		7,22 (0,745)	-0,47 (0,0638)	113		7,29 (0,857)	-0,36 (0,0639)	-0,11 (0,0903) [-0,29;0,07]
104 Wochen°									
männlich	110		7,05 (0,563)	-0,35 (0,0777)	126		7,01 (0,833)	-0,26 (0,0734)	-0,09 (0,1064) [-0,30;0,12]
weiblich	86		7,08 (0,594)	-0,33 (0,0793)	66		7,20 (0,645)	0,01 (0,0841)	-0,34 (0,1150) [-0,57;-0,11]

SD: Standardabweichung; SE: Standardfehler

*LOCF

° Repeated-Measurements Modell

** Der Interaktionstest wurde nur für die LOCF-Analyse d. h. nur nach 52 Wochen berechnet.

Tabelle 4-56: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Studienabbrüche aufgrund unzureichender glykämischer Kontrolle, n (%)
D1690C00004		
Nach 52 Wochen		
männlich		
Metformin+Dapagliflozin	178	0 (0,0)
Metformin+Sulfonylharnstoff	203	8 (3,9)
OR [95%-KI]		0,06 [0,00;1,12]
weiblich		
Metformin+Dapagliflozin	140	1 (0,7)
Metformin+Sulfonylharnstoff	151	6 (4,0)
OR [95%-KI]		0,17 [0,02;1,46]
Interaktionstest		I ² =0%, p=0,58
Nach 104 Wochen		
männlich		
Metformin+Dapagliflozin	178	32 (18,0)
Metformin+Sulfonylharnstoff	203	47 (23,2)
OR [95%-KI]		0,73 [0,44;1,20]
weiblich		
Metformin+Dapagliflozin	140	19 (13,6)
Metformin+Sulfonylharnstoff	151	33 (21,9)
OR [95%-KI]		0,56 [0,30;1,04]
Interaktionstest		I ² =0%, p=0,52

4.3.1.3.2.2 Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach dem Alter – RCT

Für den Endpunkt „Ausmaß der Blutzuckerkontrolle (HbA1c)“ ergab sich hinsichtlich des Alters der Patienten (<65 versus \geq 65 Jahre) kein Hinweis auf eine Effektmodifikation, der Interaktionstest war zum Niveau 0,2 nicht signifikant (Interaktionstest Alter: $I^2=0\%$; $p=0,83$). Auch für die Operationalisierung der Studienabbrüche aufgrund unzureichender glykämischer Kontrolle ergab sich kein Hinweis auf eine Effektmodifikation durch Alter ($I^2=0\%$; $p=0,62$ Interaktionstest).

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-57: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangs- wert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert % (SE)	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert % (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
<65 Jahre	257	7,74 (0,888)	7,20 (0,804)	-0,54 (0,0532)	257	7,74 (0,913)	7,25 (1,117)	-0,49 (0,0532)	-0,05 (0,0753) [-0,20;0,10]
≥65 Jahre	61	7,88 (0,784)	7,25 (0,660)	-0,58 (0,0888)	97	7,69 (0,780)	7,15 (0,921)	-0,56 (0,0703)	-0,01 (0,1136) [-0,24;0,21]
Interaktionstest**									I ² =0%;p=0,83
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen°									
<65 Jahre	218		7,14 (0,768)	-0,54 (0,0550)	211		7,14 (0,998)	-0,47 (0,0562)	-0,07 (0,0786) [-0,22;0,09]
≥65 Jahre	54		7,22 (0,652)	-0,58 (0,0918)	73		7,12 (0,797)	-0,54 (0,0756)	-0,05 (0,1195) [-0,28;0,19]
104 Wochen°									
<65 Jahre	161		7,05 (0,570)	-0,31 (0,0645)	145		7,06 (0,803)	-0,11 (0,0666)	-0,19 (0,0922) [-0,38;-0,01]
≥65 Jahre	35		7,11 (0,609)	-0,40 (0,1262)	47		7,13 (0,696)	-0,16 (0,1086)	-0,24 (0,1667) [-0,57;-0,10]

SD: Standardabweichung; SE: Standardfehler

*LOCF

° Repeated-Measurements-Modell

** Der Interaktionstest wurde nur für die LOCF-Analyse d. h. nur nach 52 Wochen berechnet

Tabelle 4-58: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Studienabbrüche aufgrund unzureichender glykämischer Kontrolle, n (%)
D1690C00004		
Nach 52 Wochen		
<65 Jahre		
Metformin+Dapagliflozin	257	1 (0,4)
Metformin+Sulfonylharnstoff	257	11 (4,3)
OR [95%-KI]		0,09 [0,01;1,68]
≥65 Jahre		
Metformin+Dapagliflozin	61	0 (0,0)
Metformin+Sulfonylharnstoff	97	3 (3,1)
OR [95%-KI]		0,22 [0,01;4,32]
Interaktionstest		I ² =0%, p=0,62
Nach 104 Wochen		
<65 Jahre		
Metformin+Dapagliflozin	257	41 (16,0)
Metformin+Sulfonylharnstoff	257	64 (24,9)
OR [95%-KI]		0,57 [0,37;0,89]
≥65 Jahre		
Metformin+Dapagliflozin	61	10 (16,4)
Metformin+Sulfonylharnstoff	97	16 (16,5)
OR [95%-KI]		0,99 [0,42;2,36]
Interaktionstest		I ² =19,4%, p=0,27

4.3.1.3.2.3 Hypoglykämien stratifiziert nach dem Geschlecht – RCT

In der Subgruppenanalyse der Hypoglykämien getrennt nach Geschlecht ergab sich anhand von Interaktionstests kein Hinweis auf eine Effektmodifikation durch das Geschlecht der Patienten. Dies galt für alle Kategorien von Hypoglykämien: Alle Interaktionstest zeigten p-Werte >0,2.

Tabelle 4-59: Ergebnisse für Hypoglykämien stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie D1690C00004		Zu bewertendes Arzneimittel N / n (%)		Vergleichstherapie N / n (%)		Odds Ratio [95%-KI]
		Metformin+ Dapagliflozin		Metformin+ Sulfonylharnstoff		
52 Wochen						
männlich	Patienten mit mindestens einer Hypoglykämie	178	6 (3,4)	206	89 (43,2)	0,05 [0,02;0,11]
	schwer		0		2 (1,0)	0,23 [0,01;4,8]
	leicht		3 (1,7)		81 (39,3)	0,03 [0,01;0,09]
	andere		3 (1,7)		20 (9,7)	0,16 [0,05;0,55]
weiblich	Bestätigte Hypoglykämien		1 (0,6)	203	26 (12,8)	0,04 [0,01;0,29]
	Patienten mit mindestens einer Hypoglykämie	140	5 (3,6)	153	61 (39,9)	0,06 [0,02;0,14]
	schwer		0		1 (0,7)	0,36 [0,01;8,95]
	leicht		2 (1,4)		56 (36,6)	0,03 [0,01;0,11]
	andere		3 (2,1)		15 (9,8)	0,20 [0,06;0,71]
	Bestätigte Hypoglykämien		2 (1,4)	151	22 (14,6)	0,08 [0,02;0,37]
Interaktions- test	Patienten mit mindestens einer Hypoglykämie					I ² =0%, p=0,76
	schwer					I ² =0%, p=0,84
	leicht					I ² =0%, p=0,96
	andere					I ² =0%, p=0,79
	Bestätigte Hypoglykämien					I ² =0%, p=0,52
104 Wochen						
männlich	Patienten mit mindestens einer Hypoglykämie	178	8 (4,5)	206	101 (49,0)	0,05 [0,02;0,10]
	schwer		0		2 (1,0)	0,23 [0,01;4,80]
	leicht		5 (2,8)		94 (45,6)	0,03 [0,01;0,09]
	andere		3 (1,7)		23 (11,2)	0,14 [0,04;0,46]
	Bestätigte Hypoglykämien		1 (0,6)	203	30 (14,8)	0,03 [0,00;0,24]

Studie D1690C00004		Zu bewertendes Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
		Metformin+ Dapagliflozin	Metformin+ Sulfonylharnstoff	
weiblich	Patienten mit mindestens einer Hypoglykämie	140 6 (4,3)	153 72 (47,1)	0,05 [0,02;0,12]
	schwer	0	1 (0,7)	0,36 [0,01;8,95]
	leicht	3 (2,1)	67 (43,8)	0,03 [0,01;0,09]
	andere	3 (2,1)	16 (10,5)	0,19 [0,05;0,66]
	Bestätigte Hypoglykämien	2 (1,4)	151 26 (17,2)	0,07 [0,02;0,30]
Interaktions- test	Patienten mit mindestens einer Hypoglykämie			I ² =0%, p=0,96
	schwer			I ² =0%, p=0,84
	leicht			I ² =0%, p=0,79
	andere			I ² =0%, p=0,72
	Bestätigte Hypoglykämien			I ² =0%, p=0,54

4.3.1.3.2.4 Hypoglykämien stratifiziert nach dem Alter – RCT

Für Hypoglykämien zeigte sich hinsichtlich des Alters der Patienten kein Hinweis auf eine relevante Interaktion. Der zum Niveau 0,2 signifikante Interaktionstest für leichte Hypoglykämien zu Woche 52 deutet auf eine qualitative Interaktion hin, die nicht als fazitrelevant gewertet wird. Zu Woche 104 ist keiner der durchgeführten Interaktionstests signifikant.

Tabelle 4-60: Ergebnisse für Hypoglykämien stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie D1690C00004		Zu bewertendes Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
		Metformin+ Dapagliflozin	Metformin+ Sulfonylharnstoff	
52 Wochen				
<65 Jahre	Patienten mit mindestens einer Hypoglykämie	257 8 (3,1)	258 110 (42,6)	0,04 [0,02;0,09]
	schwer	0	1 (0,4)	0,33 [0,01;8,22]
	leicht	3 (1,2)	102 (39,5)	0,02 [0,01;0,06]
	andere	5 (1,9)	25 (9,7)	0,18 [0,07;0,49]
	Bestätigte Hypoglykämien	3 (1,2)	257 39 (15,2)	0,07 [0,02;0,22]

Studie D1690C00004		Zu bewertendes Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
		Metformin+ Dapagliflozin	Metformin+ Sulfonylharnstoff	
≥65 Jahre	Patienten mit mindestens einer Hypoglykämie	61	101	0,08 [0,02;0,27]
	schwer	3 (4,9)	40 (39,6)	
	leicht	0	2 (2,0)	0,32 [0,02;6,85]
	andere	2 (3,3)	35 (34,7)	0,06 [0,01;0,28]
	Bestätigte Hypoglykämien	1 (1,6)	10 (9,9)	0,15 [0,02;1,22]
		0 (0,0)	97	0,08 [0,00;1,33]
Interaktions-test	Patienten mit mindestens einer Hypoglykämie			I ² =0%, p=0,41
	schwer			I ² =0%, p=0,99
	leicht			I ² =42,8%, p=0,19
	andere			I ² =0%, p=87
	Bestätigte Hypoglykämien			I ² =0%; p=0,93
104 Wochen				
<65 Jahre	Patienten mit mindestens einer Hypoglykämie	257	258	0,05 [0,02;0,09]
	schwer	11 (4,3)	127 (49,2)	
	leicht	0	1 (0,4)	0,33 [0,01;8,22]
	andere	6 (2,3)	119 (46,1)	0,03 [0,01;0,07]
	Bestätigte Hypoglykämien	5 (1,9)	26 (10,1)	0,18 [0,07;0,47]
		3 (1,2)	257	0,06 [0,02;0,19]
≥65 Jahre	Patienten mit mindestens einer Hypoglykämie	61	101	0,06 [0,02;0,21]
	schwer	3 (4,9)	46 (45,5)	
	leicht	0	2 (2,0)	0,32 [0,02;6,85]
	andere	2 (3,3)	42 (41,6)	0,05 [0,01;0,21]
	Bestätigte Hypoglykämien	1 (1,6)	13 (12,9)	0,11 [0,01;0,89]
		0 (0,0)	97	0,06 [0,00;0,96]
Interaktions-test	Patienten mit mindestens einer Hypoglykämie			I ² =0%, p=0,68
	schwer			I ² =0%, p=0,99
	leicht			I ² =0%, p=0,54
	andere			I ² =0%, p=0,70
	Bestätigte Hypoglykämien			I ² =0%, p=0,99

4.3.1.3.2.5 Gewichtsveränderung stratifiziert nach dem Geschlecht – RCT

Für den Endpunkt Gewichtsveränderung ergab sich kein Hinweis auf eine Interaktion zwischen Behandlung und Geschlecht unabhängig von der betrachteten Operationalisierung

(absolute Veränderung bzw. Anteil der Patienten mit einer Gewichtsreduktion von mindestens 5%).

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-61: Ergebnisse für Gewichtsveränderung stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangs- wert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert kg (SE)	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert kg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
männlich	178	92,09 (16,553)	88,75 (15,893)	-3,34 (0,2616)	203	92,28 (16,928)	93,93 (17,500)	1,65 (0,2450)	-4,99 (0,3584) [-5,70;-4,29]
weiblich	140	84,59 (16,240)	80,90 (16,648)	-3,70 (0,3101)	151	80,83 (15,601)	82,22 (16,148)	1,4 (0,2985)	-5,10 (0,4319) [-5,95;-4,25]
Interaktionstest**									I ² =0%;p=0,77
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen°									
männlich	157		89,00 (16,380)	-3,42 (0,2800)	171		93,58 (17,143)	1,66 (0,2671)	-5,09 (0,3869) [-5,85;-4,32]
weiblich	117		81,71 (16,170)	-3,81 (0,3352)	113		82,28 (15,990)	1,70 (0,3356)	-5,51 (0,4771) [-6,45;-4,57]
104 Wochen°									
männlich	110		89,49 (16,662)	-3,61 (0,3153)	126		94,74 (16,846)	1,51 (0,2967)	-5,11 (0,4327) [-5,96;-4,26]
weiblich	87		80,8 (15,177)	-4,17 (0,4092)	69		82,25 (15,280)	1,68 (0,4297)	-5,85 (0,5968) [-7,03;-4,67]

*LOCF

SD: Standardabweichung; SE: Standardfehler

° Repeated-Measurements-Modell

** Der Interaktionstest wurde nur für die LOCF-Analyse d. h. nur nach 52 Wochen berechnet

Tabelle 4-62: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie D1690C00004	Metformin+ Dapagliflozin N / n (%)		Metformin+ Sulfonylharnstoff N / n (%)		Odds Ratio [95%-KI]
52 Wochen					
männlich	178	54 (30,3)	203	4 (2,0)	21,67 [7,66;61,30]
weiblich	140	62 (44,3)	151	5 (3,3)	23,21 [8,96;60,12]
Interaktionstest					$I^2=0\%,p=0,92$
104 Wochen					
männlich	178	38 (21,3)	203	6 (3,0)	8,91 [3,67;21,65]
weiblich	140	43 (30,7)	151	3 (2,0)	21,87 [6,60;72,47]
Interaktionstest					$I^2=28,2\%,p=0,92$

4.3.1.3.2.6 Gewichtsveränderung stratifiziert nach dem Alter – RCT

Für den Endpunkt Gewichtsveränderung ergab sich kein Hinweis auf eine Interaktion zwischen Behandlung und Alter der Patienten unabhängig von der betrachteten Operationalisierung (absolute Veränderung bzw. Anteil der Patienten mit einer Gewichtsreduktion von mindestens 5%).

Tabelle 4-63: Ergebnisse für Gewichtsveränderung stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangs- wert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert kg (SE)	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert kg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
<65 Jahre	257	89,09 (17,415)	85,67 (17,346)	-3,43 (0,2286)	257	89,12 (17,995)	90,80 (18,568)	1,68 (0,2286)	-5,11 (0,3232) [-5,75;-4,48]
≥65 Jahre	61	87,49 (14,023)	83,74 (13,443)	-3,70 (0,4161)	97	82,82 (14,454)	83,99 (14,900)	1,13 (0,3290)	-4,83 (0,5336) [-5,88;-3,77]
Interaktionstest**									I ² =0%;p=0,65
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen°									
<65 Jahre	220		86,30 (17,392)	-3,51 (0,2464)	211		90,41 (18,388)	1,79 (0,2522)	-5,30 (0,3527) [-5,99;-4,61]
≥65 Jahre	54		84,19 (13,250)	-3,74 (0,4349)	73		85,26 (14,361)	1,34 (0,3626)	-5,08 (0,5686) [-6,21;-3,96]
104 Wochen°									
<65 Jahre	162		86,47 (17,474)	-3,75 (0,2910)	148		91,79 (18,155)	1,67 (0,3001)	-5,42 (0,4180) [-6,24;-4,59]
≥65 Jahre	35		81,89 (10,805)	-4,09 (0,4895)	47		85,70 (13,616)	1,21 (0,4187)	-5,30 (0,6460) [-6,58;-4,01]

*LOCF

SD: Standardabweichung; SE: Standardfehler

° Repeated-Measurements-Modell

** Der Interaktionstest wurde nur für die LOCF-Analyse d. h. nur nach 52 Wochen berechnet

Tabelle 4-64: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) stratifiziert nach dem Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie D1690C00004	Metformin+ Dapagliflozin N / n (%)		Metformin+ Sulfonylharnstoff N / n (%)		Odds Ratio [95%-KI]
52 Wochen					
<65 Jahre	257	92 (35,8)	257	5 (1,9)	28,10 [11,19;70,60]
≥ 65 Jahre	61	24 (39,3)	97	4 (4,1)	14,11 [4,57;43,51]
Interaktionstest					$I^2=0\%$; $p=0,35$
104 Wochen					
<65 Jahre	257	66 (25,7)	257	7 (2,7)	12,34 [5,54;27,51]
≥ 65 Jahre	61	15 (24,6)	97	2 (2,1)	14,51 [3,18;66,20]
Interaktionstest					$I^2=0\%$; $p=0,85$

4.3.1.3.2.7 Therapiezufriedenheit stratifiziert nach dem Geschlecht – RCT

Hinsichtlich der Therapiezufriedenheit ergab sich ein Hinweis auf eine Effektmodifikation durch das Geschlecht der Patienten. Die Interaktion war quantitativer Art, sowohl für Männer als auch für Frauen war der Behandlungseffekt nicht signifikant. Somit ergibt sich keine Änderung des für die gesamte Zielpopulation gezogenen Fazits.

Tabelle 4-65: Ergebnisse für Therapiezufriedenheit (DTSQ) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppen- unterschied Mittelwert (SE)
	N	Ausgangswert DTSQs Mittelwert (SD)	Endwert DTSQc Mittelwert* (SE)	N	Ausgangswert DTSQs Mittelwert (SD)	Endwert DTSQc Mittelwert (SE)	
52 Wochen	Metformin+Dapagliflozin			Metformin+Sulfonylharnstoff			
männlich	152	30,3 (5,62)	13,8 (0,411)	164	31,3 (5,52)	13,8 (0,396)	0,0 (0,572) [-1,1;1,2]
weiblich	114	31,8 (5,22)	14,5 (0,498)	110	31,7 (4,61)	13,3 (0,507)	1,2 (0,710) [-0,2;2,6]
Inter- aktionstest							$I^2=42,3\%$; $p=0,19$

DTSQ: Disease Treatment Satisfaction Questionnaire

SD: Standardabweichung; SE: Standardfehler

4.3.1.3.2.8 Therapiezufriedenheit stratifiziert nach dem Alter – RCT

Für die Therapiezufriedenheit ergab sich kein Hinweis auf eine signifikante Variation des Behandlungsunterschiedes für ältere und jüngere Patienten.

Tabelle 4-66: Ergebnisse für Therapiezufriedenheit (DTSQ) stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppen- unterschied Mittelwert (SE)
	N	Ausgangswert DTSQs Mittelwert (SD)	Endwert DTSQc Mittelwert (SE)	N	Ausgangswert DTSQs Mittelwert (SD)	Endwert DTSQc Mittelwert (SE)	
52 Wochen	Metformin+Dapagliflozin			Metformin+Sulfonylharnstoff			
<65 Jahre	212	31,1 (5,20)	14,1 (0,350)	201	31,2 (5,29)	13,6 (0,360)	0,4 (0,502) [-0,6;1,4]
≥65 Jahre	54	30,3 (6,51)	14,2 (0,725)	73	32,1 (4,78)	13,6 (0,623)	0,6 (0,962) [-1,4;2,5]
Inter- aktionstest							I ² =0%, p=0,93

DTSQ: Disease Treatment Satisfaction Questionnaire

SD: Standardabweichung; SE: Standardfehler

4.3.1.3.2.9 Blutdrucksenkung stratifiziert nach dem Geschlecht – RCT

Hinsichtlich der Blutdrucksenkung ergab sich anhand von Interaktionstests kein Hinweis auf eine Effektmodifikation durch das Geschlecht der Patienten.

Tabelle 4-67: Ergebnisse für Blutdrucksenkung stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert mmHg (SE) [95%-KI]
	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
männlich systolisch	178	135,0 (14,49)	130,1 (12,56)	-4,5 (0,847)	203	133,3 (13,70)	135,2 (12,99)	1,4 (0,793)	-5,9 (1,162) [-8,2;-3,6]
diastolisch	178	81,9 (8,96)	80,7 (8,85)	-0,9 (0,569)	203	80,8 (8,34)	80,7 (8,77)	-0,3 (0,532)	-0,6 (0,780) [-2,1;1,0]
weiblich systolisch	140	129,4 (13,72)	126,0 (13,51)	-4,1 (1,002)	148	132,6 (14,13)	132,0 (14,45)	0,1 (0,974)	-4,2 (1,401) [-7,0;-1,5]
diastolisch	140	80,0 (7,89)	77,9 (8,56)	-2,4 (0,601)	148	80,9 (8,21)	80,0 (8,40)	-0,7 (0,584)	-1,7 (0,838) [-3,3;-0,0]
Interaktions- test**									I ² =0%, p=0,35 I ² =0%, p=0,34
52 Wochen°									
männlich systolisch	157		130,73 (12,689)	-3,99 (0,9008)	171		135,08 (13,169)	1,43 (0,8603)	-5,41 (1,2471) [-7,86;-2,96]
diastolisch	157		80,97 (8,822)	-0,59 (0,5993)	171		80,59 (8,831)	-0,30 (0,5718)	-0,29 (0,8293) [-1,92;1,34]

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert mmHg (SE) [95%-KI]
	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
weiblich systolisch	117		127,13 (12,847)	-3,81 (1,0450)	112		132,21 (14,064)	0,33 (1,0625)	-4,14 (1,4925) [-7,08;-1,20]
diastolisch	117		78,47 (8,393)	-2,18 (0,6389)	112		80,01 (8,275)	-0,71 (0,6476)	-1,47 (0,9099) [-3,26;0,32]
104 Wochen°									
männlich systolisch	110		130,17 (12,093)	-4,40 (1,0031)	126		135,71 (12,766)	1,57 (0,9386)	-5,97 (1,3745) [-8,68;-3,26]
diastolisch	110		79,51 (8,232)	-2,17 (0,6461)	126		79,97 (8,419)	-1,09 (0,6046)	-1,08 (0,8859) [-2,83;0,66]
weiblich systolisch	87		128,89 (14,285)	-1,54 (1,3003)	69		133,30 (13,639)	0,67 (1,4346)	-2,21 (1,9404) [-6,03;1,62]
diastolisch	87		78,38 (8,319)	-2,23 (0,7490)	69		78,97 (7,723)	-1,74 (0,8233)	-0,49 (1,1138) [-2,69;1,70]

* LOCF

SD: Standardabweichung; SE: Standardfehler

° Repeated-Measurements-Modell

** Der Interaktionstest wurde nur für die LOCF-Analyse d. h. nur nach 52 Wochen berechnet

4.3.1.3.2.10 Blutdrucksenkung stratifiziert nach dem Alter – RCT

Hinsichtlich der Blutdrucksenkung ergab sich anhand von Interaktionstests ein Hinweis auf eine Effektmodifikation durch das Alter der Patienten. Der Behandlungseffekt war in der Subgruppe der älteren Patienten (≥ 65 Jahre) größer als bei jüngeren Patienten (< 65 Jahre). Da die Interaktion für den systolischen Blutdruck nicht qualitativer sondern quantitativer Art war, wird sie als nicht fazitrelevant gesehen.

Tabelle 4-68: Ergebnisse für Blutdrucksenkung stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert mmHg (SE) [95%-KI]
	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
52 Wochen*									
<65 Jahre systolisch	257	131,5 (14,44)	128,3 (13,60)	-3,4 (0,725)	255	132,5 (14,07)	132,9 (13,90)	0,7 (0,728)	-4,1 (1,027) [-6,1;-2,1]
diastolisch	257	81,0 (8,96)	79,8 (9,02)	-1,4 (0,471)	255	81,7 (8,02)	80,9 (8,62)	-0,7 (0,473)	-0,7 (0,667) [-2,0;0,6]
≥65 Jahre systolisch	61	137,2 (13,42)	128,3 (10,95)	-7,8 (1,435)	96	134,4 (13,30)	136,3 (12,91)	1,3 (1,143)	-9,1 (1,839) [-12,7;-5,5]
diastolisch	61	81,2 (6,56)	77,9 (7,82)	-2,6 (0,885)	96	78,6 (8,55)	79,2 (8,50)	0,2 (0,703)	-2,8 (1,137) [-5,1;-0,6]
Interaktions- test**									I ² =82,3%;p=0,02 I ² =60,9%;p=0,11
52 Wochen°									
<65 Jahre systolisch	220		129,32 (13,394)	-3,03 (0,7659)	210		132,88 (13,550)	0,73 (0,7833)	-3,75 (1,0956) [-5,91;-1,60]
diastolisch	220		80,38 (8,861)	-1,06 (0,4982)	210		80,73 (8,527)	-0,66 (0,5092)	-0,40 (0,7124) [-1,80;1,00]

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert mmHg (SE) [95%-KI]
	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert mmHg (SE)	
	Metformin+Dapagliflozin				Metformin+Sulfonylharnstoff				
≥65 Jahre systolisch	54		128,70 (10,485)	-7,29 (1,4871)	73		137,00 (13,281)	1,69 (1,2668)	-8,98 (1,9561) [-12,85;-5,11]
diastolisch	54		77,97 (7,870)	-2,47 (0,9352)	73		79,30 (8,800)	0,27 (0,7871)	-2,74 (1,2299) [-5,17;-0,30]
104 Wochen°									
<65 Jahre systolisch	162		129,14 (12,538)	-3,05 (0,8665)	148		133,96 (13,550)	0,81 (0,9026)	-3,86 (1,2513) [-6,32;-1,40]
diastolisch	162		79,36 (8,176)	-2,09 (0,5253)	148		80,70 (7,851)	-0,98 (0,5461)	-1,11 (0,7580) [-2,60;0,38]
≥65 Jahre systolisch	35		131,74 (15,401)	-3,43 (2,0452)	47		137,70 (11,222)	1,70 (1,7471)	-5,13 (2,6912) [-10,48;0,22]
diastolisch	35		77,42 (8,628)	-2,54 (1,2529)	47		76,21 (8,316)	-2,58 (1,0788)	0,04 (1,6588) [-3,25;3,34]

SD: Standardabweichung; SE: Standardfehler

* LOCF

° Repeated-Measurements-Modell

** Der Interaktionstest wurde nur für die LOCF-Analyse d. h. nur nach 52 Wochen berechnet

4.3.1.3.2.11 Unerwünschte Ereignisse stratifiziert nach dem Geschlecht – RCT

Auch für die Gesamtrate unerwünschter Ereignisse und für die Gesamtrate schwerwiegender unerwünschter Ereignisse zeigten sich keine Hinweise auf relevante Unterschiede zwischen den Subgruppen hinsichtlich des Geschlechts (im Vergleich Metformin+Dapagliflozin gegenüber Metformin+Sulfonylharnstoff). Für Studienabbrüche aufgrund unerwünschter Ereignisse gab es eine signifikante Interaktion, aber bei den geringen Ereigniszahlen kann man nicht von einem validen Ergebnis sprechen. Hinsichtlich der Genital- und Harnwegsinfektionen zeigten sich keine signifikanten Interaktionen, aber generell war die Anzahl der Infektionen bei den Frauen häufiger.

Tabelle 4-69: Ergebnisse für „Unerwünschte Ereignisse“ stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie		N	Gesamtrate unerwünschter Ereignisse, n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
D1690C0 0004					
männlich	Nach 52 Wochen				
	Metformin+Dapagliflozin	178	138 (77,5)	13 (7,3)	9 (5,1)
weiblich	Metformin+Sulfonylharnstoff	206	162 (78,6)	23 (11,2)	8 (3,9)
	OR [95%-KI]		0,94 [0,58;1,52]	0,63 [0,31;1,28]	1,32 [0,50;3,49]
männlich	Metformin+Dapagliflozin	140	117 (83,6)	10 (7,1)	5 (3,6)
	Metformin+Sulfonylharnstoff	153	121 (79,1)	16 (10,5)	9 (5,9)
	OR [95%-KI]		1,35 [0,74;2,43]	0,66 [0,29;1,50]	0,59 [0,19;1,81]
Interaktionstest			I ² =0%; p=0,35	I ² =0%; p=0,93	I ² =10,4%;p=0,93
männlich	Nach 104 Wochen				
	Metformin+Dapagliflozin	178	151 (84,8)	21 (11,8)	10 (5,6)
weiblich	Metformin+Sulfonylharnstoff	206	169 (82,0)	29 (14,1)	8 (3,9)
	OR [95%-KI]		1,22 [0,71;2,11]	0,82 [0,45;1,49]	1,47 [0,57;3,82]
männlich	Metformin+Dapagliflozin	140	120 (85,7)	15 (10,7)	6 (4,3)
	Metformin+Sulfonylharnstoff	153	128 (83,7)	22 (14,4)	13 (8,5)
	OR [95%-KI]		1,17 [0,62;2,22]	0,71 [0,35;1,44]	0,48 [0,18;1,31]
Interaktionstest			I ² =0%; p=0,92	I ² =0%; p=0,78	I ² =60,4%;p=0,11

Tabelle 4-70: Ergebnisse für Unerwünschte Ereignisse – Anzeichen für Genital- und Harnwegsinfektionen stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Anzahl der Patienten mit Anzeichen für eine Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen für eine Harnwegsinfektion, n (%)
D1690C00004			
Nach 52 Wochen			
männlich			
Metformin+Dapagliflozin	178	7 (3,9)	14 (7,9)
Metformin+Sulfonylharnstoff	206	1 (0,5)	8 (3,9)
OR [95%-KI]		8,39 [1,02;68,88]	2,11 [0,87;5,16]
weiblich			
Metformin+Dapagliflozin	140	35 (25,0)	23 (16,4)
Metformin+Sulfonylharnstoff	153	8 (5,2)	14 (9,2)
OR [95%-KI]		6,04 [2,96;13,56]	1,95 [0,96;2,96]
Interaktionstest		I ² =0%; p=0,78	I ² =0%; p=0,89
Nach 104 Wochen			
männlich			
Metformin+Dapagliflozin	178	12 (6,7)	19 (10,7)
Metformin+Sulfonylharnstoff	206	1 (0,5)	10 (4,9)
OR [95%-KI]		14,82 [1,91;115,14]	2,34 [1,06;5,18]
weiblich			
Metformin+Dapagliflozin	140	39 (27,9)	29 (20,7)
Metformin+Sulfonylharnstoff	153	9 (5,9)	23 (15,0)
OR [95%-KI]		6,18 [2,87;13,32]	1,48 [0,81;2,70]
Interaktionstest		I ² =0%; p=0,43	I ² =0%; p=0,36

4.3.1.3.2.12 Unerwünschte Ereignisse stratifiziert nach dem Alter – RCT

Bei der Gesamtrate unerwünschter Ereignisse zeigte sich nach 52 Wochen eine signifikante Interaktion bezüglich des Alters, jedoch änderte sich nichts an der Richtung des Effekts und die Interaktion war nach 104 Wochen nicht mehr vorhanden.

Tabelle 4-71: Ergebnisse für Unerwünschte Ereignisse stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie		N	Gesamtrate unerwünschter Ereignisse, n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
D1690C 00004					
<65 Jahre	Nach 52 Wochen				
	Metformin+Dapagliflozin	257	208 (80,9)	17 (6,6)	9 (3,5)
	Metformin+Sulfonylharnstoff OR [95%-KI]	258	196 (76,0) 1,34 [0,88;2,05]	20 (7,8) 0,84 [0,43;1,65]	6 (2,3) 1,52 [0,53;4,35]
≥65 Jahre	Metformin+Dapagliflozin	61	47 (77,0)	6 (9,8)	5 (8,2)
	Metformin+Sulfonylharnstoff	101	87 (86,1)	19 (18,8)	11 (10,9)
	OR [95%-KI]		0,54 [0,24;1,23]	0,47 [0,18;1,25]	0,73 [0,24;2,21]
Inter-aktions-test			I ² =73,2%,p=0,05	I ² =0%,p=0,34	I ² =0%,p=0,34
<65 Jahre	Nach 104 Wochen				
	Metformin+Dapagliflozin	257	218 (84,8)	25 (9,7)	11 (4,3)
	Metformin+Sulfonylharnstoff OR [95%-KI]	258	207 (80,2) 1,38 [0,87;2,18]	26 (10,1) 0,96 [0,54;1,71]	8 (3,1) 1,40 [0,55;3,53]
≥65 Jahre	Metformin+Dapagliflozin	61	53 (86,9)	11 (18,0)	5 (8,2)
	Metformin+Sulfonylharnstoff	101	90 (89,1)	25 (24,8)	13 (12,9)
	OR [95%-KI]		0,81 [0,31;2,14]	0,67 [0,30;1,48]	0,60 [0,20;1,79]
Inter-aktions-test			I ² =0%,p=0,33	I ² =0%,p=0,47	I ² =24,5%,p=0,25

Tabelle 4-72: Ergebnisse für Unerwünschte Ereignisse – Anzeichen für Genital- und Harnwegsinfektionen stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Anzahl der Patienten mit Anzeichen für eine Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen für eine Harnwegsinfektion, n (%)
D1690C00004			
Nach 52 Wochen			
<65 Jahre			
Metformin+Dapagliflozin	257	36 (14,0)	29 (11,3)
Metformin+Sulfonylharnstoff	258	8 (3,1)	14 (5,4)
OR [95%-KI]		5,09 [2,32;11,18]	2,22 [1,14;4,30]
≥65 Jahre			
Metformin+Dapagliflozin	61	6 (9,8)	8 (13,1)
Metformin+Sulfonylharnstoff	101	1 (1,0)	8 (7,9)
OR [95%-KI]		10,91 [1,28;92,95]	1,75 [0,62;4,95]
Interaktionstest		I ² =0%; p=0,51	I ² =0%; p=0,71
Nach 104 Wochen			
<65 Jahre			
Metformin+Dapagliflozin	257	43 (16,7)	40 (15,6)
Metformin+Sulfonylharnstoff	258	9 (3,5)	21 (8,1)
OR [95%-KI]		5,56 [2,65;11,67]	2,08 [1,19;3,64]
≥65 Jahre			
Metformin+Dapagliflozin	61	8 (13,1)	8 (13,1)
Metformin+Sulfonylharnstoff	101	1 (1,0)	12 (11,9)
OR [95%-KI]		15,09 [1,84;123,93]	1,12 [0,43;2,92]
Interaktionstest		I ² =0%; p=0,38	I ² =16,7%; p=0,27

4.3.1.3.2.13 Kardiale Ereignisse stratifiziert nach dem Geschlecht – RCT

Hinsichtlich der Inzidenz kardialer Ereignisse ergab sich anhand von Interaktionstests kein Hinweis auf eine Effektmodifikation durch das Geschlecht der Patienten, weder nach 52 noch nach 104 Wochen.

Tabelle 4-73: Ergebnisse für Kardiale Ereignisse stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie D1690C00004	N	Gesamtrate kardiale Ereignisse, n (%)
Nach 52 Wochen		
männlich		
Metformin+Dapagliflozin	178	13 (7,3)
Metformin+Sulfonylharnstoff	206	15 (7,3)
OR [95%-KI]		1,00 [0,46;2,17]
weiblich		
Metformin+Dapagliflozin	140	11 (7,9)
Metformin+Sulfonylharnstoff	153	11 (7,2)
OR [95%-KI]		1,10 [0,46;2,63]
Interaktionstest		I ² =0%; p=0,88
Nach 104 Wochen		
männlich		
Metformin+Dapagliflozin	178	17 (9,6)
Metformin+Sulfonylharnstoff	206	19 (9,2)
OR [95%-KI]		1,04 [0,52;2,07]
weiblich		
Metformin+Dapagliflozin	140	14 (10,0)
Metformin+Sulfonylharnstoff	153	17 (11,1)
OR [95%-KI]		0,89 [0,42;1,88]
Interaktionstest		I ² =0%; p=0,76

4.3.1.3.2.14 Kardiale Ereignisse stratifiziert nach dem Alter – RCT

Hinsichtlich kardialer Ereignisse ergab sich im Interaktionstest ein Hinweis auf eine signifikante Interaktion durch das Alter der Patienten. Allerdings war für beide Altersgruppen der Unterschied in der Häufigkeit kardialer Ereignisse zwischen den Behandlungsgruppen nicht signifikant, so dass der Hinweis als nicht fazitrelevant angesehen wird.

Tabelle 4-74: Ergebnisse für Kardiale Ereignisse stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie D1690C00004	N	Gesamtrate kardiale Ereignisse, n (%)
Nach 52 Wochen		
<65 Jahre		
Metformin+Dapagliflozin	257	18 (7,0)
Metformin+Sulfonylharnstoff	258	10 (3,9)
OR [95%-KI]		1,87 [0,84;4,13]
≥65 Jahre		
Metformin+Dapagliflozin	61	6 (9,8)
Metformin+Sulfonylharnstoff	101	16 (15,8)
OR [95%-KI]		0,58 [0,21;1,57]
Interaktionstest		I ² =69,1%; p=0,07
Nach 104 Wochen		
<65 Jahre		
Metformin+Dapagliflozin	257	25 (9,7)
Metformin+Sulfonylharnstoff	258	17 (6,6)
OR [95%-KI]		1,53 [0,80;2,90]
≥65 Jahre		
Metformin+Dapagliflozin	61	6 (9,8)
Metformin+Sulfonylharnstoff	101	19 (18,8)
OR [95%-KI]		0,47 [0,18;1,25]
Interaktionstest		I ² =74,2%; p=0,05

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen aus randomisierten kontrollierten Studien geben. Die Zusammenfassung soll Aussagen zu allen in Abschnitt 4.3.1.3 präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus randomisierten kontrollierten Studien zusammen.

Die in die Nutzenbewertung eingeschlossene RCT D1690C00004 war eine Zulassungsstudie für Dapagliflozin. Im Rahmen des Zulassungsprozesses ergab sich eine Zielpopulation für Dapagliflozin, die Patienten unter 75 Jahren mit einer eGFR ≥ 60 mL/min/1,73 m² bzw. Kreatinin-clearance ≥ 60 mL/min und nicht mit Schleifendiuretika behandelt umfasst. Die Population der Patienten, für die Dapagliflozin zugelassen ist, ist die für die Nutzenbewertung primäre Population. Die Resultate dieser Population werden im Folgenden zusammengefasst. Die Resultate sind konsistent mit denen der Studienpopulation.

Vermeidung von Hypoglykämien bei ähnlich effektiver Blutzuckerkontrolle

Bezüglich des **Ausmaßes der Blutzuckerkontrolle** ergab sich **nach 52 Wochen** kein statistisch signifikanter Unterschied zwischen der Metformin+Dapagliflozin-Gruppe und der Metformin+Sulfonylharnstoff-Gruppe. Ausgehend von mittleren Ausgangswerten von 7,77% bzw. 7,72% im HbA1c in der Metformin+Dapagliflozin- bzw. der Metformin+Sulfonylharnstoff-Gruppe wurde der HbA1c in der Metformin+Dapagliflozin-Gruppe um durchschnittlich 0,55% gesenkt, in der Metformin+Sulfonylharnstoff-Gruppe betrug die durchschnittliche HbA1c-Senkung 0,51%. Somit war in beiden Behandlungsgruppen eine ähnliche, im Ausmaß klinisch relevante HbA1c-Senkung (Reduktion des HbA1c >0,5%) zu beobachten.

Die HbA1c-Senkung unter Metformin+Dapagliflozin blieb über die Zeit erhalten, während die glykämische Kontrolle unter Metformin+Sulfonylharnstoff nachließ, sodass sich **nach 104 Wochen** (*Repeated-Measurements-Analyse*) eine statistisch signifikant größere HbA1c-Senkung in der Metformin+Dapagliflozin-Gruppe feststellen ließ (MWD [95%-KI]: -0,19% [-0,34;-0,03]). Sowohl nach 52 als auch nach 104 Wochen brachen in der Metformin+Dapagliflozin-Gruppe statistisch signifikant weniger Patienten die Studie wegen unzureichender Blutzuckerkontrolle ab (OR [95%-KI]: 0,08 [0,01;0,59] / 0,65 [0,44;0,97]).

Hinsichtlich der in Subgruppenanalysen untersuchten Faktoren Alter und Geschlecht ergaben sich keine Hinweise auf Effektmodifikation, alle Interaktionstests zeigten p-Werte >0,2 (Interaktionstest für HbA1c nach 52 Wochen Alter: $I^2=0\%$; $p=0,83$ und Geschlecht: $I^2=0\%$; $p=0,52$).

Die Resultate bzgl. des Ausmaßes der Blutzuckerkontrolle wurden zur Interpretation des patientenrelevanten Endpunktes Hypoglykämien herangezogen.

Hinsichtlich der **Anzahl der Patienten mit mindestens einer Hypoglykämie** ergab sich nach 52 Wochen ein statistisch signifikanter Vorteil von Metformin+Dapagliflozin gegenüber Metformin+Sulfonylharnstoff. Nur 11/318 Patienten (3,4%) in der Metformin+Dapagliflozin-Gruppe hatten im Beobachtungszeitraum von 52 Wochen mindestens eine Hypoglykämie; in der Metformin+Sulfonylharnstoff-Gruppe waren es 150/354 (42,1%) (OR [95%-KI]: 0,05 [0,03;0,09]). Schwere Hypoglykämien traten insgesamt bei 3 Patienten auf, alle in den ersten 52 Wochen und alle unter Metformin+Sulfonylharnstoff (0% unter Metformin+Dapagliflozin versus 0,8% unter Metformin+Sulfonylharnstoff, OR [95%-KI]: 0,16 [0,01;3,16]). Auch gab es in der Metformin+Dapagliflozin-Gruppe gegenüber der Metformin+Sulfonylharnstoff-Gruppe signifikant weniger **bestätigte Hypoglykämien** sowohl nach 52 als auch nach 104 Wochen (OR[95%-KI]: 0,06 [0,02;0,20] nach 52 Wochen und 0,05 [0,02;0,16] nach 104 Wochen).

Bezüglich der für die Zielpopulation untersuchten Anzahl der Studienabbrüche aufgrund von Hypoglykämien ergaben sich in der Metformin+Sulfonylharnstoff-Gruppe insgesamt 5 (1,4%) Studienabbrüche aufgrund von Hypoglykämien innerhalb der ersten 52 Wochen bzw.

6 (1,7%) nach 104 Wochen, wohingegen es in der Metformin+Dapagliflozin-Gruppe keine gab.

Auch für Hypoglykämien ergab sich kein Hinweis auf eine Effektmodifikation durch Alter oder Geschlecht.

Aus der gemeinsamen Betrachtung von Blutzuckerkontrolle und Anzahl der Hypoglykämien ergibt sich somit eine Überlegenheit für Metformin+Dapagliflozin gegenüber Metformin+Sulfonylharnstoff.

Signifikante Gewichtsreduktion mit Metformin+Dapagliflozin gegenüber Metformin+Sulfonylharnstoff

Patienten in der Metformin+Dapagliflozin-Gruppe erzielten nach 52 Wochen eine Gewichtsreduktion von im Mittel -3,49 kg, während Patienten unter Metformin+Sulfonylharnstoff im Mittel 1,54 kg an Körpergewicht zunahm. Der statistisch signifikante Behandlungsunterschied von -5,02 kg (95%-KI: [-5,57; -4,48]) ist für Patienten im Indikationsgebiet klinisch relevant. Die klinische Relevanz des beobachteten Unterschieds wurde auch in einer präspezifizierten Responder-Analyse gezeigt, in der Response definiert wurde als eine Gewichtsreduktion von zumindest 5% gegenüber dem Ausgangsgewicht. Mit Metformin+Dapagliflozin erreichten 36,6% der Patienten eine klinisch relevante Gewichtsreduktion von >5% gegenüber 2,6% in der Metformin+Sulfonylharnstoff-Gruppe (OR 22,01 [10,93;44,33]).

Der Effekt von Metformin+Dapagliflozin im Hinblick auf die Gewichtsreduktion hielt auch nach 104 Wochen an. Der mittlere Unterschied zwischen Patienten in der Metformin+Dapagliflozin-Gruppe und Patienten, die mit Metformin+Sulfonylharnstoff behandelt wurden lag bei -5,38 kg (95%-KI: [-6,07; -4,68]). Der Anteil der Patienten mit einer Gewichtsreduktion von >5% nach 104 Wochen lag bei 25,5% in der Metformin+Dapagliflozin-Gruppe, während nur 2,6% der Patienten in der Metformin+Sulfonylharnstoff-Gruppe eine solche Gewichtsreduktion erreichten (OR [95%-KI]: 13,10 [6,45;26,60]).

Für den Endpunkt „Gewichtsveränderungen“ ergab sich kein Hinweis auf eine Effektmodifikation durch Alter oder Geschlecht.

Kein Unterschied in der Therapiezufriedenheit

Sowohl unter Metformin+Dapagliflozin als auch unter Metformin+Sulfonylharnstoff zeigten sich die Patienten zufrieden mit der Behandlung. Ausgehend von ähnlich hohen Ausgangswerten im DTSQs gaben Patienten unter Metformin+Dapagliflozin im Mittel einen DTSQc Skalenwert von 14,1 (SE: 0,317) an, während Patienten unter Metformin+Sulfonylharnstoff durchschnittlich einen Skalenwert von 13,6 Punkten (SE: 0,312) auf der von -18 bis +18 Punkte umfassenden DTSQc Skala berichteten. Die DTSQc misst relative Veränderung in der Therapiezufriedenheit. Positive Werte sind als Verbesserung zu interpretieren, der Wert 0 auf der Skala bedeutet keine Veränderung.

Zwischen den Gruppen war kein statistisch signifikanter Unterschied vorhanden (MWD [95%-KI]: 0,5 [-0,3;1,4]).

Für den Endpunkt „Therapiezufriedenheit“ ergaben sich keine fazitrelevanten Hinweise auf eine Effektmodifikation durch Alter oder Geschlecht.

Statistisch signifikante Reduktion des systolischen Blutdrucks um 5 mmHg nach 52 Wochen

Mit Metformin+Dapagliflozin konnten die Patienten im Mittel eine Reduktion ihres systolischen Blutdrucks von -4,4 mmHg (SE: 0,649) erreichen. Patienten in der Metformin+Sulfonylharnstoff-Gruppe wiesen nach 52 Wochen nur geringe Veränderungen ihres systolischen Blutdrucks auf (+1,0 mmHg (SE: 0,618)). Der Unterschied zwischen den beiden Behandlungsgruppen von -5,3 mmHg war statistisch signifikant (95%-KI: [-7,1;-3,6]) und hielt auch nach 104 Wochen an (MWD [95%-KI]: -4,46 mmHg; [-6,67;-2,26]). Der Unterschied in der Reduktion des diastolischen Blutdrucks mit Metformin+Dapagliflozin gegenüber Metformin+Sulfonylharnstoff wies ebenfalls diesen Trend auf, war aber im Ausmaß geringer und erreichte keine statistische Signifikanz. Hinsichtlich der Reduktion des systolischen Blutdrucks ergab sich ein Hinweis auf eine Effektmodifikation durch das Alter der Patienten: bei älteren Patienten (≥ 65 Jahre) war eine größere Blutdrucksenkung mit Metformin+Dapagliflozin vorhanden als bei Patienten unter 65 Jahren (MWD [95%-KI]: -9,1 mmHg [-12,7;-5,5] gegenüber -4,1 mmHg [-6,1;-2,1]) (Interaktionstest: $p=0,02$; $I^2=82,3\%$). Allerdings war der Effekt über alle Altersgruppen hinweg statistisch signifikant zugunsten von Metformin+Dapagliflozin. Somit wurde die Interaktion nicht als fazitrelevant erachtet.

Keine statistisch signifikanten Unterschiede für unerwünschte Ereignisse

Die Gesamtrate der Patienten, die ein unerwünschtes (80,2% versus 78,8%) oder ein schwerwiegendes unerwünschtes (7,2% versus 10,9%) Ereignis hatten, sowie der Anteil der Patienten, die aufgrund unerwünschter Ereignisse die Studie abbrachen (4,4% versus 4,7%) war in beiden Gruppen ähnlich. Die Gruppenunterschiede nach Woche 52 waren nicht statistisch signifikant: OR [95%-KI]: 1,09 [0,75;1,58], OR [95%-KI]: 0,64 [0,37;1,10], OR [95%-KI]: 0,93 [0,45;1,91]. Auch nach 104 Wochen ergaben sich keine signifikanten Unterschiede zwischen den Behandlungsoptionen bei der Gesamtrate unerwünschter Ereignisse (OR [95%-KI]: 1,20 [0,80;1,82]), bei der Gesamtrate schwerwiegender unerwünschter Ereignisse (OR [95%-KI]: 0,77 [0,49;1,22]) und bei den Studienabbrüchen aufgrund unerwünschter Ereignisse (OR [95%-KI]: 0,85 [0,44;1,66]).

Dabei gilt zu erwähnen, dass unter Metformin+Dapagliflozin Anzeichen für Genitalinfektionen (13,2% versus 2,5%) sowie Anzeichen für Harnwegsinfektionen (11,6% versus 6,1%) häufiger auftraten als bei Patienten, die mit Metformin+Sulfonylharnstoff behandelt wurden. Der Unterschied in den Inzidenzen war sowohl nach 52 Wochen (OR [95%-KI]: 5,92 [2,83;12,37] bzw. 2,02 [1,16;3,50]) als auch nach 104 Wochen (OR [95%-KI]: 6,67 [3,32;13,37] bzw. 1,76 [1,10;2,81]) statistisch signifikant. Die aufgetretenen

Anzeichen für Genital- und Harnwegsinfektionen waren allerdings meist nur mild oder moderat in ihrer Intensität und mit den gängigen Behandlungsmethoden einfach zu behandeln.

Da sich die Gesamtrate der unerwünschten Ereignisse in den beiden Behandlungsgruppen nicht voneinander unterschied, ist nicht von einem Schaden auszugehen, da Patienten unter einer Behandlung mit Metformin+Sulfonylharnstoff andere unerwünschte Ereignisse häufiger erleiden als Patienten unter Metformin+Dapagliflozin.

Für den Endpunkt „unerwünschte Ereignisse“ ergaben sich keine fazitrelevanten Hinweis auf eine Effektmodifikation durch Alter oder Geschlecht.

Kardiale Ereignisse: kein statistisch signifikanter Unterschied zwischen Metformin+Dapagliflozin und Metformin+Sulfonylharnstoff.

Kardiale Ereignisse traten über den Zeitraum von 52 Wochen in beiden Behandlungsgruppen in vergleichbarer Häufigkeit auf (7,5% versus 7,2%). Nach 104 Wochen hatten in der Metformin+Dapagliflozin-Gruppe 31 (9,7%) Patienten ein kardiales Ereignis, wohingegen in der Metformin+Sulfonylharnstoff-Gruppe 36 (10,0%) Patienten ein solches Ereignis hatten. Signifikante Unterschiede zwischen den Behandlungsoptionen fanden sich nicht.

Sowohl nach 52 Wochen als auch nach 104 Wochen zeigten sich Hinweise auf eine Interaktion durch das Alter der Patienten ($I^2=69,1\%$; $p=0,07$ bzw. $I^2=74,2\%$; $p=0,05$). Nach 104 Wochen lag die Häufigkeit kardialer Ereignisse bei den jüngeren Patienten (<65 Jahre) bei 9,7% versus 6,6% in der Metformin+Dapagliflozin bzw. Metformin+Sulfonylharnstoff Gruppe. In der Subgruppe der älteren Patienten erlebten 9,8% versus 18,8% der Patienten kardiale Ereignisse. Keiner der beobachteten Unterschiede war statistisch signifikant.

Zusammenfassend gelten die Aussagen, die für die gesamte Zielpopulation getroffen wurden, auch für die untersuchten Subgruppen bezüglich Alter und Geschlecht. Für einzelne Endpunkte ergaben sich Hinweise auf eine Interaktion durch das Alter der Patienten. Diese Interaktionen wurden in der Regel als nicht fazitrelevant eingestuft. Signifikante Unterschiede waren über Subgruppen hinweg signifikant bzw. nicht signifikante blieben nicht signifikant – in der Regel variiert lediglich die Größe des Effekts.

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Netzwerk-Meta-Analyse (MTC) für Dapagliflozin in der Add-on Kombinationstherapie mit Metformin

In der MTC werden die wichtigsten Wirksamkeits- und Sicherheitsendpunkte (HbA1c, Gewicht, systolischer Blutdruck und Hypoglykämien) der klinischen Studien mit Dapagliflozin in Bezug gesetzt zu denen anderer antidiabetischer Arzneimittel, die zur Behandlung des Diabetes mellitus Typ 2 eingesetzt werden. Ziel dieser Analyse war es, die Wirksamkeit und Sicherheit von Metformin+Dapagliflozin gegenüber anderen mit Metformin kombinierten Medikamenten zu untersuchen.

Die Analyse enthält zwei Komponenten: einen systematischen Review der dazu verfügbaren Literatur (*Systematic Review of Antidiabetic Agents in Type 2 Diabetes mellitus: Metformin add-on therapy* [(Oxford Outcomes Ltd., 2011b)]) sowie eine Bayesianische MTC der antidiabetischen Arzneimittel (*Network Meta-Analysis of Anti Diabetic Agents in Type 2 Diabetes Mellitus: Metformin add on therapy* (Oxford Outcomes Ltd., 2011a; Oxford Outcomes Ltd., 2012b)). Die Literaturrecherchen und Registersuchen des systematischen Reviews wurde zwei Mal aktualisiert (Oxford Outcomes Ltd., 2012a; Oxford Outcomes Ltd., 2012c).

Im systematischen Review wurden mittels Literaturrecherche und der Suche in den Studienregistern RCT identifiziert, die die Kombination von Metformin mit anderen antidiabetischen Arzneimitteln bei Patienten, die zuvor keine ausreichende Blutzuckerkontrolle mit Metformin-Monotherapie hatten, untersuchen.

Im Bericht zur MTC werden zwei Netzwerke beschrieben, die sich hinsichtlich ihrer Studiendauer unterscheiden. Im Netzwerk A sind alle relevanten RCT mit einer Studiendauer von 24±6 Wochen, im Netzwerk B alle relevanten RCT mit einer Studiendauer von 52±6 Wochen eingeflossen. Diese Stratifizierung wurde basierend auf der Analyse der relativen Effektmaße im zeitlichen Verlauf vorgenommen. Die mittlere Veränderung des HbA1c-Wertes variierte im zeitlichen Verlauf in inkonsistenter Weise unter den in den RCT untersuchten Komparatoren. Patienten, die in den Studien mit Sekretagoga wie Sulfonylharnstoffen oder Gliniden behandelt wurden, hatten in den ersten vier bis sechs Behandlungsmonaten eine vergleichsweise starke HbA1c-Wert-Senkung, die aber nicht

anhaltend war. Im weiteren Behandlungsverlauf kam es zu einer stetigen Erhöhung des HbA1c-Wertes bis zum Erreichen des Ausgangswertes, d. h. die kurzfristig erzielten Behandlungseffekte persistieren nicht über einen längeren Zeitraum. Dieser Verlauf wird als sog. *J-curve* bezeichnet und kommt typischerweise bei Sekretagoga, nicht aber bei anderen Wirkstoffklassen der OAD vor. Auch bei den Endpunkten „mittlere Veränderung des Körpergewichts“ sowie „Hypoglykämien“ sind Veränderungen der Effektmaße im Studienverlauf erkennbar. Hinzu kommt, dass bei Sulfonylharnstoffen Titrationsphasen mit in die Studiendauer einfließen. Es wurde daher darauf geachtet, dass ein ausreichend großer Zeitraum zwischen Titration und Endpunkterhebung bestand.

Aufgrund der beschriebenen Situation wurden zwei MTC-Modelle erstellt, wobei Sulfonylharnstoffe und Glinide aus dem 24-wöchigen Netzwerk ausgeschlossen wurden (Ausnahme: Endpunkt systolischer Blutdruck). Es wurde ein sechswöchiges Zeitfenster um die beiden für die Netzwerke relevanten Studiendauern gelegt, d. h. im 24-wöchigen Netzwerk gingen Studien mit einer Dauer von 18-30 Wochen und im 52-wöchigen Netzwerk Studien mit einer Dauer von 46-58 Wochen ein. Innerhalb dieser Zeitfenster können die relativen Effektmaße als gleich angesehen werden.

4.3.2.1.1.1 Studien des pharmazeutischen Unternehmers

Benennen Sie in der nachfolgenden Tabelle alle Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war.

Von den randomisierten kontrollierten Studien, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden, sollen die Studien, deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle randomisierten kontrollierten Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.

Benennen Sie nur randomisierte, kontrollierte Studien, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-75: Liste der Studien des pharmazeutischen Unternehmers – Studien für indirekte Vergleiche auf Basis von RCT

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
D1690C00004	ja	abgeschlossen	Behandlungsphase: 52 Wochen	Metformin (1500–2500 mg/Tag)+ Dapagliflozin (2,5–10 mg/Tag) Metformin (1500–2500 mg/Tag)+Glipizid (5 mg-20 mg/Tag)
D1690C00004 (LT I)	nein	abgeschlossen	Verlängerungsphase I: bis 104 Wochen	Metformin (1500–2500 mg/Tag)+ Dapagliflozin (2,5–10 mg/Tag) Metformin (1500–2500 mg/Tag)+Glipizid (5 mg-20 mg/Tag)
D1690C00004 (LT II)	nein	laufend	Verlängerungsphase II: bis 208 Wochen	Metformin (1500–2500 mg/Tag)+ Dapagliflozin (2,5–10 mg/Tag) Metformin (1500–2500 mg/Tag)+Glipizid (5 mg-20 mg/Tag)
MB102014	ja	abgeschlossen	Behandlungsphase: 24 Wochen Verlängerungsphase: bis 104 Wochen	Metformin (≥1500 mg/Tag)+ Dapagliflozin (2,5, 5 und 10 mg/Tag) Placebo+Metformin
D1690C00012	ja	abgeschlossen	Behandlungsphase: 24 Wochen Verlängerungsphase: bis 104 Wochen	Metformin (≥1500 mg/Tag)+ Dapagliflozin (10 mg) Placebo+Metformin
MB102003	nein	abgeschlossen	2 Wochen	Dapagliflozin (5, 25 und 100 mg/Tag)+/- Metformin Placebo+/-Metformin
MB102020	nein	abgebrochen	4 Wochen	Dapagliflozin (5 mg) +Vorbehandlung Dapagliflozin (20 mg) +Vorbehandlung

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
MB102045	nein	abgeschlossen	12 Wochen	Dapagliflozin (5 mg/Tag)+Vorbehandlung Placebo+Vorbehandlung
MB102035	nein	abgeschlossen	12 Wochen	Dapagliflozin (10 mg/Tag)+Vorbehandlung (Metformin und/oder Sulfonylharnstoffe) Placebo+Vorbehandlung (Metformin und/oder Sulfonylharnstoffe)
MB102054	nein	abgeschlossen*	24 Wochen	Dapagliflozin (5 und 10 mg/Tag)+Vorbehandlung Placebo+Vorbehandlung
MB102055	nein	laufend	24 Wochen	Dapagliflozin (5 und 10 mg/Tag)+Metformin (1500-3000 mg/Tag)+Vorbehandlung Placebo+Metformin+Vorbehandlung
MB102073	nein	laufend	12 Wochen	Dapagliflozin (10 mg/Tag)+Vorbehandlung Placebo+Vorbehandlung
MB102077	nein	laufend	12 Wochen	Dapagliflozin (10 mg/Tag)+Vorbehandlung Placebo+Vorbehandlung
D1691C00003	nein	abgeschlossen	16 Wochen	Metformin (≥ 1500 mg/Tag)+ Dapagliflozin (2,5, 5 und 10 mg/Tag) Placebo+Metformin
D1690C00018	nein	abgeschlossen	Behandlungsphase: 24 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00018 (LT I)	nein	abgeschlossen	Verlängerungsphase I: bis 52 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00018 (LT II)	nein	laufend	Verlängerungsphase II: bis 104 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
D1690C00019	nein	abgeschlossen	Behandlungsphase: 24 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00019 (LT I)	nein	abgeschlossen	Verlängerungsphase I: bis 52 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00019 (LT II)	nein	laufend	Verlängerungsphase II: bis 104 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00007	nein	Abgebrochen vor Einschluss des ersten Patienten	24 Wochen	Dapagliflozin (10 mg/Tag)+Metformin Sitagliptin (100 mg/Tag) +Metformin
CV181169	nein	laufend	24 Wochen	Dapagliflozin (10 mg)+Metformin XR (1500-2000 mg) +Saxagliptin (5 mg) Dapagliflozin (10 mg)+Metformin XR (1500-2000 mg) +Placebo Metformin XR (1500-2000 mg) +Saxagliptin (5 mg) +Placebo
CV181168	nein	laufend	52 Wochen	Dapagliflozin (10 mg)+Metformin (≥ 1500 mg) +Saxagliptin (5 mg) Dapagliflozin (10 mg)+Metformin (≥ 1500 mg) +Placebo

LT: Long term; XR; Extended release

* Es liegt noch kein Studienbericht vor.

Geben Sie an, welchen Stand die Information in Tabelle 4-75 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus sollte nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Der Informationsstand für den Status der angegebenen Studien bezieht sich auf den 29.10.2012.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-6 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-76: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
D1690C00004 (LT I)	Die Verlängerungsphase I (bis 104 Wochen) wurde nicht verwendet, da nur Studien, mit einer Studiendauer von maximal 52 Wochen in das Netzwerk eingeschlossen wurden.
D1690C00004 (LT II)	Die Verlängerungsphase II (bis 208 Wochen) läuft noch. Der Studienbericht wird für Mitte 2013 erwartet.
MB102003	Studiendauer zu gering.
MB102020	Phase-1-Studie. Die Studiendauer beträgt nur 4 Wochen.
MB102045	Studiendauer zu gering.
MB102035	Studiendauer zu gering.
MB102054	Die Studie wurde kürzlich abgeschlossen.. Es liegen noch keine Ergebnisse vor. Die Studie wurde daher in Anhang 4-E aufgenommen.
MB102055	Die Studie läuft noch. Es liegen noch keine Ergebnisse vor.
MB102073	Studiendauer zu gering. Studie läuft noch.
MB102077	Studiendauer zu gering. Studie läuft noch.
D1691C00003	Studiendauer zu gering.
D1690C00018 (inkl. LT I und II)	Die Studie ist placebokontrolliert und enthält die falsche bzw. keine Vergleichstherapie
D1690C00019 (inkl. LT I und II)	Die Studie ist placebokontrolliert und enthält die falsche bzw. keine Vergleichstherapie.
D1690C00007	Die Studie wurde vor Einschluss des ersten Patienten abgebrochen.
CV181169	Die Studie läuft noch. Es liegen noch keine Ergebnisse vor.
CV181168	Die Studie läuft noch. Es liegen keine Ergebnisse vor.

4.3.2.1.1.2 Studien aus der bibliografischen Literaturrecherche

Sofern eine bibliografische Literaturrecherche durchgeführt wurde, beschreiben Sie nachfolgend das Ergebnis dieser Recherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-C.

Falls die Recherche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel gemeinsam mit der Recherche zu einer anderen Fragestellung (z. B. zu indirekten Vergleichen) durchgeführt wurde, unterteilen Sie die Angaben zu relevanten Treffern und Studien entsprechend diesen Fragestellungen.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.

Da für die vom G-BA festgelegte Vergleichstherapie (Metformin+Sulfonylharnstoff) eine direkt vergleichende Studie vorliegt, war keine bibliografische Literaturrecherche notwendig.

Netzwerk-Meta-Analyse

Zur Bewertung des medizinischen Zusatznutzens von Dapagliflozin unter der Fragestellung „Identifikation von Studien (RCT), um die Sicherheit und Wirksamkeit von Dapagliflozin mit anderen Antidiabetika zu vergleichen“ wurde eine gemeinsame Suche nach Studien mit Dapagliflozin und Studien mit anderen Arzneimitteln zur Behandlung von Patienten mit Typ-2-Diabetes durchgeführt, mit dem Ziel eine MTC durchführen zu können.

Die umfassende bibliografische Literaturrecherche vom 11. Mai 2011 nach randomisierten, kontrollierten Studien erzielte insgesamt 8252 Treffer. Nach automatisiertem und händischem Ausschluss der Duplikate (n=5370) wurden die verbleibenden 2882 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den im Bericht (Oxford Outcomes Ltd., 2011b) definierten Kriterien gesichtet. Basierend auf den Informationen des Titels und des Abstracts wurden 2651 Publikationen begründet durch Ausschlusskriterien als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Für die verbleibenden 231 Publikationen, deren Titel und Abstract die Einschlusskriterien erfüllen, wurden im Anschluss die Volltexte hinsichtlich dieser Kriterien überprüft. Aufgrund der in Abschnitt 4.3.2.1.1 beschriebenen Situation wurden zwei MTC-Modelle erstellt, wobei Sulfonylharnstoffe und Glinide aus dem 24-wöchigen Netzwerk ausgeschlossen wurden. Daher wurden die Studien von Ristic (2006), Marre (2002), Arechavaleta (2011), Umpierrez (2006) und Moses (1999) für die MTC-Analyse ausgeschlossen. Eine Ausnahme stellt der Endpunkt systolischer Blutdruck dar, weshalb die Studien von Charpentier (2001) und Papathanassiou (2009) in der MTC nur für diesen Endpunkt berücksichtigt wurden.

Im Ergebnis entsprachen 213 nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C, Tabelle 4-153).

Somit ergab die bibliografische Literaturrecherche 18 relevante Publikationen. Dabei handelt es sich um die Studien von (Charpentier et al., 2001; DeFronzo et al., 2005; Matthews et al., 2005; Charbonnel et al., 2006; Bosi et al., 2007; Nauck et al., 2007; Bolli et al., 2008; Raz et al., 2008; Scott et al., 2008; DeFronzo et al., 2009; Kaku, 2009; Nauck et al., 2009; Bergenstal et al., 2010b; Filozof und Gautier, 2010; Göke et al., 2010; Matthews et al., 2010; Taskinen et al., 2011). Zudem wurde die Publikation von (Ferrannini et al., 2009)

identifiziert. Es handelt sich hierbei jedoch um eine Doppelpublikationen zur Publikation von Matthews et al., 2010.

Aufgrund der Literaturrecherche wurden somit 17 relevante Studien identifiziert.

Diese 17 Studien flossen in die MTC mit ein.

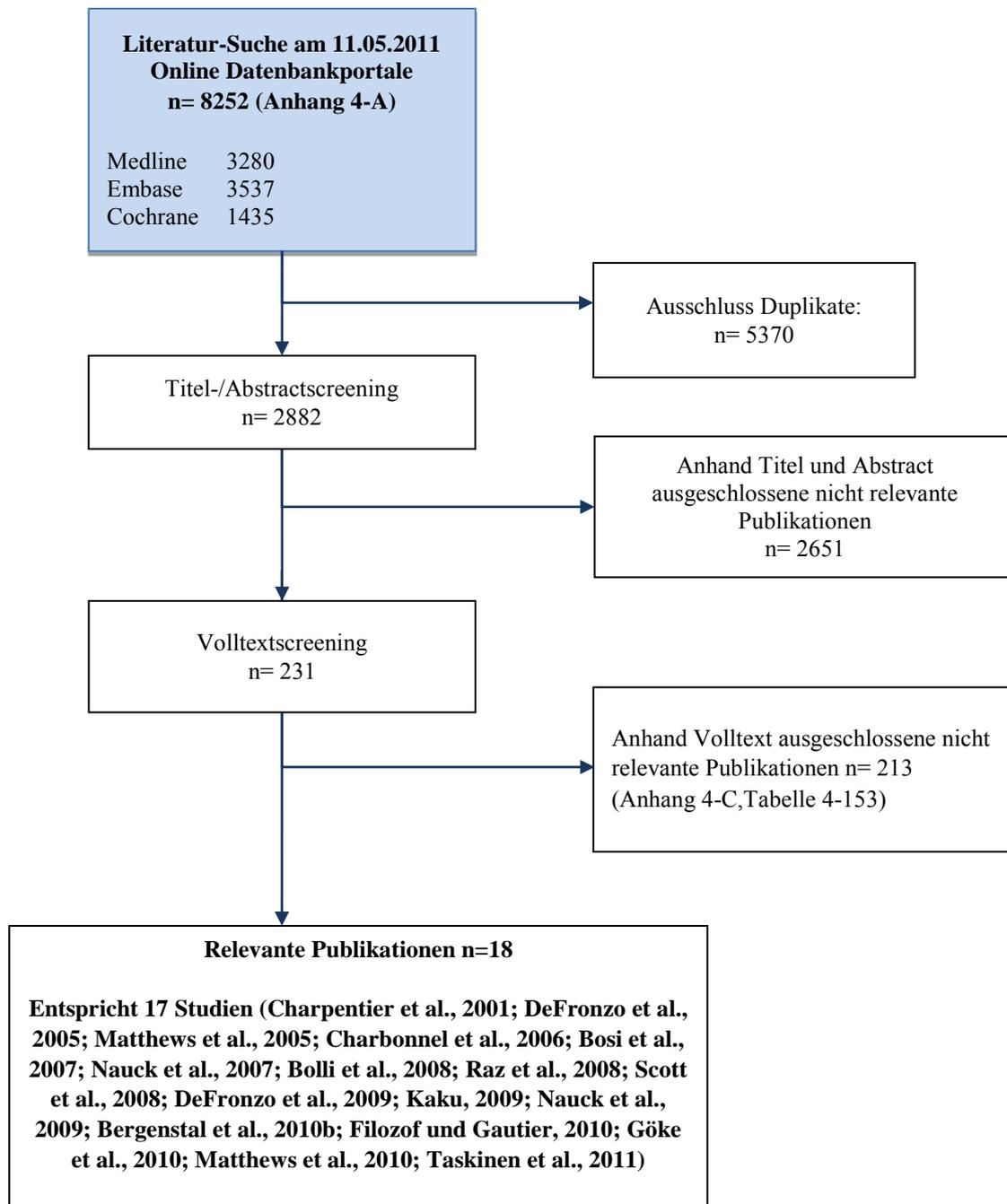


Abbildung 9: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien zur Durchführung einer MTC (11. Mai 2011)

Am 04. Juni 2012 wurde die bibliografische Literaturrecherche aktualisiert. Die umfassende bibliografische Literaturrecherche vom 04. Juni 2012 nach randomisierten, kontrollierten Studien erzielte insgesamt 9716 Treffer. Nach automatisiertem und händischem Ausschluss der Duplikate (n=6393) wurden die verbleibenden 3323 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den im Bericht (Oxford Outcomes Ltd., 2012c) definierten Kriterien gesichtet. Basierend auf den Informationen des Titels und des Abstracts wurden 3083 Publikationen begründet durch Ausschlusskriterien als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Für die verbleibenden 240 Publikationen, deren Titel und Abstract die Einschlusskriterien erfüllen, wurden im Anschluss die Volltexte hinsichtlich dieser Kriterien überprüft. Zusätzlich wurden aufgrund der in Abschnitt 4.3.2.1.1 beschriebenen Situation zwei MTC-Modelle erstellt, wobei Sulfonylharnstoffe und Glinide aus dem 24-wöchigen Netzwerk ausgeschlossen wurden. Daher wurde die Studie von Pfützner et al., 2011b für die MTC-Analyse ausgeschlossen. Im Ergebnis entsprachen 221 nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C, Tabelle 4-153 und Tabelle 4-154). Somit ergab die bibliografische Literaturrecherche 19 relevante Publikationen (davon eine Doppelpublikation [Ferrannini et al. 2009]), also 18 relevante Studien. Durch die Aktualisierung der bibliografischen Literaturrecherche ergab sich somit eine neue relevante Publikation: Yang et al., 2011.

Die Publikation von (Yang et al., 2011) wurde nicht in der MTC berücksichtigt. Im Folgenden werden die Ergebnisse jedoch getrennt dargestellt und auf Unterschiede zur MTC und auf den möglichen Einfluss auf die Ergebnisse der MTC eingegangen.

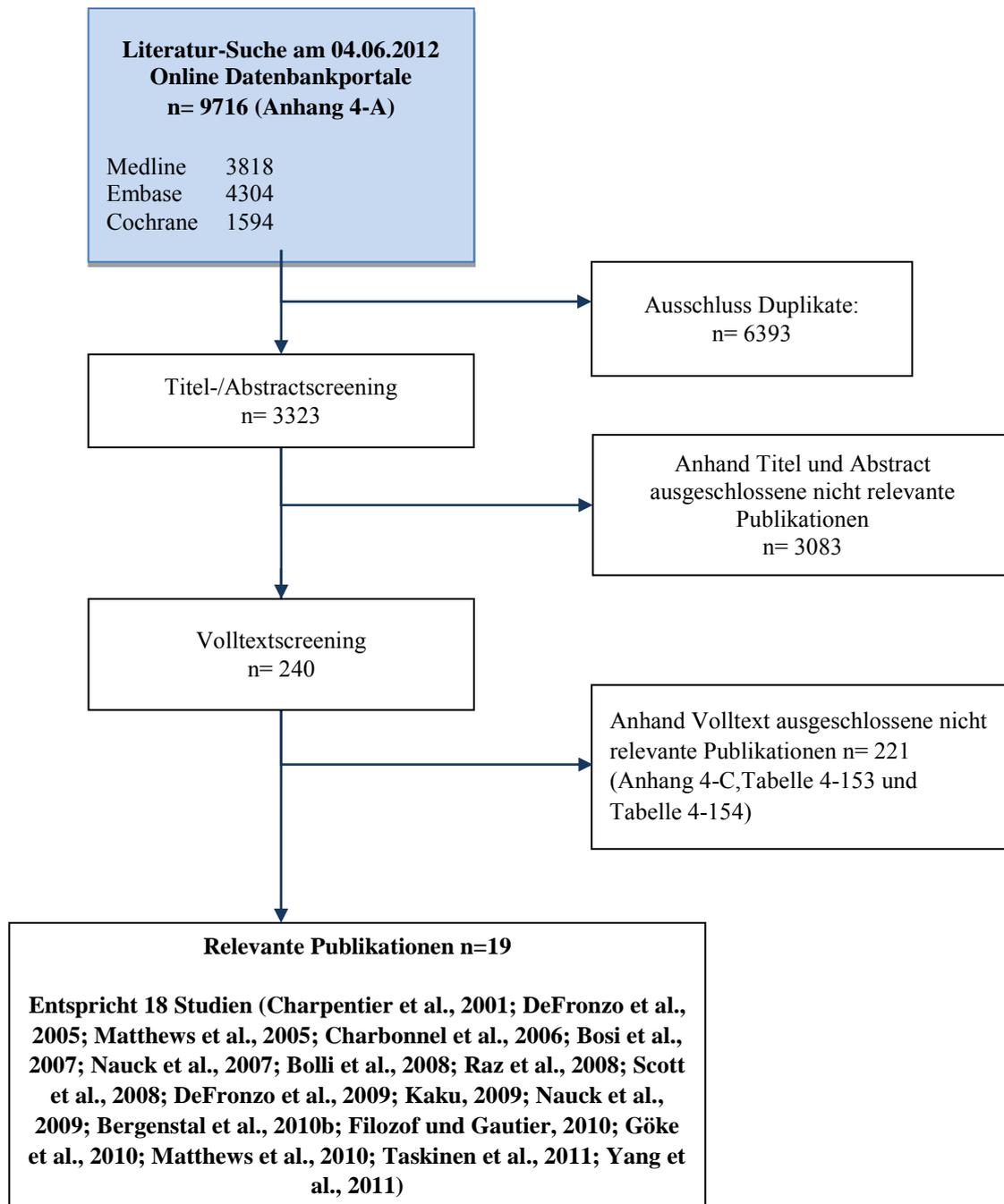


Abbildung 10: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien zur Durchführung einer MTC (04. Juni 2012)

Am 09. Oktober 2012 wurde die bibliografische Literaturrecherche ein zweites Mal aktualisiert. Die umfassende bibliografische Literaturrecherche vom 09. Oktober 2012 nach randomisierten, kontrollierten Studien erzielte insgesamt 10265 Treffer. Nach automatisiertem und händischem Ausschluss der Duplikate (n=6685) wurden die verbleibenden 3580 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den im Bericht (Oxford Outcomes Ltd., 2012a) definierten Kriterien gesichtet. Basierend auf den Informationen des Titels und des Abstracts wurden 3333 Publikationen begründet durch Ausschlusskriterien als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Für die verbleibenden 247 Publikationen, deren Titel und Abstract die Einschlusskriterien erfüllen, wurden im Anschluss die Volltexte hinsichtlich dieser Kriterien überprüft. Im Ergebnis entsprachen 226 nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C, Tabelle 4-153, Tabelle 4-154, Tabelle 4-155). Somit ergab die bibliografische Literaturrecherche 21 relevante Publikationen (davon eine Doppelpublikation [Ferrannini et al. 2009]), also 20 relevante Studien. Durch die Aktualisierung der bibliografischen Literaturrecherche ergaben sich somit zwei neue relevante Publikation: Yang et al., 2012 und Gallwitz 2012.

Die zwei Publikationen von (Gallwitz et al., 2012; Yang et al., 2012), die durch das zweite Update identifiziert wurden, wurden nicht in der MTC berücksichtigt. Im Folgenden werden die Ergebnisse jedoch getrennt dargestellt und auf Unterschiede zur MTC und auf den möglichen Einfluss auf die Ergebnisse der MTC eingegangen.

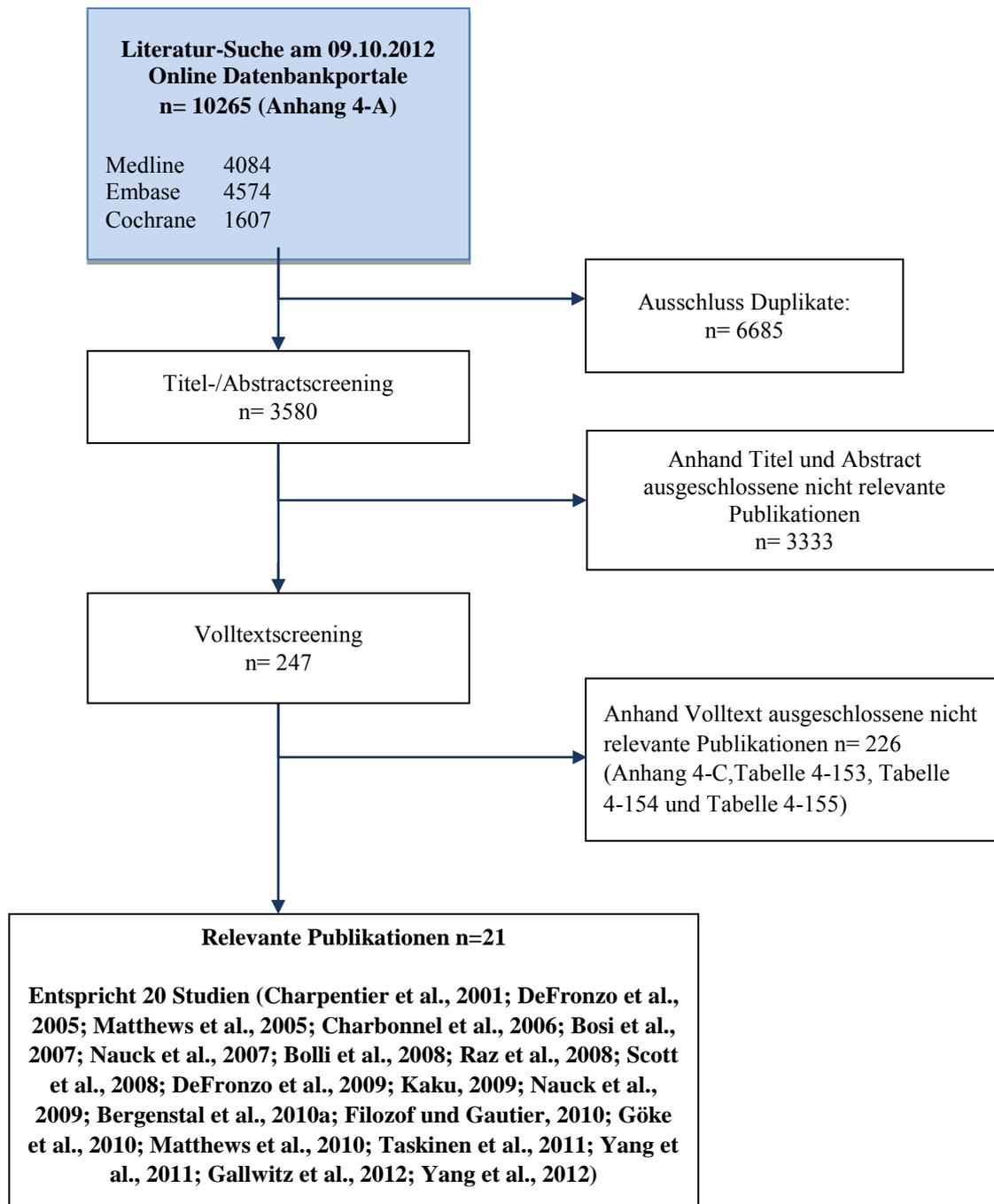


Abbildung 11: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien zur Durchführung einer MTC (09. Oktober 2012)

4.3.2.1.1.3 Studien aus der Suche in Studienregistern

Tabelle 4-77 : Relevante Studien aus der Suche in Studienregistern

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate ^{a)})	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. ^{b)})
MB102014	clinicaltrials.gov [NCT00528879] (Clinicaltrials.gov, 2012w)	ja	nein
D1690C00012	clinicaltrials.gov [NCT00855166] (Clinicaltrials.gov, 2012x) WHO ICTRP [NCT00855166] (ICTRP Search Portal, 2012b)	ja	nein
Bergental 2010a	clinicaltrials.gov [NCT00754988] (Clinicaltrials.gov, 2012z)	nein	nein
Pratley 2010	clinicaltrials.gov [NCT00700817] (Clinicaltrials.gov, 2012{) WHO ICTRP (ICTRP Search Portal, 2012f)	nein	nein
D1690C00004	clinicaltrials.gov [NCT00660907] (Clinicaltrials.gov, 2012a) WHO ICTRP [NCT00660907] (ICTRP Search Portal, 2012c) [EUCTR2007-005220-33-GB] (ICTRP Search Portal, 2012a)	ja	nein
Pan 2012	clinicaltrials.gov [NCT00822211] (Clinicaltrials.gov, 2012t)	nein	nein
Scott 2008	clinicaltrials.gov [NCT00541775] (Clinicaltrials.gov, 2012s)	nein	nein

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d.^b)
Charbonnel 2006	clinicaltrials.gov [NCT00086515] (Clinicaltrials.gov, 2012q)	nein	ja
DeFronzo 2009	clinicaltrials.gov [NCT00121667] (Clinicaltrials.gov, 2012o) WHO ICTRP [NCT00121667] (ICTRP Search Portal, 2012d)	nein	ja
Bosi 2007	clinicaltrials.gov [NCT00099892] (Clinicaltrials.gov, 2012m)	nein	ja
Taskinen 2011	clinicaltrials.gov [NCT00601250] (Clinicaltrials.gov, 2012e)	nein	ja
Raz 2008	clinicaltrials.gov [NCT00337610] (Clinicaltrials.gov, 2012l)	nein	ja
Bolli 2008	clinicaltrials.gov [NCT00237237] (Clinicaltrials.gov, 2012k)	nein	ja
Bergental 2010b	clinicaltrials.gov [NCT00637273] (Clinicaltrials.gov, 2012i)	nein	nein
DeFronzo 2005	clinicaltrials.gov [NCT00039013] (Clinicaltrials.gov, 2012g)	nein	ja
Nauck 2009	clinicaltrials.gov [NCT00318461] (Clinicaltrials.gov, 2012p) WHO ICTRP [NCT00318461] (ICTRP Search Portal, 2012e)	nein	ja
Nauck 2007	clinicaltrials.gov [NCT00094770] (Clinicaltrials.gov, 2012v)	nein	ja

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d.^b)
Göke 2010	clinicaltrials.gov [NCT00575588] (Clinicaltrials.gov, 2012n)	nein	ja
Filozof 2010	clinicaltrials.gov [NCT00102466] (Clinicaltrials.gov, 2012y)	nein	ja
Yang 2012	clinicaltrials.gov [NCT00813995] (Clinicaltrials.gov, 2012j)	nein	ja
Gallwitz 2012	clinicaltrials.gov [NCT00359762] (Clinicaltrials.gov, 2012u)	nein	nein
<p>a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt</p>			

4.3.2.1.1.4 Resultierender Studienpool

Tabelle 4-78: Studienpool – Indirekte Vergleiche auf Basis von RCT

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Zulassungsstudie des zu bewertenden Arzneimittels (ja/nein)	Gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
MB102014	Ja	Ja	Nein	Ja (Bristol-Myers Squibb Company, 2009) (Bristol-Myers Squibb Company, 2010a)	Ja clinicaltrials.gov [NCT00528879] (Clinicaltrials.gov, 2012w)	Ja (Bailey et al., 2010)
D1690C00012	Ja	Ja	Nein	Ja (AstraZeneca, 2010a; AstraZeneca, 2011a) (AstraZeneca, 2012)	Ja clinicaltrials.gov [NCT00855166] (Clinicaltrials.gov, 2012x) WHO ICTRP [NCT00855166] (ICTRP Search Portal, 2012b)	Ja (Bolinder et al., 2012)
Scott 2008	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00541775] (Clinicaltrials.gov, 2012s)	Ja (Scott et al., 2008)
Charbonnel 2006	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00086515] (Clinicaltrials.gov, 2012q)	Ja (Charbonnel et al., 2006)

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Zulassungsstudie des zu bewertenden Arzneimittels (ja/nein)	Gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
DeFronzo 2009	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00121667] (Clinicaltrials.gov, 2012o) WHO ICTRP [NCT00121667] (ICTRP Search Portal, 2012d)	Ja (DeFronzo et al., 2009)
Bosi 2007	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00099892] (Clinicaltrials.gov, 2012m)	Ja (Bosi et al., 2007)
Taskinen 2011	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00601250] (Clinicaltrials.gov, 2012e)	Ja (Taskinen et al., 2011)
Bergental 2010a	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00754988] (Clinicaltrials.gov, 2012z)	Ja (Bergental et al., 2010a)
Raz 2008	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00337610] (Clinicaltrials.gov, 2012l)	Ja (Raz et al., 2008)
Bolli 2008	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00237237] (Clinicaltrials.gov, 2012k)	Ja (Bolli et al., 2008)
Bergental 2010b ¹	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00637273] (Clinicaltrials.gov, 2012i)	Ja (Bergental et al., 2010b)

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Zulassungsstudie des zu bewertenden Arzneimittels (ja/nein)	Gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
DeFronzo 2005	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00039013] (Clinicaltrials.gov, 2012g)	Ja (DeFronzo et al., 2005)
Nauck 2009	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00318461] (Clinicaltrials.gov, 2012p) WHO ICTRP [NCT00318461] (ICTRP Search Portal, 2012e)	Ja (Nauck et al., 2009)
Pratley 2010	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00700817] (Clinicaltrials.gov, 2012j) WHO ICTRP (ICTRP Search Portal, 2012f)	Ja (Pratley et al., 2010)
Kaku 2009	Nein	Nein	Ja	Nein	Nein	Ja (Kaku, 2009)
Charpentier 2001	Nein	Nein	Ja	Nein	Nein	Ja (Charpentier et al., 2001)
Papathanassiou 2009	Nein	Nein	Nein	Nein	Nein	Ja (Papathanassiou et al., 2009)
D1690 C00004	Ja	Ja	Ja	Ja (AstraZeneca, 2010b; AstraZeneca, 2011b)	Ja (Clinicaltrials.gov, 2012a) (ICTRP Search Portal, 2012a)	Ja (Nauck et al., 2011)

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Zulassungsstudie des zu bewertenden Arzneimittels (ja/nein)	Gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
Matthews 2005	Nein	Nein	Ja	Nein	Nein	Ja (Matthews et al., 2005)
Nauck 2007	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00094770] (Clinicaltrials.gov, 2012v)	Ja (Nauck et al., 2007)
Göke 2010	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00575588] (Clinicaltrials.gov, 2012n)	Ja (Göke et al., 2010)
Filozof 2010	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00102466] (Clinicaltrials.gov, 2012y)	Ja (Filozof und Gautier, 2010)
Salvadeo 2010 ²	Nein	Nein	Ja	Nein	Nein	Ja (Salvadeo et al., 2010)
Derosa 2010 ²	Nein	Nein	Ja	Nein	Nein	Ja (Derosa et al., 2010)
Matthews 2010 ³	Nein	Nein	Ja	Nein	Nein	Ja (Ferrannini et al., 2009; Matthews et al., 2010) ³
Yang 2011*	Nein	Nein	Ja	Nein	Nein	Ja (Yang et al., 2011)
Derosa 2011a* ²	Nein	Nein	Ja	Nein	Nein	Ja (Derosa et al., 2011a)
Derosa 2011b* ⁴	Nein	Nein	Ja	Nein	Nein	Ja (Derosa et al., 2011b)

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Zulassungsstudie des zu bewertenden Arzneimittels (ja/nein)	Gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
Pan 2012*	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00822211] (Clinicaltrials.gov, 2012t)	Ja (Pan et al., 2012)
Yang 2012*	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00813995] (Clinicaltrials.gov, 2012j)	Ja (Yang et al., 2012)
De Fronzo 2012* ²	Nein	Nein	Ja	Nein	Nein	Ja (DeFronzo et al., 2012)
Gallwitz 2012*	Nein	Nein	Ja	Nein	Ja clinicaltrials.gov [NCT00359762] (Clinicaltrials.gov, 2012u)	Ja (Gallwitz et al., 2012)

¹ der Studienarm Exenatide+Metformin wurde in der Meta-Analyse nicht berücksichtigt, da die Dosierung (2 mg Exenatide [einmal wöchentlich]+Metformin 1,5 g) zum Zeitpunkt der Durchführung der Meta-Analyse nicht zugelassen war.

² Studie wurde nicht in der Literaturrecherche oder Studienregister identifiziert, sondern als Konferenz-Abstract oder aus der Referenzliste einer eingeschlossenen Studie oder Meta-Analyse

³ Ferrannini et al. 2009 entspricht der Doppelpublikation von Mathews et al. 2010

⁴ entspricht der Publikation zum Konferenz-Abstract von Salvadeo, 2010

* Studien, die durch aktualisierte Literaturrecherche bzw. Registersuche gefunden wurden

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

4.3.2.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-F zu hinterlegen.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-79: Charakterisierung der eingeschlossenen Studien – Studien für indirekte Vergleiche

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
MB102014	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	Dapagliflozin 10 mg + Metformin (n=135) Placebo+Metformin (n=137)	24 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert. Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat.
D1690C00012	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend	Dapagliflozin 10 mg+ Metformin (n=91) Placebo+Metformin (n=91)	24 Wochen	Europaweit	Primärer Endpunkt: Gewichtsveränderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
		kontrolliert werden kann.				<p>sowie weitere Endpunkte: HbA1c-Änderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert. Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat</p>
Scott 2008	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Sitagliptin 100 mg +≥Metformin (n=94) Placebo+Metformin (n=91)	18 Wochen	International	<p>Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert</p> <p>Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie</p>

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						auftrat
Charbonnel 2006	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Sitagliptin 100 mg + Metformin (n=464) Metformin+Placebo (n=237)	24 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
DeFronzo 2009	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Saxagliptin 5 mg+ Metformin (n=191) Placebo+Metformin (n=192)	24 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						Vergleich zum Ausgangswert. Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Bosi 2007	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Vildagliptin 100 mg + Metformin (n=183) Placebo + Metformin (n=181)	24 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Taskinen 2011	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht	Linagliptin 5 mg + Metformin (n=523) Placebo + Metformin (n=177)	24 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
		kontrolliert werden kann.				sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Bergental 2010a	RCT, doppelblind, placebokontrolliert	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Sitagliptin 100 mg+ Metformin (n=190)	24 Wochen	N/R	Patientenrelevante Endpunkte: HbA1c-Änderung im Vergleich zum Ausgangswert Anteil der Patienten, die einen HbA1c-Wert von <7,0% erreichen Gewichtsveränderung im Vergleich zum Ausgangswert
Raz 2008	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend	Sitagliptin 100 mg+ Metformin (n=96) Placebo+Metformin (n=94)	30 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
		kontrolliert werden kann.				sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Bolli 2008	RCT, doppelblind, aktivkontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Vildagliptin 100 mg+ Metformin (n=295) Pioglitazone+ Metformin (n=280)	24 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Bergenstal 2010b	RCT, doppelblind, aktivkontrolliert,	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren	Exenatide 2 mg +Metformin (n=160) ¹ Sitagliptin 100 mg	26 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	multizentrisch	Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	+Metformin (n=166) Pioglitazone 45 mg +Metformin (n=165)			Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert. Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
DeFronzo 2005	RCT, doppelblind, placebokontrolliert	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Exenatide (5 µg BID) (n=110) Exenatide (10 µg BID) (n=113) Placebo (n=113)	30 Wochen	USA	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Nauck 2009	RCT, doppelblind, aktiv- und placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Liraglutide 1,2 mg+ Metformin (n=241) Liraglutide 1,8 mg+ Metformin (n=242) Glimepirid 4 mg +Metformin (n=244) Placebo+Metformin (n=122)	26 Wochen	International	<p>Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert</p> <p>Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert. Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat</p>
Pratley 2010	RCT, aktivkontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren	Liraglutide 1,2 mg +Metformin (n=255) Liraglutide 1,8 mg	26 Wochen	N/R	<p>Primärer Endpunkt: HbA1c-Änderung im Vergleich zum</p>

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
		Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	+Metformin (n=221) Sitagliptin 100 mg +Metformin (n=219)			Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert. Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Kaku 2009	RCT, doppelblind, placebokontrolliert multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Pioglitazon 30 mg +Metformin (n=83) Placebo+Metformin (n=86)	28 Wochen	Japan	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Anteil der Patienten, die einen HbA1c-Wert von <6,5% erreichen

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Charpentier 2001	RCT, doppelblind, placebokontrolliert	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Glimperid 1-6 mg+ Metformin (n=147) Placebo+2,55 g Metformin (n=75)	20 Wochen	Frankreich	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert. Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
Papathanassiou 2009	RCT, aktivkontrolliert	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Pioglitazon 30 mg +Metformin (n=14) Glimepirid 4 mg +Metformin (n=14)	26 Wochen	Griechenland	Patientenrelevante Endpunkte HbA1c-Änderung im Vergleich zum Ausgangswert Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert.
D1690C00004	RCT, doppelblind, aktivkontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Dapagliflozin 10 mg+ Metformin (n=400) Glipizid 5-20 mg+ Metformin (n=408)	52 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Matthews 2005	RCT, doppelblind, aktivkontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann.	Pioglitazon 30-45 mg+ Metformin (n=313) Gliclazid 160-320 mg+ Metformin (n=317)	52 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Anteil der Patienten, die einen HbA1c-Wert von <7,0% erreichen Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Nauck 2007	RCT, doppelblind, aktivkontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert	Sitagliptin 100 mg +Metformin (n=584) Glipizid 5-20 mg +Metformin (n=588)	52 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
		alleine mit Metformin nicht ausreichend kontrolliert werden kann	(insgesamt wurden 1172 Patienten randomisiert)			Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Göke 2010	RCT, doppelblind, aktivkontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	Saxagliptin 5 mg+ Metformin (n=430) Glipizid 5-20 mg +Metformin (n=428)	52 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert Anzahl und Anteil

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Filozof 2010	RCT, doppelblind, aktivkontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	Vildagliptin 100 mg Metformin Gliclazid 80-320 mg+ Metformin	52 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert
Salvadeo 2010	RCT, aktivkontrolliert	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	Exenatide 20 µg +Metformin Glimepirid 6 mg+ Metformin	52 Wochen	N/R	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert
Derosa 2010	RCT, einfachblind, aktivkontrolliert	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert	Exenatide 20 µg+ Metformin Glibenclamid 15 mg+ Metformin	52 Wochen	Italien	Patientenrelevante Endpunkte: HbA1c-Änderung im Vergleich zum

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
		alleine mit Metformin nicht ausreichend kontrolliert werden kann	(insgesamt wurden 128 Patienten randomisiert)			Ausgangswert Gewichtsveränderung im Vergleich zum Ausgangswert
Matthews 2010	RCT, doppelblind, aktivkontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	2-6 mg Glimepirid+≥1,5 g Metformin (1383) 100 mg Vildagliptin+≥1,5 g Metformin(1389) (insgesamt wurden 3118 Patienten randomisiert)	104 Wochen	International	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat
Yang 2011*	RCT, doppelblind,	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden	Saxagliptin 5 mg +Metformin (n=283) Placebo+Metformin (n=287)	24 Wochen	Asien (einschließlich Indien)	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
		kann				Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat Blutdruckänderung im Vergleich zum Ausgangswert Unerwünschte Ereignisse
Derosa 2011a*	RCT, doppelblind, aktivkontrolliert	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	Pioglitazone 45 mg +Metformin (n=146) Glibenclamide 15 mg +Metformin (n=142)	52 Wochen	Italien	Primärer Endpunkt: HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Unerwünschte Ereignisse
Derosa 2011b* ²	RCT, einfachblind,	Erwachsene Patienten mit	Exenatide 20 µg+ Metformin (n=57)	52 Wochen	Italien	Primärer Endpunkt:

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	aktivkontrolliert	Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	Glimepirid 6 mg +Metformin (n=54)			HbA1c-Änderung im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtsveränderung im Vergleich zum Ausgangswert Unerwünschte Ereignisse
Pan 2012*	RCT, doppelblind, placebo-kontrolliert	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	Vildagliptin 50 mg + Metformin (n=148) Vildagliptin 100 mg + Metformin (n=146) Placebo+Metformin (n=144)	24 Wochen	China	Primärer Endpunkt: Veränderung des HbA1c-Ausgangswertes bis Woche 24 Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Veränderung des FPG-Ausgangswertes bis Woche 24 Unerwünschte Ereignisse
DeFronzo 2012* ³	RCT, doppelblind, placebo-	Erwachsene Patienten mit Diabetes mellitus	Alogliptin (12,5 mg bzw. 25 mg)+ Metformin (n=128 /	26 Wochen	20 Länder	Primärer Endpunkt: Veränderung des

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	kontrolliert; multizentrisch (327)	Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	129) Pioglitazon (15mg / 30 mg / 45 mg)+ Metformin (n=130 / 129 /129) Pioglitazon (15mg)+Alogliptin (12,5mg / 25 mg)+Metformin (n=130 / 130) Pioglitazon (30 mg)+Alogliptin (12,5 mg / 25 mg)+Metformin (n=130 / 130) Pioglitazon (45 mg)+Alogliptin (12,5 mg / 25 mg)+Metformin (n=130 / 130) Placebo+Metformin (n=129)			HbA1c- Ausgangswertes bis Woche 26 Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Hypoglykämien Gewichtsveränderung im Vergleich zum Ausgangswert Unerwünschte Ereignisse
Gallwitz 2012*	RCT, doppelblind, aktiv-kontrolliert; multizentrisch (209)	Erwachsene Patienten mit Diabetes mellitus Typ 2, die mit Monoformin oder Metformin+ein weiteres OAd erhalten.	Linagliptin+ Metformin (n=777) Glimepirid +Metformin (n=775)	2 Jahre	12.Feb 2008 – 21.Dez. 2010 16 Ländern	Primärer Endpunkt: Veränderung des HbA1c- Ausgangswertes bis Woche 104 Patientenrelevante sekundäre Ziele sowie weitere

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						Endpunkte: Hypoglykämien Gewichtsveränderung im Vergleich zum Ausgangswert Unerwünschte Ereignisse
Yang 2012*	RCT, doppelblind, placebo-kontrolliert; multizentrisch (17)	Erwachsene Patienten aus China mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit Metformin nicht ausreichend kontrolliert werden kann	Sitagliptin 100mg + Metformin (n=197) Placebo+Metformin (n=198)	24 Wochen	13. Jan 2009 – 9. Aug 2010 In China	Primärer Endpunkt: Veränderung des HbA1c-Ausgangswertes bis Woche 104 Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Hypoglykämien Unerwünschte Ereignisse

BID: Zweimal täglich

¹ der Studienarm Exenatide+Metformin wurde in der Meta-Analyse nicht berücksichtigt, da die Dosierung (2 mg Exenatide [einmal wöchentlich]+Metformin 1,5 g) zum Zeitpunkt der Durchführung der Meta-Analyse nicht zugelassen war.

² entspricht der Publikation zum Konferenz-Abstract von Salvadeo, 2010

* Studien, die durch aktualisierte Studienrecherche gefunden wurden

³ Es wurden nur die Pioglitazon 30 mg and 45 mg Behandlungsgruppen für die Nutzenbewertung herangezogen. Im Folgenden werden daher nur noch diese Behandlungsgruppen dargestellt.

Tabelle 4-80: Charakterisierung der Interventionen – Studien für indirekte Vergleiche

Studie	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
MB102014	10 mg Dapagliflozin+ 1,8 g Metformin	Placebo+1,8 g Metformin			Vorbehandlung: 2-wöchige einfachblinde Placebo-lead-in-Phase 10 mg Dapagliflozin, 1,8 g Metformin täglich
D1690C00012	10 mg Dapagliflozin+ 1,5-2,5 g Metformin	Placebo+ 1,5-2,5 g Metformin			Vorbehandlung: 2-wöchige einfachblinde Placebo-lead-in-Phase 10 mg Dapagliflozin, 1,5-2,5 g Metformin täglich
Scott 2008	100 mg Sitagliptin+ ≥1,5 g Metformin	Placebo+≥1,5 g Metformin			100 mg Sitagliptin, ≥1,5 g Metformin
Charbonnel 2006	100 mg Sitagliptin+ ≥1,5 g Metformin	≥1,5 g Metformin+ Placebo			Vorbehandlung: 2-wöchige einfachblinde Placebo-lead-in-Phase 100 mg Sitagliptin, ≥1,5 g Metformin
DeFronzo 2009	5 mg Saxagliptin+ 1,5-2,5 g Metformin	Placebo+ 1,5-2,5 g Metformin			Vorbehandlung: 2-wöchige einfachblinde Placebo-lead-in-Phase 5 mg Saxagliptin, 1,5-2,5 g Metformin (gleiche Dosis wie bei Einschluss).
Bosi 2007	100 mg Vildagliptin+ 2 g Metformin	Placebo+ 2 g Metformin			100 mg Vildagliptin, 2 g Metformin
Taskinen 2011	5 mg Linagliptin+ ≥1,5 g Metformin	Placebo+≥1,5 g Metformin			Vorbehandlung: 2-wöchige open-label lead-in-Phase 5 mg Linagliptin, ≥1,5 g Metformin

Studie	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
Bergental 2010a	100 mg Sitagliptin+ ≥1,5 g Metformin	Placebo+≥1,5 g Metformin			100 g Sitagliptin, ≥1,5 g Metformin
Raz 2008	100 g Sitagliptin+ 1,5-2,5 g Metformin	Placebo+ 1,5-2,5 g Metformin			Vorbehandlung: 2-wöchige <i>open-label lead-in</i> Phase 100 g Sitagliptin, 1,5-2,5 g Metformin
Bolli 2008	100 g Vildagliptin+ ≥1,5 g Metformin	Pioglitazon+ ≥1,5 g Metformin			100 mg Vildagliptin, ≥1,5 g Metformin,
Bergental 2010b	2 mg Exenatide (einmal wöchentlich) +1,5 g Metformin ¹	100 mg Sitagliptin +1,5 g Metformin	45 mg Pioglitazon +1,5 g Metformin		100 mg Sitagliptin oder 45 mg Pioglitazon+ 1,5 g Metformin
DeFronzo 2005	5 µg Exenatide (BID)	10 µg Exenatide (BID)	Placebo		4-wöchige einfachblinde <i>lead-in</i> -Phase mit subkutaner Injektion von Placebo BID 10 bzw. 20 µg Exenatide und ≥1,5 g Metformin
Nauck 2009	1,2 mg Liraglutide+ 2 g Metformin	1,8 mg Liraglutide+ 2 g Metformin	4 mg Glimepirid+ 2 g Metformin	Placebo+ 2 g Metformin	Titrationphase: 2 Wochen für 1,2 mg Liraglutide, 3 Wochen für 1,8 mg Liraglutid, sowie 4 mg Glimepirid 1,2 bzw. 1,8 mg Liraglutid oder 4 mg Glimepirid mit 2 g Metformin
Pratley 2010	1,2 mg Liraglutide+ Metformin	1,8 mg Liraglutide+ Metformin	100 mg Sitagliptin+ Metformin		Titrationphase: 2 Wochen für 1,2 mg Liraglutide, 3 Wochen für 1,8 mg Liraglutide Metformin-Dosis: N/R

Studie	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	ggf. weitere Spalten mit Behandlungs- charakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
Kaku 2009	30 mg Pioglitazon+ 500-750 mg Metformin	Placebo+ 500-750 mg Metformin			Titrationsphase: 12 Wochen für Pioglitazon 500-750 mg Metformin
Charpentier 2001	1-6 mg Glimepirid+ 2,55 g Metformin	Placebo+2,55 g Metformin			Titrationsphase: 12 Wochen
Papathanassiou 2009	30 mg Pioglitazon+ Metformin	4 mg Glimepirid+ Metformin			
D1690C00004	10 mg Dapagliflozin+ 1,5-2,5 g Metformin	5-20 mg Glipizid+ 1,5-2,5 g Metformin			Vorbehandlung: 2- wöchige <i>open-label</i> <i>lead-in</i> -Phase Titrationsphase: 18 Wochen 10 mg Dapagliflozin oder 5-20 mg Glipizid und 1,5-2,5 g Metformin (wie bei Einschluss)
Matthews 2005	30-45 mg Pioglitazon+ 0,3-0,5 g Metformin	160-320 mg Gliclazid+ 0,3-0,5 g Metformin			Titrationsphase: 16 Wochen 30-45 mg Pioglitazon oder 160-320 mg Gliclazide+0,3-0,5 g Metformin (wie bei Einschluss)
Nauck 2007	100 mg Sitagliptin +Metformin	5-20 mg Glipizid +Metformin			Titrationsphase: 18 Wochen Vorbehandlung: 2- wöchige <i>open-label</i> <i>lead-in</i> -Phase mit Sitagliptin
Göke 2010	5 mg Saxagliptin +1,5-3 g Metformin	5-20 mg Glipizid+1,5-3 g Metformin			Titrationsphase: 18 Wochen Vorbehandlung: 2- wöchige <i>open-label</i> <i>lead-in</i> -Phase mit Saxagliptin
Filozof 2010	100 mg Vildagliptin+ ≥1,5 g Metformin	80-320 mg Gliclazid+ ≥1,5 g Metformin			Titrationsphase: 12 Wochen

Studie	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	ggf. weitere Spalten mit Behandlungs- charakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
Salvadeo 2010	20 µg Exenatide +1,5-2 g Metformin	6 mg Glimepirid+ 1,5-2 g Metformin			Titrationsphase: 4 Wochen
Derosa 2010	20 µg Exenatide+ 1-2 g Metformin	15 mg Glibenclamid+ 1-2 g Metformin			Titrationsphase: 4 Wochen
Matthews 2010	2-6 mg Glimepirid+ ≥1,5 g Metformin	100 mg Vildagliptin+ ≥1,5 g Metformin 1			Titration während der gesamten Studiendauer
Yang 2011*	5 mg Saxagliptin+ 1,5-3 g Metformin	Placebo+ 1,5-3 g Metformin			Placebo <i>Lead-in</i> Phase: 2 Wochen
Derosa 2011a*	45 mg Pioglitazon +Metformin	15 mg Glibenclamid+ Metformin			
Derosa 2011b* ²	20 µg Exanatide +Metformin (1-2g)	6 mg Glimepirid +Metformin (1-2g)			Titrationsphase: 4 Wochen
Pan 2012*	Vildagliptin 50 mg + Metformin 1500 mg	Vildagliptin 100 mg + Metformin 1500 mg	Placebo + Metformin		<i>Lead-in</i> Phase: 2 Wochen
DeFronzo 2012* ³	Pioglitazon (30 mg / 45 mg) + Metformin (≥1500 mg)	Placebo+ Metformin (≥1500 mg)			<i>Screening</i> -Periode: 2 Wochen <i>Run in</i> / Stabilisierungsperiode ;4 Wochen Behandlungsdauer: 26 Wochen <i>Follow-up</i> -Periode; 2 Wochen
Gallwitz 2012*	Linagliptin 5mg + Metformin (≥1500 mg)	Glimepirid +Metformin (≥1500 mg)			<i>Run-in</i> Phase: 2 Wochen (mit oder ohne ein er 6 wöchigen Wash-out- Phase) Behandlungsdauer:

Studie	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	ggf. weitere Spalten mit Behandlungs- charakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
					104 Wochen Glimepirid war die Anfangsdosis 1 mg; dann wurde diese schrittweise erhöht bis zu 4 mg) Notfallmedikation: Pioglitanzone Follow-up-Periode; 1 Woche
Yang 2012*	100 mg Sitagliptin+ Metformin (1000 oder 1700 mg)	Placebo+ Metformin (1000 oder 1700 mg)			Screening-Phase: 1 Woche Bis zu 9 Wochen Metformin Hochtitrierungs- oder Stabilisierungsphase und Diät und Exercise-Phase

BID: zweimal täglich;

¹ der Studienarm Exenatide+Metformin wurde in der Meta-Analyse nicht berücksichtigt, da die Dosierung (2 mg Exenatide [einmal wöchentlich] +Metformin 1,5g) zum Zeitpunkt der Durchführung der Meta-Analyse nicht zugelassen war.

² entspricht der Publikation zum Konferenz-Abstract von Salvadeo, 2010

* Studien, die durch aktualisierte Studienrecherche gefunden wurden

³ Es wurden nur die Pioglitazon 30 mg and 45 mg Behandlungsgruppen für die Nutzenbewertung herangezogen.

Tabelle 4-81: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD/SE)	Geschlecht w / m (%)	BMI (kg/m²) Mittelwert (SD/SE)
MB102014				
Dapagliflozin+Metformin	135	52,7 (9,9)	43/57	31,2 (5,1)
Placebo+Metformin	137	53,7 (10,3)	45/55	31,8 (5,3)
D1690C00012				
Dapagliflozin+Metformin	91	60,6 (8,16)	45/55	32,06 (3,887)
Placebo+Metformin	91	60,8 (6,82)	44/56	31,68 (3,89)
Scott 2008				
Sitagliptin+Metformin	94	55,2 (9,8)	45/55	30,3 (4,7)
Placebo+Metformin	91	55,3 (9,3)	41/59	30,0 (4,5)
Charbonnel 2006				
Sitagliptin+Metformin	464	54,4 (10,4)	44/56	30,9 (5,3)
Placebo+Metformin	237	54,7 (9,7)	41/59	31,5 (4,9)
DeFronzo 2009				
Saxagliptin+Metformin	191	54,7 (9,6)	46/54	31,2 (4,7)
Placebo+Metformin	192	54,8 (10,2)	46/54	31,6 (4,8)
Bosi 2007				
Vildagliptin+Metformin	183	53,9 (9,5)	38/62	32,9 (5,0)
Placebo+Metformin	181	54,5 (10,3)	47/53	33,2 (6,1)
Taskinen 2011				
Linagliptin+Metformin	523	56,5 (10,1)	47/53	29,58 (4,84)
Placebo+Metformin	177	56,6 (10,9)	43/57	30,05 (5,01)
Bergental 2010a				
Sitagliptin+Metformin	190	56 (gesamte Studien- population)	n. b.	25-45 (Untergrenze 23 für Asiaten)
Placebo+Metformin	95		n. b.	32 (gesamte Studien- population)
Raz 2008				
Sitagliptin+Metformin	96	53,6 (9,5)	49/51	30,1 (4,4)
Placebo+Metformin	94	56,1 (9,5)	59/41	30,4 (5,3)
Bolli 2008				
Vildagliptin+Metformin	295	56,3 (9,3)	38/62	32,2 (5,6)
Placebo+Metformin	280	57,0 (9,7)	36/64	32,1 (5,1)
Bergental 2010b				
Sitagliptin+Metformin	166	52 (11)	48,2/51,8	32 (5)
Pioglitazone+Metformin	165	53 (10)	52,1/47,9	32 (6)
Exenatide+Metformin ¹	160	52 (10)	44,4/55,6	32 (5)

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD/SE)	Geschlecht w / m (%)	BMI (kg/m²) Mittelwert (SD/SE)
DeFronzo 2005				
Exenatide BID (5 µg)	110	53 (11)	48/52	34 (6)
Exenatide BID (10 µg)	113	52 (11)	40/60	34 (6)
Placebo	113	54 (9)	41/59	34 (6)
Nauck 2009				
Liraglutide (1,2 mg)+Metformin	241	57 (9)	46/54	31,1 (4,8)
Liraglutide (1,2 mg)+Metformin	242	57 (9)	41/59	30,9 (4,6)
Glimepirid+Metformin	244	57 (9)	43/57	31,2 (4,6)
Placebo+Metformin	122	56 (9)	40/60	31,6 (4,4)
Pratley 2010				
Liraglutide (1,2 mg)+Metformin	255	55,9 (9,6)	48/52	32,6 (5,2)
Liraglutide (1,2 mg)+Metformin	221	55 (9,1)	48/52	33,1 (5,1)
Sitagliptin+Metformin	219	55 (9,0)	45/55	32,6 (5,4)
Kaku 2009				
Pioglitazon+Metformin	83	52 (8,6)	34/66	25,6 (4,2)
Placebo+Metformin	86	53 (7,5)	43/57	25,4 (3,6)
Charpentier 2001				
Glimepirid+Metformin	147	56,8	41/59	29,5
Placebo+Metformin	75	56,7	40/60	29,2
Papathanassiou 2009				
Glimepirid+Metformin	14	63,6 (7,3)	79/21	31,9 (5,5)
Pioglitazone+Metformin	14	62,8 (7,2)	79/21	33,9 (7,0)
D1690C00004				
Dapagliflozin+Metformin	400	58,1 (9,37)	45/55	31,71 (5,10)
Glipizid+Metformin	408	58,6 (9,80)	45/55	31,23 (5,05)
Matthews 2005				
Gliclazid+Metformin	313	57 (9,0)	51/49	32,6 (5,8)
Pioglitazon+Metformin	317	56 (9,2)	49/51	32,6 (5,0)
Nauck 2007				
Glipizid+Metformin	584	56,6 (9,8)	39/61	31,3 (5,2)
Sitagliptin+Metformin	588	56,8 (9,3)	43/57	31,2 (5,0)
Göke 2010				
Glipizid+Metformin	430	57,6 (10,37)	46/54	31,3 (6,17)
Saxagliptin+Metformin	428	57,5 (10,26)	50/50	31,5 (5,70)
Filozof 2010				
	1007			
Gliclazid+Metformin		59,7 (10,2)	48/52	30,8 (5,0)
Vildagliptin+Metformin		59,2 (9,9)	48/52	31,2 (5,0)

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD/SE)	Geschlecht w / m (%)	BMI (kg/m²) Mittelwert (SD/SE)
Salvadeo 2010 Exenatide+Metformin Glimepirid+Metformin		n. b.	n. b.	28,4 (1,3) 28,5 (1,4)
Derosa 2010 Exenatide+Metformin Glibenclamid+Metformin	128	>18 57 (8) 56 (7)	52/48 49/51	28,7 (1,5) 28,5 (1,4)
Matthews 2010 Glimepirid+Metformin Vildagliptin+Metformin	1383 1389	57,5 (9,19) 57,5 (9,07)	46/54 47/53	31,7 (5,3) 31,9 (5,3)
Yang 2011* Saxagliptin+Metformin Placebo+Metformin	283 287	53,8 (10,4) 54,4 (10,1)	51,9/48,1 51,6/48,4	26,3 (3,6) 26,1 (3,5)
Derosa 2011a* Pioglitazone+Metformin Glibenclamid+Metformin	102 99	55,8 (7,9) 56,9 (8,8)	50/50 48,5/51,5	27,8 (2,4) 28,2 (3,1)
Derosa 2011b* ² Exenatid+Metformin Glimepirid+Metformin	57 54	56 (7) 55 (6)	50,9/49,1 51,9/48,1	28,4 (1,3) 28,5 (1,4)
Pan 2012* Vildagliptin 50 mg+Metformin Vildagliptin 100 mg+Metformin Placebo+Metformin	148 146 144	53,7 (10,0) 54,2 (9,62) 54,5 (9,68)	55,4/44,6 50/50 54,2/45,8	25,03 (3,09) 26,01 (3,26) 25,46 (3,09)
DeFronzo 2012* ³ Pioglitazone 30 mg+Metformin Pioglitazone 45 mg+Metformin Placebo+Metformin	129 129 129	56,1 (9,4) 54,5 (9,7) 55,2 (9,9)	51,2 / 48,8 58,9 / 41,1 52,7 /47,3	31,4 (5,4) 30,7 (4,7) 30,6 (4,8)
Gallwitz 2012* Linagliptin+Metformin Glimepirid+Metformin	776 775	59,8 (9,4) 59,8 (9,4)	60 / 40 61 / 39	30,2 (4,8) 30,3 (4,6)
Yang 2012* Sitagliptin+Metformin Placebo+Metformin	197 198	54,1 (9,0) 55,1 (9,8)	53 / 47 45 / 55	25,3 (3,1) 25,3 (3,6)

SD: Standardabweichung; SE: Standardfehler; n. b.: nicht berichtet;

¹ der Studienarm Exenatide+Metformin wurde in der Meta-Analyse nicht berücksichtigt, da die Dosierung (2 mg Exenatide [einmal wöchentlich] +Metformin 1,5 g) zum Zeitpunkt der Durchführung der Meta-Analyse nicht zugelassen war.

² entspricht der Publikation zum Konferenz-Abstract von Salvadeo, 2010

* Studien, die durch aktualisierte Studienrecherche gefunden wurden

³ Es wurden nur die Pioglitazon 30 mg and 45 mg Behandlungsgruppen für die Nutzenbewertung herangezogen..

Tabelle 4-82: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Gewicht (kg) Mittelwert (SD/SE)	Baseline HbA1c (%) Mittelwert (SD/SE)	Dauer der Diabetes- Erkrankung (Jahre) Mittelwert (SD/SE)
MB102014				
Dapagliflozin+Metformin	135	N/R	7,92 (0,82)	6,1 (5,4)
Placebo+Metformin	137	N/R	8,11 (0,96)	5,8 (5,1)
D1690C00012				
Dapagliflozin+Metformin	91	92,06 (14,128)	7,19 (0,443)	6,03 (4,534)
Placebo+Metformin	91	90,91 (13,716)	7,16 (0,531)	5,52 (5,266)
Scott 2008				
Sitagliptin+Metformin	94	83,1 (17,1)	7,75 (0,99)	4,9 (3,5)
Placebo+Metformin	91	84,6 (16,5)	7,68 (0,88)	5,4 (3,7)
Charbonnel 2006				
Sitagliptin+Metformin	464	86,7 (17,8)	7,96 (0,81)	6,0 (5,0)
Placebo+Metformin	237	89,6 (17,5)	8,03 (0,82)	6,6 (5,5)
DeFronzo 2009				
Saxagliptin+Metformin	191	87,3 (17,0)	8,1 (0,8)	6,4 (4,7)
Placebo+Metformin	192	87,1 (17,8)	8,1 (0,9)	6,7 (5,6)
Bosi 2007				
Vildagliptin+Metformin	183	95,3 (20,4)	8,4 (1,0)	5,8 (4,7)
Placebo+Metformin	181	94,8 (24,5)	8,3 (0,9)	6,2 (5,3)
Taskinen 2011				
Linagliptin+Metformin	523	82,2 (17,2)	8,09 (0,86)	>5
Placebo+Metformin	177	83,3 (16,6)	8,02 (0,88)	>5
Bergenstal 2010a				
Sitagliptin+Metformin	190	92,48	7,94	n. b.
Placebo+Metformin	95	91,09	8,03	n. b.
Raz 2008				
Sitagliptin+Metformin	96	81,5 (16,8)	9,3 (0,9)	8,4 (6,5)
Placebo+Metformin	94	81,2 (19,4)	9,1 (0,8)	7,3 (5,3)
Bolli 2008				
Vildagliptin+Metformin	295	91,8 (18,5)	8,4 (1,0)	6,4 (4,9)
Placebo+Metformin	280	91,2 (16,9)	8,4 (0,9)	6,4 (5,2)
Bergenstal 2010b				
Sitagliptin+Metformin	166	87 (20)	8,5 (1,2)	5 (4)
Pioglitazon+Metformin	165	88 (20)	8,5 (1,1)	6 (5)
Exenatide+Metformin ¹	160	89 (20)	8,6 (1,2)	6 (5)

Studie Gruppe	N	Gewicht (kg) Mittelwert (SD/SE)	Baseline HbA1c (%) Mittelwert (SD/SE)	Dauer der Diabetes- Erkrankung (Jahre) Mittelwert (SD/SE)
DeFronzo 2005				
Exenatide BID (5 µg)	110	100 (22)	8,3 (1,1)	6.2 (5.9)
Exenatide BID (10 µg)	113	101 (20)	8,2 (1,0)	4.9 (4.7)
Placebo	113	100 (19)	8,2 (1,0)	6.6 (6.1)
Nauck 2009				
Liraglutide (1,2 mg)+Metformin	241	n. b.	8,3 (1,0)	7 (5)
Liraglutide (1,2 mg)+Metformin	242		8,4 (1,0)	8 (5)
Glimepirid+Metformin	244		8,4 (1,0)	8 (5)
Placebo+Metformin	122		8,4 (1,1)	8 (6)
Pratley 2010				
Liraglutide (1,2 mg)+Metformin	255	93,7 (18,4)	8,4 (0,8)	6.0 (4.5)
Liraglutide (1,2 mg)+Metformin	221	94,6 (18,1)	8,4 (0,7)	6.4 (5.4)
Sitagliptin+Metformin	219	93,1 (18,9)	8,5 (0,7)	6.3 (5.4)
Kaku 2009				
Pioglitazon+Metformin	83	n. b.	7,58 (1,0)	4.5 (3.7)
Placebo+Metformin	86		7,55 (0,9)	5.6 (5.0)
Charpentier 2001				
Glimepirid+Metformin	147	81,2	6,4 (1,1)	5,6
Placebo+Metformin	75	82,2	6,8 (1,2)	7,0
Papathanassiou 2009	28			
Glimepirid+Metformin	14	81,4 (15,3)	7,4 (0,8)	5,3 (6,5)
Pioglitazone+Metformin	14	85,9 (18,7)	7,7 (0,7)	5,3 (3,6)
D1690C00004				
Dapagliflozin+Metformin	400	88,44 (16,323)	7,69 (0,86)	6,08 (4,61)
Glipizid+Metformin	408	87,60 (16,97)	7,74 (0,89)	6,55 (5,90)
Matthews 2005				
Gliclazid+Metformin	313	92,7 (17,4)	8,53 (0,89)	5,5 (5,1)
Pioglitazon+Metformin	317	91,8 (16,2)	8,71 (1,00)	5,8 (5,1)
Nauck 2007				
Glipizid+Metformin	584	89,7 (17,5)	7,6 (0,9)	6,2 (5,4)
Sitagliptin+Metformin	588	89,5 (17,4)	7,7 (0,9)	6,5 (6,1)
Göke 2010				
Glipizid+Metformin	430	88,6 (19,64)	7,7 (0,9)	5,4 (4,7)
Saxagliptin+Metformin	428	88,7 (18,61)	7,7 (0,9)	5,5 (4,5)
Filozof 2010	1007			
Gliclazid+Metformin		84,2 (17,9)	8,5 (1,0)	6,8 (5,3)
Vildagliptin+Metformin		85,7 (16,6)	8,5 (1,0)	6,4 (5,1)

Studie Gruppe	N	Gewicht (kg) Mittelwert (SD/SE)	Baseline HbA1c (%) Mittelwert (SD/SE)	Dauer der Diabetes- Erkrankung (Jahre) Mittelwert (SD/SE)
Salvadeo 2010 Exenatide+Metformin Glimepirid+Metformin		n. b.	n. b.	n. b.
Derosa 2010 Exenatide+Metformin Glibenclamid+Metformin	128	82,0 (8,3) 82,4 (9,1)	8,8 (0,7) 8,9 (0,8)	n. b. n. b.
Matthews,2010 Glimepirid+Metformin Vildagliptin+Metformin	1383 1389	88,9(17,8) 89,5 (18,1)	7,3 (0,7) 7,3 (0,7)	5,7 (5,0) 5,7 (5,2)
Yang 2011* Saxagliptin+Metformin Placebo+Metformin	283 287	68,9 (12,5) 69,0 (11,9)	7,9 (0,8) 7,9 (0,8)	5,1 (5,0) 5,1 (4,0)
Derosa 2011a* Pioglitazon+Metformin Glibenclamid+Metformin	102 99	79,6 (8,9) 78,9 (8,3)	7,4 (1,1) 7,5 (1,2)	n. b.
Derosa 2011b* ² Exenatide+Metformin Glimepirid+Metformin	57 54	80,2 (7,5) 81,4 (8,1)	8,7 (0,7) 8,8 (0,8)	n. b.
Pan 2012* Vildagliptin 50 mg+Metformin Vildagliptin 100 mg+Metformin Placebo+Metformin	148 146 144	68,36 (11,10) 71,58 (11,93) 69,83 (11,18)	8,05 (0,84) 8,09 (0,85) 8,01 (0,82)	5,02 (4,42) 4,92 (4,80) 5,15 (4,58)
DeFronzo 2012* ³ Pioglitazone 30mg+Metformin Pioglitazone 45mg+Metformin Placebo+Metformin	129 129 129	81,6 – 85,6	8,5 (0,7) 8,5 (0,7) 8,5 (0,6)	7,6 (7,1) 5,7 (4,2) 6,0 (5,0)
Gallwitz 2012* Linagliptin+Metformin Glimepirid+Metformin	776 775	86,1 (17,6) 86,8 (16,7)	7,7 (0,9) 7,7 (0,9)	n. b.
Yang 2012* Sitagliptin+Metformin Placebo+Metformin	197 198	67,9 (10,7) 68,9 (13,3)	8,5 (0,9) 8,5 (0,9)	6,4 (4,4) 7,3 (4,6)

SD: Standardabweichung; SE: Standardfehler; n. b.: nicht berichtet;

¹ der Studienarm Exenatide+Metformin wurde in der Meta-Analyse nicht berücksichtigt, da die Dosierung (2 mg Exenatide [einmal wöchentlich]+Metformin 1,5 g) zum Zeitpunkt der Durchführung der Meta-Analyse nicht zugelassen war.

² entspricht der Publikation zum Konferenz-Abstract von Salvadeo, 2010

* Studien, die durch aktualisierte Studienrecherche gefunden wurden

³ Es wurden nur die Pioglitazon 30 mg and 45 mg Behandlungsgruppen für die Nutzenbewertung herangezogen.

Beschreiben Sie die Studien zusammenfassend. Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

In die vorliegende Netzwerk-Meta-Analyse wurden ausschließlich RCT einbezogen, da diese mit der geringsten Ergebnisunsicherheit behaftet sind. Die Studien waren größtenteils doppelt verblindet. Eine Studie war einfachblind (Derosa et al., 2010) und bei einer Studie wurde die Verblindung nicht angegeben (Salvadeo et al., 2010). Zwei Studien waren unverblindet (Papathanassiou et al., 2009; Pratley et al., 2010); hiervon stellt nur die Studie von Pratley (2010) einen Vergleich mit einem DPP-4-Inhibitor dar. Es wurden sowohl aktivkontrollierte als auch placebokontrollierte Studien ausgewertet. In allen Studien wurden Antidiabetika als *Add-on*-Therapie zu Metformin untersucht. Die Studienteilnehmer waren erwachsene Männer und Frauen mit Typ-2-Diabetes, deren Blutzuckerwert mit Metformin allein nur unzureichend kontrolliert werden konnte. Der Bereich der HbA1c-Werte, mit denen Patienten eingeschlossen wurden, war in den verschiedenen Studien unterschiedlich, der Gesamtbereich aller Studien lag zwischen 6,5 und 11%.

Alle Studienteilnehmer waren mit Metformin vorbehandelt. Die Metformin-Behandlung wurde während der Studien fortgesetzt, wobei in den meisten Fällen die Dosierung der Vorbehandlung beibehalten wurde. Die tägliche Metformin-Dosis lag zwischen 500 und 2550 mg.

Folgende Antidiabetika wurden in den berücksichtigten Studien untersucht:

- Sulfonylharnstoffe (Glimepirid, Glibenclamid, Glipizid, Gliclazid) nur in der 52-wöchigen Netzwerkanalyse (Ausnahme: Endpunkt systolischer Blutdruck)
- Pioglitazon
- DPP-4-Inhibitoren (Sitagliptin, Vildagliptin, Saxagliptin, Linagliptin)
- GLP-1-Analoga (Exenatide, Liraglutide)
- Dapagliflozin

Endpunkte, die in die Netzwerk-Meta-Analyse einbezogen wurden, sind:

- HbA1c-Wert im Vergleich zum Ausgangswert (23 Studien)
- Anteil der Patienten, die einen HbA1c-Wert von <7,0% erreichen (18 Studien)
- Anteil der Patienten, die einen HbA1c-Wert von <6,5% erreichen (12 Studien)
- Gewichtsveränderung im Vergleich zum Ausgangswert (20 Studien)
- Veränderung des systolischen Blutdrucks im Vergleich zum Ausgangswert (10 Studien)

- Anzahl und Anteil der Patienten, bei denen mindestens eine Hypoglykämie auftrat (18 Studien)
- Gesamtrate unerwünschter Ereignisse, Gesamtrate schwerwiegender unerwünschter Ereignisse, Studienabbrüche aufgrund unerwünschter Ereignisse

Das Durchschnittsalter der Patienten lag zwischen 52 und 64 Jahren. Im Allgemeinen war das Durchschnittsalter der Behandlungsgruppen recht ähnlich (maximal 2,5 Jahre Unterschied, in den meisten Fällen maximal ein Jahr).

In den meisten Studien wurden etwas mehr Männer als Frauen oder etwa gleich viele Männer wie Frauen untersucht. Hiervon weicht nur die Studie von Papathanassiou (2009) mit einem Frauenanteil von 79% auffallend ab. Innerhalb der meisten Studien war das Verhältnis der Geschlechter zwischen den Behandlungsgruppen ungefähr gleich. Nur in den Studien von Raz (2008), Bosi (2007), DeFronzo (2005) und Kaku (2009) gibt es größere Unterschiede zwischen den Gruppen.

In der Studie von Kaku (2009) lag der durchschnittliche BMI der Patienten in der Nähe des Normalgewichts, ansonsten lag der BMI der Patienten um die 30 oder darüber. Es gab keine auffallenden Unterschiede im BMI zwischen den Behandlungsgruppen.

Die Diabetesdauer der Patienten lag meist bei etwa 5-8 Jahren. Es gab keine auffälligen Unterschiede zwischen den Studien oder zwischen den Behandlungsgruppen der einzelnen Studien.

In die Netzwerk-Meta-Analyse wurden nur Studien mit einer Dauer von mindestens 18 Wochen einbezogen.

4.3.2.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-83: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
MB102014	ja	ja	ja	ja	nein	nein	niedrig
D1690C00012	ja	ja	ja	ja	nein	nein	niedrig
Scott 2008	unklar	unklar	unklar	unklar	nein	nein	niedrig
Charbonnel 2006	unklar	unklar	unklar	unklar	nein	nein	niedrig
DeFronzo 2009	ja	ja	ja	ja	nein	nein	niedrig
Bosi 2007	unklar	unklar	unklar	unklar	nein	nein	hoch
Taskinen 2011	unklar	unklar	unklar	unklar	nein	ja	niedrig
Bergenstal 2010a	unklar	unklar	ja	ja	unklar	ja	hoch
Raz 2008	unklar	unklar	unklar	unklar	nein	nein	niedrig
Bolli 2008	ja	ja	ja	ja	nein	nein	niedrig
Bergenstal 2010b	ja	ja	ja	ja	nein	nein	niedrig
DeFronzo 2005	unklar	unklar	unklar	unklar	nein	nein	niedrig
Nauck 2009	ja	unklar	unklar	unklar	nein	nein	niedrig
Pratley 2010	ja	ja	nein	nein	nein	nein	niedrig
Kaku 2009	unklar	unklar	unklar	unklar	unklar	ja	hoch
Charpentier 2001	ja	ja	unklar	unklar	unklar	nein	hoch
Papathanassiou 2009	nein	unklar	nein	nein	unklar	ja	hoch
D1690C00004	ja	ja	ja	ja	nein	nein	niedrig
Matthews 2005	ja	ja	ja	ja	nein	nein	niedrig
Nauck 2007	unklar	unklar	ja	ja	nein	nein	niedrig
Göke 2010	ja	ja	ja	ja	nein	nein	niedrig
Filozof 2010	unklar	unklar	ja	ja	nein	nein	niedrig
Salvadeo 2010	unklar	unklar	unklar	unklar	unklar	ja	hoch
Derosa 2010	ja	ja	ja	nein	nein	nein	niedrig
Matthews 2010	unklar	unklar	ja	ja	nein	nein	niedrig
Yang 2011*	ja	ja	ja	ja	nein	nein	niedrig
Derosa 2011a*	ja	ja	ja	ja	nein	ja	niedrig
Derosa 2011b* ^o	ja	ja	ja	ja	nein	ja	niedrig
Pan 2012*	unklar	unklar	ja	ja	nein	nein	niedrig

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
DeFronzo 2012*	unklar	unklar	unklar	unklar	nein	nein	niedrig
Gallwitz 2012*	ja	ja	ja	ja	nein	nein	niedrig
Yang 2012*	ja	unklar	unklar	unklar	nein	nein	niedrig

*aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

°Publikation zum Abstract von Salvadeo 2010

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Publikationen DeFronzo 2009, Bolli 2008, Bergenstal 2010b, Matthews 2005, Göke, 2010 sowie die Studienberichte MB102014, D1690C00012, D1690C00004 und Yang 2011 beschreiben alle die Ergebnisse randomisierter kontrollierter Studien. In diesen Publikationen war jeweils die Erzeugung der Randomisierungssequenz adäquat beschrieben, die Verdeckung der Gruppenzugehörigkeit sowie die Verblindung der Patienten und der Behandler gewährleistet. Auch gab es in diesen Publikationen weder Hinweise auf ergebnisgesteuerte Berichterstattung noch finden sich sonstige das Verzerrungspotenzial beeinflussende Punkte.

Bei Scott 2008, Charbonnel 2006, Bosi 2007, Raz 2008 und DeFronzo 2005 handelte es sich ebenfalls jeweils um eine randomisierte kontrollierte Studie. Es wurden jedoch keine Angaben zur Randomisierung und Verblindung gemacht, es wurde jedoch jeweils von einer doppelblinden Studie berichtet. Es wurde für diese Publikation angenommen, dass das Verzerrungspotenzial niedrig ist.

Bei Taskinen 2011 wurden ebenfalls keine Angaben zur Radomisierung und Verblindung der Studie gemacht. Es wurde jedoch eine randomisierte kontrollierte doppelblinde Studie beschrieben. Aus der Publikation ging nicht hervor, wie mit den Werten der Patienten umgegangen worden ist, die eine Notfallmedikation erhalten hatten. Die Studie war trotz des unklaren Punktes als niedrig verzerrt anzusehen.

Bei Bergenstal 2010a wurden nicht genügend Information angegeben, um eine Verzerrung auszuschließen. Daher wurde diese Publikation als hoch verzerrt bewertet.

Pratley 2010 war eine offene Studie, die Randomisierungssequenz war jedoch adäquat beschrieben und die Gruppenzuteilung erfolgte verdeckt. Weitere Punkte, die eine Verzerrung verursachen können, waren nicht beschrieben, daher wurde die Verzerrung der Publikation als niedrig bewertet.

Kaku 2009 stellte keine Information zur Randomisierung und Verblindung zur Verfügung. Ergebnisstesteuerte Berichtserstattung konnte nicht ausgeschlossen werden, da positive Subgruppenergebnisse selektiv in der Publikation beschrieben wurden. Auch waren Patienten in der Behandlungsgruppe signifikant kürzer an Diabetes erkrankt, was ebenfalls eine Verzerrung verursachen kann. Die Publikation wurde daher als hoch verzerrt angenommen.

Bei Charpentier 2001 fehlten adäquate methodische Information, die relevant für die Abschätzung des Verzerrungspotenzials sind (z. B. Verblindung, Kalkulation der Studiengröße). Deshalb konnte eine Verzerrung nicht ausgeschlossen werden.

Auch bei der offenen Studie von Papathanassiou 2009 fehlten adäquate methodische Informationen, die relevant für die Abschätzung des Verzerrungspotenzials sind. Daher konnte ebenfalls eine Verzerrung nicht ausgeschlossen werden. Bei Nauck 2007, Filozof 2010, Matthews 2010 und Pan 2012 fehlten Information zur Erzeugung der Randomisierungssequenz und zur Verdeckung der Gruppenzugehörigkeit. Ansonsten gab es keine Hinweise auf ein Verzerrungspotenzial.

Salvadeo 2010 war die Publikation des Abstracts der Studie. Die erhältlichen Informationen waren nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht, daher konnte eine Verzerrung nicht ausgeschlossen werden. Derosa 2011b ist die Publikation von dem Konferenzabstract von Salvadeo 2010. Die Publikation von Derosa 2011b bestätigt bzw. ergänzt die Ergebnisse und Angaben zur Studie von Salvadeo 2010. Durch die ergänzende Information kann daher ein niedriges Verzerrungspotenzial angenommen werden.

Derosa 2010 war eine einfachblinde Studie (Patienten sind verblindet, jedoch nicht der Behandler). Weitere Hinweise auf das Verzerrungspotenzial beeinflussende Punkte konnten nicht gefunden werden, daher wurde ein niedriges Verzerrungspotenzial angenommen.

Derosa 2011a war eine doppelblinde Studie. Weitere Hinweise auf das Verzerrungspotenzial beeinflussende Punkte konnten nicht gefunden werden, daher wurde ein niedriges Verzerrungspotenzial angenommen.

In der Studie von DeFronzo 2012 waren keine Information über die Erzeugung der Randomisierungssequenz sowie über die Verdeckung der Gruppenzugehörigkeit angegeben. Auch wurde die Studie zwar als doppelblind beschrieben, aber es wurden keine Angaben gemacht, wer genau verblindet gewesen ist. Es wurden aber keine Hinweise auf das Verzerrungspotenzial beeinflussende Punkte gefunden. Daher kann ein niedriges Verzerrungspotenzial angenommen werden.

Die Studie von Gallwitz 2012 beschreibt die Ergebnisse einer randomisierten aktiv-kontrollierten Studie. In dieser Publikation war die Erzeugung der Randomisierungssequenz adäquat beschrieben und die Verdeckung der Gruppenzugehörigkeit sowie die Verblindung der Patienten und der Behandler gewährleistet. Auch gab es weder Hinweise auf

ergebnisgesteuerte Berichtserstattung noch finden sich sonstige das Verzerrungspotenzial beeinflussende Punkte. Die Studie kann daher als niedrig verzerrt bewertet werden.

Bei Yang 2012 war die Erzeugung der Randomisierungssequenz adäquat beschrieben, jedoch gab es keine Angaben zur Verdeckung der Gruppenzugehörigkeit. Des weiteren wurde die Studie als doppelblind beschrieben, jedoch gab es keine expliziten Angaben wer verblindet worden ist. Auch gab es weder Hinweise auf ergebnisgesteuerte Berichtserstattung noch fanden sich sonstige das Verzerrungspotenzial beeinflussende Punkte. Daher wird von einer niedrig verzerrten Studie ausgegangen.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-84: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	Hypo- glykämien	Gewichts- veränder- ung	Un- erwünschte Ereignisse	Folge- komplika- tionen	Gesamt- mortalität	Systolischer Blutdruck
MB102014	Ja	Ja	Ja	Nein	Nein	Ja
D1690C00012	Ja	Ja	Ja	Nein	Nein	Nein
Scott 2008	Ja	Ja	Ja	Nein	Nein	Nein
Charbonnel 2006	Ja	Nein	Ja	Nein	Nein	Nein
DeFronzo 2009	Ja	Ja	Ja	Nein	Nein	Ja
Bosi 2007	Ja	Ja	Ja	Nein	Nein	Nein
Taskinen 2011	Ja	Ja	Ja	Nein	Nein	Nein
Bergental 2010a	Ja	Ja	Ja ²	Nein	Nein	Nein
Raz 2008	Ja	Ja	Ja	Nein	Nein	Nein
Bolli 2008	Ja	Ja	Ja	Nein	Nein	Nein
Bergental 2010b	Nein	Ja	Nein	Nein	Nein	Ja
DeFronzo 2005	Ja	Ja	Ja ¹	Nein	Nein	Nein
Nauck 2009	Ja	Ja	Nein	Nein	Nein	Ja
Pratley 2010	Ja	Ja	Ja	Nein	Nein	Ja
Kaku 2009	Ja	Ja	Ja	Nein	Nein	Nein
Charpentier 2001	Nein	Nein	Nein	Nein	Nein	Ja
Papathanassiou 2009	Nein	Nein	Nein	Nein	Nein	Ja
D1690C00004	Ja	Ja	Ja	Nein	Nein	Ja
Matthews 2005	Ja	Ja	Ja	Nein	Nein	Nein
Nauck, 2007	Ja	Nein	Ja	Nein	Nein	Nein
Göke 2010	Ja	Ja	Ja	Nein	Nein	Ja
Filozof 2010	Nein	Ja	Ja	Nein	Nein	Nein
Salvadeo 2010	Nein	Nein	Nein	Nein	Nein	Nein
Derosa 2010	Nein	Ja	Ja ¹	Nein	Nein	Nein
Matthews 2010	Ja	Ja	Nein	Nein	Nein	Nein
Yang 2011*	Ja	Ja	Ja ¹	Nein	Nein	Ja
Derosa 2011a*	Nein	Nein	Ja ¹	Nein	Nein	Nein
Derosa 2011b* °	Nein	Nein	Ja ¹	Nein	Nein	Nein

Studie	Hypo- glykämien	Gewichts- veränder- ung	Un- erwünschte Ereignisse	Folge- komplika- tionen	Gesamt- mortalität	Systolischer Blutdruck
Pan 2012*	Ja	Ja	Ja ¹	Nein	Nein	Nein
DeFronzo 2012*	Ja	Ja	Ja	Nein	Nein	Nein
Gallwitz 2012*	Ja	Ja	Ja	Nein	Nein	Nein
Yang 2012*	Ja	Ja	Ja	Nein	Nein	Nein

¹: Studie berichtet nur Studienabbrüche,

²: Studie berichtet nur schwerwiegend unerwünschte Ereignisse und Studienabbrüche

*aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

^oPublikation zum Abstract von Salvadeo 2010

Tabelle 4-85: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	Hyperosmolare Komata	Stationäre Behandlungen	Gesundheitsbezogene Lebensqualität	Therapiezufriedenheit	Ausmaß der Blutzuckersenkung (HbA1c)
MB102014	Nein	Nein	Nein	Nein	Ja
D1690C00012	Nein	Nein	Nein	Nein	Ja
Scott 2008	Nein	Nein	Nein	Nein	Ja
Charbonnel 2006	Nein	Nein	Nein	Nein	Ja
DeFronzo 2009	Nein	Nein	Nein	Nein	Ja
Bosi 2007	Nein	Nein	Nein	Nein	Ja
Taskinen 2011	Nein	Nein	Nein	Nein	Ja
Bergental 2010a	Nein	Nein	Nein	Nein	Ja
Raz 2008	Nein	Nein	Nein	Nein	Ja
Bolli 2008	Nein	Nein	Nein	Nein	Ja
Bergental 2010b	Nein	Nein	Nein	Nein	Ja
DeFronzo 2005	Nein	Nein	Nein	Nein	Ja
Nauck 2009	Nein	Nein	Nein	Nein	Ja
Pratley 2010	Nein	Nein	Nein	Nein	Ja
Kaku 2009	Nein	Nein	Nein	Nein	Ja
Charpentier 2001	Nein	Nein	Nein	Nein	Nein
Papathanassiou 2009	Nein	Nein	Nein	Nein	Nein
D1690C00004	Nein	Nein	Nein	Nein	Ja
Matthews 2005	Nein	Nein	Nein	Nein	Ja
Nauck 2007	Nein	Nein	Nein	Nein	Ja
Göke 2010	Nein	Nein	Nein	Nein	Ja
Filozof 2010	Nein	Nein	Nein	Nein	Ja
Salvadeo 2010	Nein	Nein	Nein	Nein	Ja
Derosa 2010	Nein	Nein	Nein	Nein	Ja
Matthews 2010	Nein	Nein	Nein	Nein	Ja
Yang 2011*	Nein	Nein	Nein	Nein	Ja
Derosa 2011a*	Nein	Nein	Nein	Nein	Ja
Derosa 2011b* ^o	Nein	Nein	Nein	Nein	Ja

Studie	Hyperosmolare Komata	Stationäre Behandlungen	Gesundheitsbezogene Lebensqualität	Therapiezufriedenheit	Ausmaß der Blutzuckersenkung (HbA1c)
Pan 2012*	Nein	Nein	Nein	Nein	Ja
De Fronzo 2012*	Nein	Nein	Nein	Nein	Ja
Gallwitz 2012*	Nein	Nein	Nein	Nein	Ja
Yang 2012*	Nein	Nein	Nein	Nein	Ja

*aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

°Publikation zum Abstract von Salvadeo 2010

4.3.2.1.3.1 Ausmaß der Blutzuckerkontrolle (HbA1c) – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-86: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin +Placebo	Metformin +TZD	Metformin+ DPP-4-Inhibitoren	Metformin+ GLP-1-Analoga
2	MB102014 D1690C00012	•	•			
1	Kaku 2009		•	•		
10	Charbonnel 2006 DeFronzo 2009 Raz 2008 Bosi 2007 Taskinen 2011 Scott 2008 Bergenstal 2010a Yang 2011° Pan 2012° Yang 2012°		•		•	
1	DeFronzo 2005*		•			•
1	Bolli 2008			•	•	
1	Bergenstal 2010b			•	•	
1	Pratley 2010*				•	•
1	Nauck 2009*		•			•

*GLP-1-Analoga jeweils mit 2 verschiedenen Dosierungen

°aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Tabelle 4-87: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin+ Sulfonylharnstoff	Metformin+ DPP-4-Inhibitoren	Metformin +TZD	Metformin +GLP-1-Analoga
1	D1690C000 04	•	•			
2	Matthews 2005 Derosa 2011a °^ **		•		•	
5	Matthews 2010 Nauck 2007 Göke 2010 Filozof 2010 Gallwitz 2012°		•	•		
2	Salvadeo 2010** Derosa 2010**		•			•

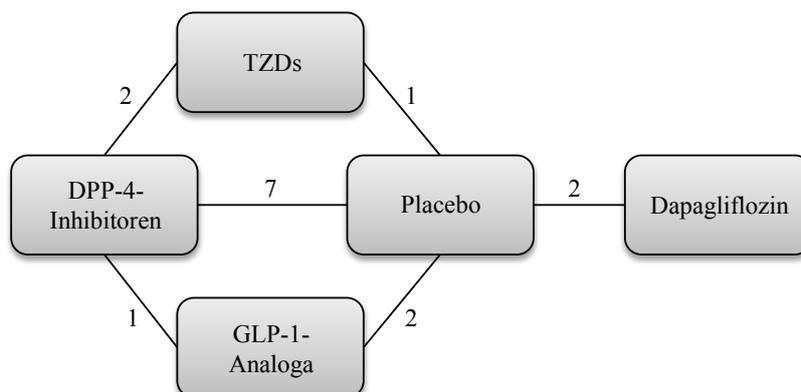
** werden nicht berücksichtigt, da ein intensives Diätprogramm Bestandteil der Studien waren → Sensitivitätsanalyse

°aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen) ^Publikation zum Abstract von Salvadeo 2010

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

A



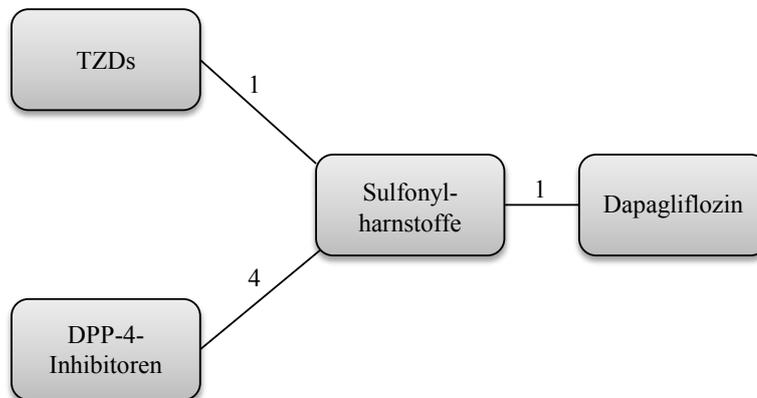
B

Abbildung 12: Netzwerk der Studien für Endpunkt „Ausmaß der Blutzuckerkontrolle“.

A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk

*die Zahlen repräsentieren die Anzahl der RCT

Es wurden 2 Netzwerke beschrieben (Abbildung 12), die aufgrund der Verfügbarkeit der Endpunkte und der Studiendauer für 24 (Netzwerk A) und 52 (Netzwerk B) Wochen unterschiedlich waren. Wie in 4.3.2.1.1 erwähnt, wurden Sulfonylharnstoffe und Glinide aus dem 24-Wochen Netzwerk wegen nicht konstanter Effektmaße innerhalb einer Studiendauer von 18-30 Wochen und einer Titrationsperiode von bis zu 18 Wochen (bei 24 Wochen Interim-Ergebnissen 52-wöchiger Studien), mit Ausnahme für den Endpunkt „Blutdrucksenkung“ (um für diesen Endpunkt die Evidenz zu erhöhen und da nicht von einem *J-curve* Verlauf dieses Endpunkts ausgegangen wird), ausgeschlossen. Im Netzwerk A sind alle RCT mit einer Studiendauer von 24±6 Wochen, im Netzwerk B alle RCT mit einer Studiendauer von 52±6 Wochen dargestellt. Für den Endpunkt Ausmaß der Blutzuckerkontrolle wurden 15 RCT nach 24 Wochen und 8 RCT nach 52 Wochen identifiziert. 2 RCT (Derosa 2010, Salvadeo 2010) des 52-wöchigen Netzwerks wurden aus der Analyse ausgeschlossen (Ausschlussgrund: hoher Gewichtsverlust aufgrund intensiven Diät- und Bewegungsprogramms) und lediglich bei den Sensitivitätsanalysen eingeschlossen.

Im Folgenden werden die Einzelergebnisse aller relevanten Studien tabellarisch beschrieben und gegebenenfalls, falls die Studien zusammenfassbar waren, ein Gesamtschätzer angegeben. Da DPP-4-Inhibitoren die zweckmäßige Vergleichstherapie für Patienten, die eine Unverträglichkeit/Kontraindikation gegenüber Sulfonylharnstoffen haben, darstellen, wurde der indirekte Vergleich für Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren beschrieben (jeweils nach 24 und nach 52 Wochen).

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-88: Operationalisierung von Ausmaß der Blutzuckerkontrolle (HbA1c)

Studie	Operationalisierung
MB102014	Anhand der Veränderung des HbA1c nach 24 Wochen
D1690C00012	Anhand der Veränderung des HbA1c nach 24 Wochen
Kaku 2009	Anhand der Veränderung des HbA1c nach 28 Wochen
Charbonnel 2006	Anhand der Veränderung des HbA1c nach 24 Wochen
DeFronzo 2009	Anhand der Veränderung des HbA1c nach 24 Wochen
Raz 2008	Anhand der Veränderung des HbA1c nach 30 Wochen
Bosi 2007	Anhand der Veränderung des HbA1c nach 24 Wochen
Taskinen 2011	Anhand der Veränderung des HbA1c nach 24 Wochen
Scott 2008	Anhand der Veränderung des HbA1c nach 18 Wochen
Bergenstal 2010a	Anhand der Veränderung des HbA1c nach 24 Wochen
DeFronzo 2005	Anhand der Veränderung des HbA1c nach 30 Wochen
Bolli 2008	Anhand der Veränderung des HbA1c nach 24 Wochen
Bergenstal 2010b	Anhand der Veränderung des HbA1c nach 26 Wochen
Pratley 2010	Anhand der Veränderung des HbA1c nach 26 Wochen
Nauck 2009	Anhand der Veränderung des HbA1c nach 26 Wochen
D1690C00004	Anhand der Veränderung des HbA1c nach 52 Wochen
Matthews 2005	Anhand der Veränderung des HbA1c nach 52 Wochen
Matthews 2010	Anhand der Veränderung des HbA1c nach 52 Wochen
Nauck 2007	Anhand der Veränderung des HbA1c nach 52 Wochen
Göke 2010	Anhand der Veränderung des HbA1c nach 52 Wochen
Filozof 2010	Anhand der Veränderung des HbA1c nach 52 Wochen
Salvadeo 2010	Anhand der Veränderung des HbA1c nach 52 Wochen
Derosa 2010	Anhand der Veränderung des HbA1c nach 52 Wochen
Yang 2011*	Anhand der Veränderung des HbA1c nach 24 Wochen
Derosa 2011a*	Anhand der Veränderung des HbA1c nach 52 Wochen
Derosa 2011b*°	Anhand der Veränderung des HbA1c nach 52 Wochen
Pan 2012*	Anhand der Veränderung des HbA1c nach 24 Wochen
DeFronzo 2012*	Anhand der Veränderung des HbA1c nach 26 Wochen
Gallwitz 2012*	Anhand der Veränderung des HbA1c nach 52 Wochen
Yang 2012*	Anhand der Veränderung des HbA1c nach 24 Wochen

*aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

° Publikation zum Abstract von Salvadeo 2010

HbA1c: Glykiertes Hämoglobin

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-89: Bewertung des Verzerrungspotenzials für Ausmaß der Blutzuckerkontrolle (HbA1c) in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MB102014	niedrig	ja	ja	nein	nein	niedrig
D1690C00012	niedrig	ja	ja	nein	nein	niedrig
Kaku 2009	hoch	unklar	ja	nein	ja	hoch
Charbonnel 2006	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2009	niedrig	unklar	unklar	nein	nein	niedrig
Raz 2008	niedrig	unklar	ja	nein	nein	niedrig
Bosi 2007	hoch	ja	unklar	nein	nein	hoch
Taskinen 2011	niedrig	unklar	ja	nein	ja	hoch
Scott 2008	niedrig	ja	ja	nein	nein	niedrig
Bergenstal 2010a	hoch	ja	unklar	unklar	ja	hoch
DeFronzo 2005	niedrig	unklar	ja	nein	nein	niedrig
Bolli 2008	niedrig	ja	nein	nein	nein	niedrig
Bergenstal 2010b	niedrig	ja	ja	nein	nein	niedrig
Pratley 2010	niedrig	nein	ja	nein	nein	niedrig
Nauck 2009	niedrig	unklar	ja	nein	nein	niedrig
D1690C00004	niedrig	ja	ja	nein	nein	niedrig
Matthews 2005	niedrig	ja	ja	nein	nein	niedrig
Matthews 2010	niedrig	ja	nein	nein	nein	niedrig
Nauck 2007	niedrig	unklar	unklar	nein	nein	niedrig
Göke 2010	niedrig	ja	ja	nein	nein	niedrig
Filozof 2010	niedrig	ja	nein	nein	nein	niedrig
Salvadeo 2010	hoch	unklar	unklar	unklar	ja	hoch
Derosa 2010	niedrig	nein	unklar	nein	ja	niedrig
Yang 2011*	niedrig	ja	ja	nein	nein	niedrig

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Derosa 2011a*	niedrig	ja	ja	nein	nein	niedrig
Derosa 2011b* ^o	niedrig	nein	unklar	nein	ja	niedrig
Pan 2012*	niedrig	unklar	ja	nein	nein	niedrig
DeFonzo 2012*	niedrig	unklar	ja	nein	nein	niedrig
Gallwitz 2012*	niedrig	ja	ja	nein	nein	niedrig
Yang 2012*	niedrig	unklar	ja	nein	nein	niedrig

*aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

^oPublikation zum Abstract von Salvadeo 2010

Begründen Sie für jede Studie die abschließende Einschätzung.

In allen Studien wurde der Endpunkt nahezu gleich operationalisiert. Bei den Studien MB102014, D1690C00012, Charbonnel 2006, Scott 2008, Bergenstal 2010b, D1690C00004, Matthews 2005, Göke 2010, Yang 2011 und Gallwitz 2012 (Yang 2011 und Gallwitz 2012 wurden später durch die aktualisierte Literaturrecherche identifiziert) war das Verzerrungspotenzial des Endpunktes als niedrig anzusehen. Der Endpunkterheber war verblindet, das ITT-Prinzip wurde adäquat umgesetzt, es wurden weder Hinweise auf ergebnisgesteuerte Berichterstattung noch auf sonstige das Verzerrungspotenzial beeinflussende Punkte gefunden.

Die Studie von Kaku 2009 war bereits auf Studienebene hoch verzerrt; es lagen keine Informationen über die Verblindung des Endpunkterheber vor und die Ergebnisse des Endpunktes wurden möglicherweise aufgrund des unterschiedlichen Alters und der unterschiedlichen Diabetesdauer der Patienten in den Behandlungsgruppen verzerrt.

Bei DeFronzo 2009 und Nauck 2007 lagen weder Informationen über die Verblindung des Endpunkterheber noch über die adäquate Umsetzung des ITT-Prinzips vor; es wurde aber von einem niedrigen Verzerrungspotenzial ausgegangen.

Die Studie von Bosi 2007 war bereits auf Studienebene hoch verzerrt, die Verblindung des Endpunkterhebers war gewährleistet, jedoch wurde das ITT-Prinzip unzureichend dargestellt, so dass dieser Endpunkt als hoch verzerrt anzusehen ist.

Bei den Publikationen von Taskinen 2011 und DeFronzo 2012 lagen keine Information über die Verblindung des Endpunkterhebers vor, jedoch wurde das ITT-Prinzip adäquat umgesetzt. Weiterhin war bei Taskinen 2012 unklar, ob die LOCF-Methode bei Patienten, die ein

Notfallmedikament erhalten hatten, angewendet worden ist. Daher wurde dieser Endpunkt bei Taskinen 2012 als hoch verzerrt angesehen. Bei DeFronzo 2012 kann dieser Endpunkt als niedrig verzerrt angesehen werden.

In der Studie von Bergenstal 2010a wurden nicht genügend Informationen zur Verfügung gestellt, um eine Verzerrung ausschließen zu können, daher wurde der Endpunkt als hoch verzerrt zu bewerten.

In den Studien von Raz 2008, DeFronzo 2005, Nauck 2009, Pan 2012 und Yang 2012 (Pan 2012 und Yang 2012 wurden später durch die aktualisierte Literaturrecherche identifiziert) lagen keine Information über die Verblindung des Endpunkterhebers vor, aber da es sonst keine Hinweise auf eine Verzerrung gab, konnte dieser Endpunkt als niedrig bewertet werden.

In der Publikation von Bolli 2008 wurde keine adäquate Umsetzung des ITT-Prinzips durchgeführt, die Analyse des Endpunktes wurde mit der per Protokoll-Analyse durchgeführt, da sie aber ausführlich beschrieben wurde und es sonst keine Hinweise auf eine Verzerrung gab, wurde dieser Endpunkt als niedrig verzerrt eingestuft.

In der Publikation von Pratley 2010 gab es keine Verblindung des Endpunkterhebers. Es wurden keine weiteren Hinweise, die auf eine Verzerrung hindeuten, gefunden.

Bei Matthews 2010 wurde die per Protokoll-Population zur Berechnung dieses Endpunktes genutzt. Die ITT-Population wurde für die Sensitivitätsanalyse verwendet. Daher wurde das ITT-Prinzip nicht adäquat umgesetzt. Auch in der Studie von Filozof 2010 beruhte die Analyse des Endpunkts auf der per Protokoll-Population. Die ITT-Population wurde für die Sensitivitätsanalyse (Ergebnisse wurden als vergleichbar angegeben, aber nicht berichtet) verwendet. Aber da es keine weiteren Hinweise auf eine Verzerrung gab, wurde der Endpunkt jeweils als niedrig verzerrt eingestuft.

Da in der Studie von Salvadeo 2010 nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um das Verzerrungspotenzial der Ergebnisse zu beurteilen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht, daher wurde das Verzerrungspotenzial als hoch bewertet. In der aktualisierten Literaturrecherche wurde die Publikation von Derosa 2011b identifiziert. Diese stellt die Publikation zu dem Konferenz-Abstract von Salvadeo 2010 dar. Daher stehen mehr Informationen zur Verfügung, so dass das Verzerrungspotenzial als niedrig einzustufen ist.

In der Publikation von Derosa 2010 handelte sich um eine einfachblinde Studie, der Endpunkterheber war nicht verblindet. Patienten mussten ein intensives Diät- und Sportprogramm durchführen, daher ist der Vergleich mit den anderen Studien mit Vorsicht zu betrachten. Da es aber keine weiteren Hinweise auf eine Verzerrung gab, wurde dieser Endpunkt als niedrig verzerrt eingestuft.

Bei der Studie von Derosa 2011a handelt es sich um eine doppelblinde Studie. Wie in den Studien von Derosa 2010 und Salvadeo 2010 mussten Patienten ein intensives Diät- und Sportprogramm durchführen, daher ist der Vergleich mit den anderen Studien mit Vorsicht zu

betrachten. Da es aber keine weiteren Hinweise auf eine Verzerrung gab, wurde dieser Endpunkt als niedrig verzerrt eingestuft.

In der Studie von DeFonzo 2012 war das Ausmaß der Blutzuckerkontrolle der primäre Endpunkt. Es ergaben sich keine Hinweise auf eine Verzerrung der Studie.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-90: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT für indirekte Vergleiche (nach 24 Wochen)

Studien	N	Veränderung des HbA1c Mittelwert (SD) %	N	Veränderung des HbA1c Mittelwert (SD) %	Gruppenunterschied Mittelwert (SE) [95%-KI]
	Metformin+DPP-4-Inhibitoren		Metformin+Placebo		
Charbonnel 2006	453	-0,67 (0,05)	224	-0,02 (0,06)	-0,65 [-0,80;-0,50]
DeFronzo 2009	186	-0,69 (0,07)	175	0,13 (0,07)	-0,82 [-1,01;-0,63]
Raz 2008	95	-1,00 (0,2)	95	0,00 (0,2)	-1,00 [-1,55;-0,45]
Bosi 2007	143	-0,9 (0,1)	130	0,2 (0,1)	-1,10 [-1,38;-0,82]
Taskinen 20011	513	-0,49 (0,04)	175	0,15 (0,06)	-0,64 [-0,78;-0,50]
Scott 2008	91	-0,73 (0,07)	88	-0,22 (0,07)	-0,51 [-0,70;-0,32]
Bergental 2010a	177	-0,889 (0,057)	90	-0,1 (0,079)	-0,79 [-0,98;-0,60]
Gesamtschätzer (mit Heterogenitätsmaß)^o					-0,74 [-0,87;-0,61] I ² =62,9%; p=0,013 ^o
	Metformin+TZD		Metformin+Placebo		
Kaku 2009	83	-0,67 (0,08781)	86	0,25 (0,09921)	-0,92 [-1,18;-0,66]
	Metformin+Dapagliflozin		Metformin+Placebo		
MB102014	132	-0,84 (0,07)	134	-0,3 (0,07)	-0,54 [-0,73;-0,35]
D1690C00012	88	-0,39 (0,0485)	91	-0,1 (0,0477)	-0,29 [-0,42;-0,16]
Gesamtschätzer (mit Heterogenitätsmaß)^o					-0,40 [-0,85;-0,16] I ² =76,9%; p=0,037 ^o
	Metformin+GLP-1-Analoga		Metformin+Placebo		
DeFronzo 2005*	110	-0,4 (0,1)	113	0,1 (0,1)	-0,70 [-0,98;-0,42]
Nauck 2009*	240	-0,97 (0,1)	121	0,09 (0,1)	-1,08 [-1,35;-0,80]
Gesamtschätzer (mit Heterogenitätsmaß)^o					-0,89 [-1,25;-0,52] I ² =71,0%; p=0,063 ^o
	Metformin+TZD		Metformin+DPP-4-Inhibitoren		
Bolli 2008	246	-0,98 (0,06)	264	-0,88 (0,05)	-0,10 [-0,25;0,05]
Bergental 2010b	165	-1,2 (0,1)	166	-0,9 (0,1)	-0,30 [-0,58;-0,02]

Studien	N	Veränderung des HbA1c Mittelwert (SD) %	N	Veränderung des HbA1c Mittelwert (SD) %	Gruppenunterschied Mittelwert (SE) [95%-KI]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					-0,17 [0,35;0,02] I ² =34,8%; p=0,216 [°]
	Metformin+DPP-4-Inhibitoren		Metformin+GLP-1-Analoga		
Pratley 2010*	246	-0,9 (0,07)	219	-0,9 (0,07)	-0,47 [-0,66;-0,28]

* Zwei verschiedene Dosierungen jeweils von GLP-1-Analoga (die Mittelwerte der beiden Dosierungsgruppen wurden gebildet)

[°] anhand Modell mit zufälligen Effekten

KI: Konfidenzintervall, SE: Standardfehler, SD: Standardabweichung, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4, HbA1c: Glykiertes Hämoglobin

Tabelle 4-91: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT für indirekte Vergleiche (nach 52 Wochen)

Studien	N	Veränderung des HbA1c Mittelwert (SE) %	N	Veränderung des HbA1c Mittelwert (SE) %	Gruppenunterschied Mittelwert [95%-KI]
	Metformin+DPP-4-Inhibitor		Metformin+Sulfonylharnstoff		
Matthews 2010	1118	-0,44 (0,02)	1072	-0,53 (0,02)	0,09 [0,03;0,15]
Nauck 2007	576	-0,51 (0,04)	559	-0,56 (0,05)	0,05 [-0,08;0,18]
Göke 2010	423	-0,57 (0,039)	423	-0,66 (0,039)	0,09 [-0,02;0,20]
Filozof 2010	386	-0,81 (0,06)	393	-0,85 (0,06)	0,04 [-0,13;0,21]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					0,08 [0,04;0,13] I ² =0,0% p=0,898
	Metformin+TZD		Metformin+Sulfonylharnstoff		
Matthews 2005	317	-0,99 (0,05)	313	-1,01 (0,03)	0,02 [-0,09;0,13]
	Metformin+Dapagliflozin		Metformin+Sulfonylharnstoff		
D1690C00004	400	-0,52 (0,04)	401	-0,52 (0,04)	0,00 [-0,11;0,11]
	Metformin+GLP-1-Analoga		Metformin+Sulfonylharnstoff		
Derosa 2011b** °°	65	-1,2 (0,06)	65	-1,4 (0,05)	0,20 [0,05;0,35]
Derosa 2010**	63	-1,5 (0,0959)	65	-1,8 (0,1023)	0,30 [0,03;0,57]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					0,22 [0,09;0,36] I ² =0%; p=0,53

** werden nicht in die primäre Analyse nicht miteinbezogen → Sensitivitätsanalyse

[°] anhand Modell mit zufälligen Effekten; ^{°°} entspricht dem Abstract von Salvadeo 2010

KI: Konfidenzintervall, SE: Standardfehler, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4, HbA1c: Glykiertes Hämoglobin

Patienten, die eine Placebo-Behandlung zu ihrer existierenden Metformin+Behandlung erhielten, wiesen keine bedeutsame Veränderung ihres HbA1c-Wertes nach 24 Wochen auf.

Wurde eine aktive Behandlung mit der Placebo-Behandlung verglichen, war jede aktive Behandlung signifikant besser als eine Placebo-Behandlung jeweils in Kombination mit Metformin. Gesamtschätzer wurden mittels eines Modells mit zufälligen Effekten berechnet, auffällig ist die hohe Heterogenität zwischen den Studien, die Ergebnisse zeigten immer den gleichen Effekt, die Heterogenität war daher nur quantitativer Natur.

In der aktualisierten Literaturrecherche wurden zusätzlich weitere Publikationen (Yang 2011, Pan 2012, Derosa 2011a, Derosa 2011b, Gallwitz 2012, De Fronzo 2012 und Yang 2012) gefunden, welche diesen Endpunkt „Ausmaß der Blutzuckerkontrolle“ beschreiben. Yang 2011, Pan 2012 und Yang 2012 beschreiben den Vergleich Metformin+DPP-4-Inhibitor versus Metformin+Placebo nach 24 Wochen und liefern ähnliche Ergebnisse. Derosa 2011a beschreibt den Vergleich Metformin+TZD versus Metformin+Sulfonylharnstoff nach 52 Wochen und liefert ebenfalls ähnliche Ergebnisse. Gallwitz beschreibt den Vergleich Metformin+DPP-4-Inhibitoren versus Metformin+Sulfonylharnstoff und liefert ebenfalls ein ähnliches Ergebnis wie die Studien, die in die MTC eingeschlossen wurden.

Tabelle 4-92: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT für indirekte Vergleiche – aktualisierte Literaturrecherche

Studien	N	Veränderung des HbA1c in % Mittelwert (SE) / [95%-KI]	N	Veränderung des HbA1c in % Mittelwert (SE) / [95%-KI]	Gruppenunterschied Mittelwert [95%-KI]
Nach 24 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Placebo		
Yang 2011	275	-0,78 (0,1)	279	-0,37 (0,1)	-0,41 [-0,69;-0,13]
Pan 2012	145	-1,05 (0,08)	144	-0,45 (0,08)	-0,60 [-0,82; -0,38]
Yang 2012	191	-1,0 [-1,2;-0,9]	194	-0,1 [-0,3;0,0]	-0,90 [-1,1; -0,7]
	Metformin+TZD		Metformin+Placebo		
	129	-0,9 (0,1)	129	-0,1 (0,1)	-0,85 [-1,13;-0,17]
Nach 52 Wochen					
	Metformin+TZD		Metformin+Sulfonylharnstoff		
Derosa 2011a **	102	-1,0 (0,1)	99	-0,8 (0,1)	0,20 [-0,48;0,08]
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff		
Galwitz 2012	764	-0,38 (0,03)	755	-0,60 (0,03)	0,22 [0,13;0,31]

** werden nicht in die primäre Analyse nicht miteinbezogen → Sensitivitätsanalyse

KI: Konfidenzintervall, SE: Standardfehler, TZD: Thiazolidinedione, DPP-4: Dipeptidyl-Peptidase-4, HbA1c: Glykiertes Hämoglobin

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei

an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.

Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.

Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.

Tabelle 4-93: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus indirekten Vergleiche nach 24 und 52 Wochen – Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel

Zu bewertende Arzneimittel	Komparator	Mittelwertdifferenz (95%-Kreditintervall)*
Nach 24 Wochen		
Metformin+Dapagliflozin	Metformin+DPP-4-Inhibitoren	0,16 [-0,17;0,49]
Nach 52 Wochen		
Metformin+Dapagliflozin	Metformin+DPP-4-Inhibitoren	-0,08 [-0,25;0,10]

*basierend auf dem best angepassten Modell

DPP-4: Dipeptidyl-Peptidase-4;

Es zeigte sich kein statistisch signifikantes Ergebnis im Vergleich von DPP-4-Inhibitoren zu Dapagliflozin jeweils als *Add-on* zu Metformin.

Um die Annahme der Konsistenz zu überprüfen, wurden die Resultate der *Head-to-Head*-Vergleiche mit den Resultaten aus den MTC verglichen. Es wurden keine Inkonsistenzen festgestellt.

Nimmt man die Studien von Salvadeo 2010 (bzw. Derosa 2011b), die durch die aktualisierte Literaturrecherche identifiziert wurde, und Derosa 2010, die ein intensives Diät- und Sportprogramm verlangten, in die Analyse der MTC mit auf, ergaben sich keine relevanten Unterschiede in den Effektschätzern. Die Publikation von Derosa 2011a, welche in der aktualisierten Literaturrecherche gefunden wurde, würde ebenfalls in die Sensitivitätsanalyse der MTC einfließen. Jedoch würde sie keinen Unterschied in der Gesamtinterpretation machen, da der Effekt Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren in der in die MTC einfließende Publikation (Matthews 2005) gleichgerichtet ist.

Würde die Netzwerk-Meta-Analyse nach 24 Wochen mit den Publikationen von Yang 2011, Pan 2012 und Yang 2012 aktualisiert werden, würden keine Unterschiede in der Gesamtinterpretation erwartet werden. Auch würde sich nichts an der Gesamtinterpretation der MTC nach 52 Wochen ändern, falls die Publikation von Gallwitz 2012 noch hinzugefügt werden würde.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.1.1 Hypoglykämien – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-94: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin+ Placebo	Metformin+ TZD	Metformin+ DPP-4-Inhibitoren	Metformin+ GLP-1-Analoga
2	MB102014 D1690C00012	•	•			
1	Kaku 2009		•	•		
9	Charbonnel 2006 DeFronzo 2009 Raz 2008 Bosi 2007 Taskinen 2011 Scott 2008 Yang 2011° Pan 2012° Yang 2012°		•		•	
1	DeFronzo 2005*		•			•
2	Bolli 2008 Bergenstal 2010b			•	•	
1	Pratley 2010*				•	•

*GLP-1-Analoga jeweils mit 2 verschiedenen Dosierungen

°aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

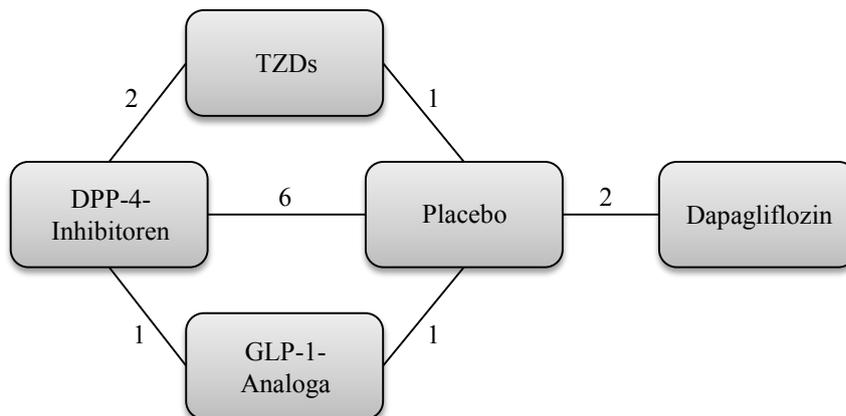
Tabelle 4-95: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin+ Sulfonylharnstoff	Metformin+ DPP-4-Inhibitoren	Metformin+ TZD	Metformin + GLP-1-Analoga
1	D1690C00004	•	•			
1	Matthews 2005		•		•	
4	Matthews 2010 Nauck 2007 Göke 2010 Gallwitz 2012		•	•		

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

A



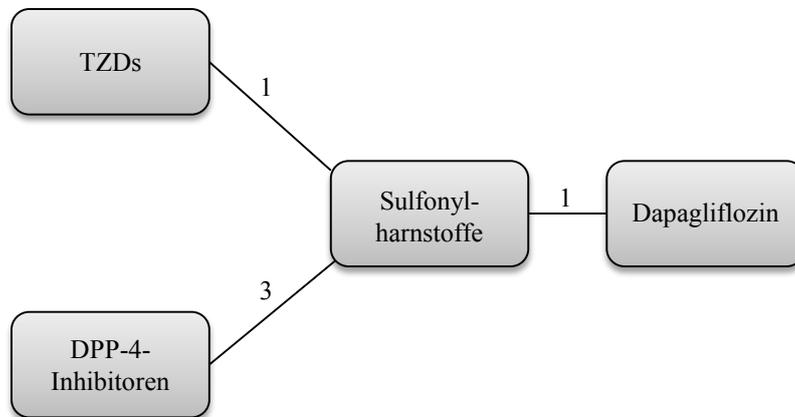
B

Abbildung 13: Netzwerk der Studien für Endpunkt Hypoglykämien. A: 24 Wochen Netzwerk.
B: 52 Wochen Netzwerk

* die Zahlen repräsentieren die Anzahl der RCT

Es wurden 2 Netzwerke beschrieben (Abbildung 13), die aufgrund der Verfügbarkeit der Endpunkte und Studiendauer unterschiedlich für 24 (Netzwerk A) und 52 (Netzwerk B) Wochen waren. Im Netzwerk A sind alle RCT mit einer Studiendauer von 24±6 Wochen, im Netzwerk B alle RCT mit einer Studiendauer von 52±6 Wochen dargestellt. Hypoglykämien wurden in 13 RCT nach 24 Wochen und in 5 RCT nach 52 Wochen berichtet.

Im Folgenden werden die Einzelergebnisse aller relevanten Studien tabellarisch beschrieben und gegebenenfalls, falls die Studien zusammenfassbar waren, ein Gesamtschätzer angegeben. Anschließend wird der indirekte Vergleich für Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren beschrieben (jeweils nach 24 und nach 52 Wochen).

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-96: Operationalisierung von Hypoglykämien

Studie	Operationalisierung
MB102014	Anzahl Hypoglykämien / Leichte Hypoglykämien: nicht definiert Schwerwiegende Hypoglykämien: externe Hilfe, Blutzuckerwert: <3 mmol/L
D1690C00012	Anzahl Hypoglykämien / leichte Hypoglykämien: symptomatische / asymptotische mit oder ohne bestätigter Blutzuckermessung <3,5 mmol/L, die nicht schwerwiegend sind Schwerwiegende Hypoglykämien: externe Hilfe, Blutzuckerwert: < 3 mmol/L
Kaku 2009	Nicht näher definiert
Charbonnel 2006	Nicht näher definiert

Studie	Operationalisierung
DeFronzo 2009	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch ohne Bestätigung, symptomatisch mit Bestätigung (Blutzuckerwert < 50 mg/dL)
Raz 2008	Nicht näher definiert
Bosi 2007	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch mit Bestätigung (<3,1 mmol/L)
Taskinen 2011	Anzahl Hypoglykämien / Leichte Hypoglykämien: asymptomatisch / symptomatisch mit Bestätigung (<3 mmol/L)
Scott 2008	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch mit Bestätigung (<3,1 mmol/L) Schwerwiegende Hypoglykämien: externe Hilfe, Blutzuckerwert: <3,1 mmol/L
DeFronzo 2005	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch ohne Bestätigung, symptomatisch mit Bestätigung Schwerwiegende: externe Hilfe
Bolli 2008	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch mit Bestätigung (<3,1 mmol/L)
Bergental 2010b	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch mit Bestätigung (<3 mmol/L) Schwerwiegende Hypoglykämien: externe Hilfe, Blutzuckerwert: <3,1 mmol/L
Pratley 2010	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch mit Bestätigung (<3,1 mmol/L) und Selbstbehandlung Schwerwiegende Hypoglykämien: externe Hilfe (ungeachtet des Blutzuckerwertes)
D1690C00004	Anzahl Hypoglykämien: symptomatisch/asymptomatisch mit oder ohne Blutzuckerwert von <3,5 mmol/L, inklusive der schwerwiegenden
Matthews 2005	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch ohne Bestätigung
Matthews 2010	Anzahl Hypoglykämien / Leichte Hypoglykämien: symptomatisch mit Bestätigung (<3,1 mmol/L)
Nauck 2007	Anzahl Hypoglykämien: symptomatisch ohne Bestätigung
Göke 2010	Anzahl Hypoglykämien: Leichte Hypoglykämien symptomatisch ohne Bestätigung / asymptomatisch mit Bestätigung (<2,8 mmol/L)
Yang 2011*	Anzahl Hypoglykämien: symptomatisch oder asymptomatisch mit oder ohne Bestätigung
Pan 2012*	Anzahl Hypoglykämien: symptomatisch oder asymptomatisch mit Bestätigung (<56 mg/dL)
DeFronzo 2012*	Nicht näher definiert
Gallwitz 2012*	Nicht näher definiert
Yang 2012*	Anzahl Hypoglykämien: symptomatisch oder asymptomatisch mit Bestätigung

*aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-97: Bewertung des Verzerrungspotenzials für Hypoglykämien in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MB102014	niedrig	ja	ja	nein	nein	niedrig
D1690C00012	niedrig	ja	ja	nein	nein	niedrig
Charbonnel 2006	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2009	niedrig	unklar	ja	nein	nein	niedrig
Raz 2008	niedrig	unklar	ja	nein	nein	niedrig
Bosi 2007	hoch	ja	ja	nein	nein	hoch
Taskinen 2011	niedrig	unklar	ja	nein	nein	niedrig
Scott 2008	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2005	niedrig	unklar	ja	nein	nein	niedrig
Bolli 2008	niedrig	ja	ja	nein	nein	niedrig
Bergenstal 2010b	niedrig	ja	ja	nein	nein	niedrig
Pratley 2010	niedrig	nein	ja	nein	nein	niedrig
D1690C00004	niedrig	ja	ja	nein	nein	niedrig
Matthews 2005	niedrig	ja	ja	nein	nein	niedrig
Matthews 2010	niedrig	ja	ja	nein	nein	niedrig
Nauck 2007	niedrig	unklar	ja	nein	nein	niedrig
Göke 2010	niedrig	ja	ja	nein	nein	niedrig
Filozof 2010	niedrig	ja	ja	nein	nein	niedrig
Yang 2011*	niedrig	ja	ja	nein	nein	niedrig
Pan 2012*	niedrig	unklar	ja	nein	nein	niedrig
DeFronzo 2012*	niedrig	unklar	ja	nein	nein	niedrig
Gallwitz 2012*	niedrig	ja	ja	nein	nein	niedrig
Yang 2012*	niedrig	unklar	ja	nein	nein	niedrig

*aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

Begründen Sie für jede Studie die abschließende Einschätzung.

Hypoglykämien wurden in den RCT unterschiedlich operationalisiert. In einigen Studien wurden nur bestätigte oder symptomatische Hypoglykämien erhoben. Auch hängt die bestätigte Hypoglykämie von unterschiedlichen Blutzuckerwerten (zwischen 2,8 und 3,5 mmol/L) ab. Der Endpunkt „Hypoglykämien“ schloss alle Hypoglykämien ein (inklusive schwerwiegende Hypoglykämien). Trotz der unterschiedlichen Operationalisierungen wurde

angenommen, dass diese Unterschiede nicht die Effektschätzer zwischen den Komparatoren beeinflussen.

In den Studien von MB102014, D1690C00012, Charbonnel 2006, Scott 2008, Bergenstal 2010b, D1690C00004, Matthews 2005, Matthews 2010, Göke 2010, Filozof 2010 und Yang 2011 (Yang 2011 wurde später durch die aktualisierte Literaturrecherche identifiziert) war das Verzerrungspotenzial des Endpunktes als niedrig anzusehen. Der Endpunkterheber war verblindet, das ITT-Prinzip wurde adäquat umgesetzt, es wurden weder Hinweise auf ergebnisgesteuerte Berichtserstattung noch sonstige Verzerrungspotenzial beeinflussende Punkte gefunden.

In den Studien von DeFronzo 2009, Raz 2008, Taskinen 2011, DeFronzo 2005, Nauck 2007 Pan 2012 und Yang 2012 (Pan 2012 und Yang 2012 wurden später durch die aktualisierte Literaturrecherche identifiziert) lagen keine Information über die Verblindung des Endpunkterhebers vor, es wurde aber von einem niedrigen Verzerrungspotenzial ausgegangen. Bei Pratley 2010 war der Endpunkterheber nicht verblindet, es gab aber keine weiteren Hinweise auf eine Verzerrung, so dass dieser Endpunkt als niedrig verzerrt eingestuft werden konnte.

Bosi 2007 war bereits auf Studienebene hoch verzerrt, die Verblindung des Endpunkterhebers war gewährleistet, jedoch wurde das ITT-Prinzip unzureichend dargestellt, so dass dieser Endpunkt als hoch verzerrt anzusehen war.

Bei Gallwitz 2012 war die Anzahl der Patienten mit mindestens einer Hypoglykämie ein Keysekundärer Endpunkt und wurde nach 104 Wochen berichtet. Der Endpunkt wurde als niedrig verzerrt betrachtet.

Bei DeFronzo 2012 wurden Hypoglykämien nur für die gepoolte Pioglitazon-Gruppe und die Kombination von Alogliptin+Pioglitazon berichtet. Dieser Vergleich ist für die Nutzenbewertung allerdings nicht relevant und wurde nicht dargestellt.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-98: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche (nach 24 Wochen im Vergleich zu Metformin+Placebo)

Studien	N	n (%)	N	n (%)	Odds Ratio [95%-KI]
		Metformin+DPP-4-Inhibitoren		Metformin+Placebo	
Charbonnel 2006	464	6 (2,2)	237	5 (2,1)	0,61 [0,18;2,01]
DeFronzo 2009	191	10 (5,2)	179	9 (5,0)	1,04 [0,41;2,63]
Raz 2008	96	1 (1,0)	94	0 (0,0)	2,97 [0,12;73,80]
Bosi 2007	183	1 (0,5)	181	1 (0,6)	0,99 [0,06;15,93]
Taskinen 20011	523	3 (0,6)	177	5 (2,8)	0,20 [0,05;0,84]
Scott 2008	94	1 (1,1)	92	2 (2,2)	0,48 [0,04;52,37]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					0,67 [0,37;1,23] I ² =0,0%; p=0,463
		Metformin+TZD		Metformin+Placebo	
Kaku 2009	83	1 (1,2)	86	0 (0,0)	3,15 [0,13;78,32]
		Metformin+Dapagliflozin		Metformin+Placebo	
MB102014	135	5 (3,7)	137	4 (2,9)	1,28 [0,34;4,87]
D1690C00012	91	2 (2,2)	91	3 (3,3)	0,66 [0,11;4,04]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					1,01 [0,35;2,97] I ² =0,0%; p=0,564
		Metformin+GLP-1-Analoga		Metformin+Placebo	
DeFronzo 2005*	223	11 (4,9)	113	6 (5,3)	0,93 [0,33;2,57]
		Metformin+TZD		Metformin+DPP-4-Inhibitoren	
Bolli 2008	280	0 (0,0)	295	1 (0,3)	0,35 [0,01;8,63]
Bergental 2010	165	1 (0,6)	166	5 (3,0)	0,20 [0,02;1,70]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					0,24 [0,04;1,41] I ² =0,0%; p=0,769
		Metformin+GLP-1-Analoga		Metformin+DPP-4-Inhibitoren	
Pratley 2010*	446	24 (5,4)	219	10 (4,6)	1,19 [0,56;2,53]

*2 verschiedene Dosierungen von GLP-1-Analoga (werden addiert)

[°]anhand Modell mit zufälligen Effekten

KI: Konfidenzintervall, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Gegenüber Metformin+Placebo hatte keine Behandlungsoption ein signifikant erhöhtes Hypoglykämienrisiko. Nauck 2009 berichtete den Endpunkt in der Publikation, wurde aber nicht in die Analyse miteingeschlossen, da dieser nicht die genaue Anzahl der Hypoglykämien berichtete.

Tabelle 4-99: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoff)

Studien	N	n (%)	N	n (%)	Odds Ratio [95%-KI]
		Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff	
Matthews 2010	1389	23 (1,7)	1383	224 (16,2)	0,09 [0,06;0,13]
Nauck 2007	588	29 (4,9)	584	187 (32,0)	0,11 [0,07;0,17]
Göke 2010	428	48 (11,2)	430	156 (36,3)	0,06 [0,03;0,10]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					0,08 [0,06;0,12] I ² =44,8%; p=0,163
		Metformin+TZD		Metformin+Sulfonylharnstoff	
Matthews 2005	317	4 (1,3)	313	35 (11,2)	0,10 [0,04;0,29]
		Metformin +Dapagliflozin		Metformin+Sulfonylharnstoff	
D1690C00004	406	14 (3,4)	408	162 (39,7)	0,05 [0,03;0,10]

[°] jeweils der Gesamtschätzer mittels Modell mit zufälligen Effekten mit Heterogenitätstest

KI: Konfidenzintervall, TZD: Thiazolidinedione, DPP-4: Dipeptidyl-Peptidase-4

Gegenüber Metformin+Sulfonylharnstoff zeigt sich für alle aktiven Komparatoren ein signifikant erhöhtes Hypoglykämienrisiko unter Metformin+Sulfonylharnstoff.

In der aktualisierten Literaturrecherche wurden zusätzlich weitere Publikationen gefunden, die den Endpunkt „Hypoglykämien“ beschreiben. Diese Publikationen beschreiben den Vergleich Metformin+DPP-4-Inhibitor versus Metformin+Placebo bzw. versus Metformin+Sulfonylharnstoff nach 24 Wochen / nach 104 Wochen und liefern ähnliche Ergebnisse:

Tabelle 4-100: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche – aktualisierte Literaturrecherche

Studien	N	n (%)	N	n (%)	Odds Ratio [95%-KI]
Nach 24 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Placebo		
Yang 2011	283	4 (1,4)	287	4 (1,4)	1,01 [0,25;4,10]
Pan 2012	146	1 (0,7)	144	0 (0,0)	2,98 [0,12;73,74]
Yang 2012	197	1 (0,5)	198	3 (1,5)	0,33 [0,03;3,22]
Nach 104 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff		
Gallwitz 2012	776	58 (7)	775	280 (36)	0,14 [0,11;0,19]

KI: Konfidenzintervall, DPP-4: Dipeptidyl-Peptidase-4

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.

Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.

Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.

Tabelle 4-101: Ergebnisse für Hypoglykämien aus indirekten Vergleiche nach 24 und 52 Wochen - Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel

Zu bewertende Arzneimittel	Komparator	Odds Ratio (95%-Kredibilitätsintervall*)
Nach 24 Wochen		
Metformin+Dapagliflozin	Metformin+DPP-4-Inhibitoren	1,75 [0,31;5,73]
Nach 52 Wochen		
Metformin+Dapagliflozin	Metformin+DPP-4-Inhibitoren	0,81 [0,18;2,59]

*basierend auf dem best angepassten Modell; DPP-4: Dipeptidyl-Peptidase-4

Die MTC zeigte sowohl nach 24 Wochen als auch nach 52 Wochen für Metformin+Dapagliflozin gegenüber DPP-4-Inhibitoren keine signifikante Erhöhung des Hypoglykämierisikos.

Um die Annahme der Konsistenz zu überprüfen, wurden die Resultate der *Head-to-Head*-Vergleiche mit den Resultaten aus den MTC verglichen. Es wurden keine Inkonsistenzen festgestellt.

Würde die Netzwerk-Meta-Analyse nach 24 Wochen mit den Publikationen von Yang 2011, Pan 2012 und Yang 2012 aktualisiert werden, würden keine Unterschiede in der Gesamtinterpretation erwartet werden. Auch würde sich nichts an der Gesamtinterpretation der MTC nach 52 Wochen ändern, falls die Publikation von Gallwitz 2012 noch hinzugefügt werden würde.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.1.2 Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-102: Operationalisierung von gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

Studie	Operationalisierung
	Siehe 4.3.2.1.3.1 und 4.3.2.1.3.1.1

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-103: Bewertung des Verzerrungspotenzials für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Siehe 4.3.2.1.3.1 und 4.3.2.1.3.1.1						

Begründen Sie für jede Studie die abschließende Einschätzung.

Da für alle Studien das Verzerrungspotenzial für die einzelnen Endpunkte „Ausmaß der Blutzuckerkontrolle“ und „Hypoglykämien“ als niedrig bewertet wurde, kann auch für die gemeinsame Betrachtung der Endpunkte das Verzerrungspotenzial niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Im Folgenden werden zunächst die in den vorherigen Kapiteln 4.3.2.1.3.1 und 4.3.2.1.3.1.1 dargestellten Endpunkte „Ausmaß der Blutzuckerkontrolle“ und „Hypoglykämien“ gemeinsam betrachtet, da das Auftreten von Hypoglykämien nur unter Berücksichtigung der Blutzuckerkontrolle sinnvoll beurteilt werden kann im. Es wird nur der Vergleich Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren betrachtet.

Tabelle 4-104: Ergebnisse für „gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien“ aus MTC mit dem zu bewertenden Arzneimittel

Zeitpunkt	Alle Studien in der MTC	Blutzuckerkontrolle	Hypoglykämien	Beurteilung
24 Wochen	Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren	kein Unterschied	kein Unterschied	kein Zusatznutzen
52 Wochen	Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren	kein Unterschied	kein Unterschied	kein Zusatznutzen

Kein Unterschied: kein Vor- bzw. Nachteil für eine der beiden Behandlungsoptionen.

Hinsichtlich der gemeinsamen Betrachtung der Blutzuckerkontrolle und der Hypoglykämien ergaben sich keine Unterschiede zwischen den Behandlungsoptionen.

4.3.2.1.3.1.3 Gewichtsveränderung – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-105: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin+ Placebo	Metformin+ TZD	Metformin+ DPP-4-Inhibitoren	Metformin+ GLP-1-Analoga
2	MB102014 D1690C0012	•	•			
2	Kaku 2009 DeFronzo 2012°		•	•		
9	DeFronzo 2009 Raz 2008 Bosi 2007 Taskinen 2011 Scott 2008 Bergenstal 2010a Yang 2011° Pan 2012° Yang 2012°		•		•	
2	DeFronzo 2005* Nauck 2009*		•			•
2	Bolli 2008 Bergenstal 2010b			•	•	
1	Pratley 2010*				•	•

*GLP-1-Analoga jeweils mit 2 verschiedenen Dosierungen

°aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

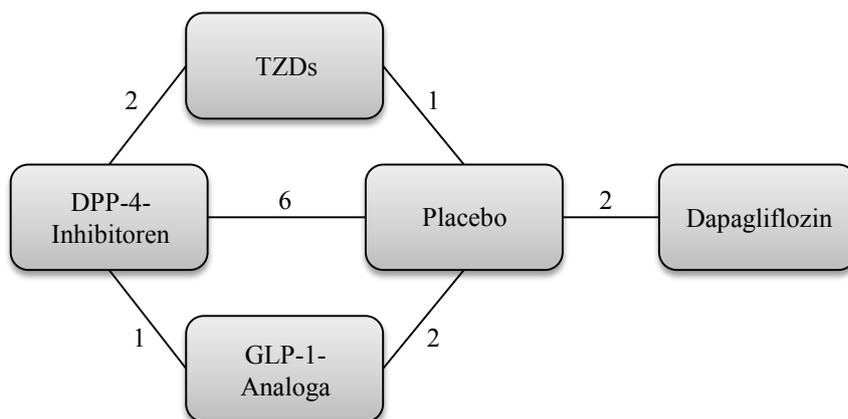
Tabelle 4-106: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin+ Sulfonylharnstoff	Metformin+ DPP-4-Inhibitoren	Metformin+ TZD	Metformin+ GLP-1-Analoga
1	D1690C00004	•	•			
5	Matthews 2010 Nauck 2007 Göke 2010 Filozof 2010 Gallwitz 2012°		•	•		
1	Derosa 2010**		•			•
1	Derosa 2011a ** °		•		•	
	Derosa 2011b ** ^ °					

** werden nicht berücksichtigt, da ein intensives Diätprogramm Bestandteil der Studien waren → Sensitivitätsanalyse
 °aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen) ^Publikation zum Abstract von Salvadeo 2010
 TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

A



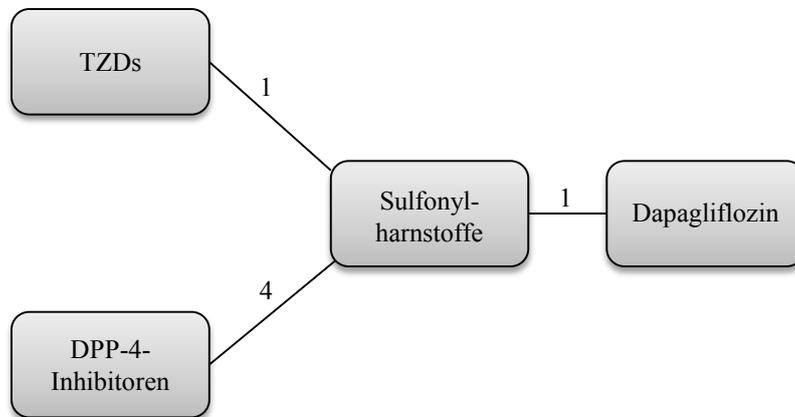
B

Abbildung 14: Netzwerk der Studien für Endpunkt Gewicht. A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk

* die Zahlen repräsentieren die Anzahl der RCT

Für den Endpunkt „Gewichtsveränderungen“ wurden 2 Netzwerke beschrieben, die aufgrund der Verfügbarkeit der Endpunkte und Studiendauer unterschiedlich für 24 und 52 Wochen waren. Im Netzwerk A sind alle RCT mit einer Studiendauer von 24±6 Wochen, im Netzwerk B alle RCT mit einer Studiendauer von 52±6 Wochen dargestellt. 14 RCT berichteten den Endpunkt „Gewichtsveränderung“ nach 24 Wochen und 6 RCT nach 52 Wochen.

Im Folgenden werden die Einzelergebnisse aller relevanten Studien tabellarisch beschrieben und gegebenenfalls, falls die Studien zusammenfassbar waren, ein Gesamtschätzer angegeben. Anschließend wird der indirekte Vergleich für Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren beschrieben (jeweils nach 24 und nach 52 Wochen).

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-107: Operationalisierung von Gewichtsveränderung

Studie	Operationalisierung
MB102014	Anhand der Veränderung des Gewichts nach 24 Wochen
D1690C00012	Anhand der Veränderung des Gewichts nach 24 Wochen
Kaku 2009	Anhand der Veränderung des Gewichts nach 28 Wochen
DeFronzo 2009	Anhand der Veränderung des Gewichts nach 24 Wochen
Raz 2008	Anhand der Veränderung des Gewichts nach 30 Wochen
Bosi 2007	Anhand der Veränderung des Gewichts nach 24 Wochen
Taskinen 2011	Anhand der Veränderung des Gewichts nach 24 Wochen
Scott 2008	Anhand der Veränderung des Gewichts nach 18 Wochen
Bergenstal 2010a	Anhand der Veränderung des Gewichts nach 24 Wochen
DeFronzo 2005	Anhand der Veränderung des Gewichts nach 30 Wochen
Bolli 2008	Anhand der Veränderung des Gewichts nach 24 Wochen
Bergenstal 2010b	Anhand der Veränderung des Gewichts nach 26 Wochen
Pratley 2010	Anhand der Veränderung des Gewichts nach 26 Wochen
Nauck 2009	Anhand der Veränderung des Gewichts nach 26 Wochen
D1690C00004	Anhand der Veränderung des Gewichts nach 52 Wochen
Matthews 2010	Anhand der Veränderung des Gewichts nach 52 Wochen
Nauck 2007	Anhand der Veränderung des Gewichts nach 52 Wochen
Göke 2010	Anhand der Veränderung des Gewichts nach 52 Wochen
Filozof 2010	Anhand der Veränderung des Gewichts nach 52 Wochen
Derosa 2010	Anhand der Veränderung des Gewichts nach 52 Wochen
Yang 2011 [°]	Anhand der Veränderung des Gewichts nach 24 Wochen
Derosa 2011a [°]	Anhand der Veränderung des Gewichts nach 52 Wochen
Derosa 2011b ^{° ^}	Anhand der Veränderung des Gewichts nach 52 Wochen
Pan 2012 [°]	Anhand der Veränderung des Gewichts nach 24 Wochen
DeFronzo 2012 [°]	Anhand der Veränderung des Gewichts nach 26 Wochen
Gallwitz 2012 [°]	Anhand der Veränderung des Gewichts nach 52 Wochen
Yang 2012 [°]	Anhand der Veränderung des Gewichts nach 26 Wochen

[°]aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

[^]Publikation zum Abstract von Salvadeo 2010, Salvadeo 2010 berichtet diesen Endpunkt nicht

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-108: Bewertung des Verzerrungspotenzials für „Gewichtsveränderung“ in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MB102014	niedrig	ja	ja	nein	nein	niedrig
D1690C00012	niedrig	ja	ja	nein	nein	niedrig
Raz 2008	niedrig	unklar	ja	nein	nein	niedrig
Bosi 2007	hoch	ja	unklar	nein	nein	hoch
Taskinen 2011	niedrig	unklar	ja	nein	ja	hoch
Scott 2008	niedrig	ja	ja	nein	nein	niedrig
Bergenstal 2010a	hoch	ja	unklar	unklar	ja	hoch
DeFronzo 2005	niedrig	unklar	ja	nein	nein	niedrig
Bolli 2008	niedrig	ja	nein	nein	nein	niedrig
Bergenstal 2010b	niedrig	ja	ja	nein	nein	niedrig
Pratley 2010	niedrig	nein	ja	nein	nein	niedrig
Nauck 2009	niedrig	unklar	ja	nein	nein	niedrig
D1690C00004	niedrig	ja	ja	nein	nein	niedrig
Matthews 2010	niedrig	ja	nein	nein	nein	niedrig
Nauck 2007	niedrig	unklar	ja	nein	nein	niedrig
Göke 2010	niedrig	ja	ja	nein	nein	niedrig
Filozof 2010	niedrig	ja	nein	nein	nein	niedrig
Derosa 2010	niedrig	nein	unklar	nein	nein	niedrig
Yang 2011°	niedrig	ja	ja	nein	nein	niedrig
Derosa 2011a°	niedrig	ja	ja	nein	nein	niedrig
Derosa 2011b °	niedrig	ja	ja	nein	nein	niedrig
Pan 2012°	niedrig	unklar	unklar	nein	nein	niedrig
DeFronzo 2012°	niedrig	unklar	ja	nein	nein	niedrig
Gallwitz 2012°	niedrig	ja	ja	nein	nein	niedrig
Yang 2012°	niedrig	unklar	ja	nein	nein	niedrig

°aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

Begründen Sie für jede Studie die abschließende Einschätzung.

In den Studien MB102014, D1690C00012, Charbonnel 2006, Scott 2008, Bergenstal 2010b, D1690C00004, Taskinen 2011, Göke 2010, Yang 2011 und Gallwitz 2012 (Yang 2011 und Gallwitz 2012 wurden später durch die aktualisierte Literaturrecherche identifiziert) ist das Verzerrungspotenzial des Endpunktes als niedrig anzusehen. Der Endpunkterheber war verblindet und das ITT-Prinzip wurde adäquat umgesetzt. Es wurden weder Hinweise auf ergebnisgesteuerte Berichtserstattung noch sonstige Verzerrungspotenzial beeinflussende Punkte gefunden.

In den Studien von Raz 2008, DeFronzo 2009, Nauck 2005, Nauck 2007, DeFronzo 2012 und Yang 2012 (DeFronzo 2012 und Yang 2012 wurden später durch die aktualisierte Literaturrecherche identifiziert) lagen keine Informationen über die Verblindung der Endpunkterheber vor, es wurde aber trotzdem von einem niedrigen Verzerrungspotenzial ausgegangen. Bei Pratley 2010 war der Endpunkterheber nicht verblindet, es gab aber keine weiteren Hinweise auf eine Verzerrung, so dass dieser Endpunkt ebenfalls als niedrig verzerrt eingestuft werden konnte.

Die Publikation von Bosi 2007 war bereits auf Studienebene hoch verzerrt, die Verblindung des Endpunkterhebers war gewährleistet, jedoch wurde das ITT-Prinzip unzureichend dargestellt, so dass dieser Endpunkt als hoch verzerrt anzusehen war.

In der Studie von Bergenstal 2010a wurde nicht genügend Informationen zur Verfügung gestellt, um eine Verzerrung ausschließen zu können, daher ist der Endpunkt als hoch verzerrt zu bewerten.

In der Studie von Bolli 2008 wurde keine adäquate Umsetzung des ITT-Prinzips durchgeführt, die Analyse des Endpunktes wurde mit der per Protokoll-Analyse durchgeführt, da sie aber ausführlich beschrieben wurde und es sonst keine Hinweise auf eine Verzerrung gab, wurde dieser Endpunkt als niedrig verzerrt eingestuft.

Bei der Publikation von Taskinen 2011 lag keine Information über die Verblindung des Endpunkterhebers vor, jedoch wurde das ITT-Prinzip adäquat umgesetzt. Weiterhin war unklar, ob die LOCF-Methode bei Patienten, die ein Notfallmedikament erhalten hatten, angewendet worden ist. Daher wurde dieser Endpunkt als hoch verzerrt angesehen.

In der Studie von Matthews 2010 wurde die per Protokoll-Population zur Berechnung dieses Endpunktes genutzt. Die ITT-Population wurde für die Sensitivitätsanalyse verwendet. Daher wurde das ITT-Prinzip nicht adäquat umgesetzt. Auch in der Publikation von Filozof 2010 beruhte die Analyse des Endpunktes auf der per Protokoll-Population. Die ITT-Population wurde für die Sensitivitätsanalyse (Ergebnisse wurden als vergleichbar angegeben, aber nicht berichtet) verwendet. Da es aber keine weiteren Hinweise auf eine Verzerrung gab, wurde der Endpunkt jeweils als niedrig verzerrt eingestuft.

Bei der Studie von Derosa 2010 handelte sich um eine einfachblinde Studie, wobei der Endpunkterheber nicht verblindet war. Die Patienten mussten ein intensives Diät- und Sportprogramm durchführen. Daher ist der Vergleich mit den anderen Studien mit Vorsicht

zu betrachten. Da es aber keine weiteren Hinweise auf eine Verzerrung gab, wurde dieser Endpunkt als niedrig verzerrt eingestuft. Derosa 2011a und Derosa 2011b sind doppelblinde Studien. Auch hier mussten die Patienten ein intensives Diät- und Sportprogramm durchführen.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-109: Ergebnisse für Gewichtsveränderung aus RCT für indirekte Vergleiche (nach 24 Wochen)

Studien	N	Veränderung des Gewichts Mittelwert (SE) kg	N	Veränderung des Gewichts Mittelwert (SE) kg	Gruppenunterschied kg Mittelwert [95%-KI]
		Metformin+DPP-4-Inhibitoren	Metformin+Placebo		
DeFronzo 2009	186	-0,92 (0,22)	175	-0,87 (0,23)	0,05 [-0,57;0,67]
Raz 2008	96	-0,5 (0,2784)	94	-0,5 (0,2959)	0,00 [-0,80;0,80]
Bosi 2007	143	0,2 (0,3)	130	-1,0 (0,3)	1,20 [0,37;2,03]
Taskinen 2011	513	-0,4 (0,2784)	175	-0,5 (0,2959)	0,10 [-0,70;0,90]
Scott 2008	94	-0,4 (0,2)	91	-0,8 (0,2)	0,40 [-0,15;0,95]
Bergental 2010a	177	-0,091 (0,204)	90	-0,488 (0,392)	0,40 [-0,47;1,26]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					0,33 [0,01;0,66] I ² =17,3%; p=0,302
		Metformin+TZD	Metformin+Placebo		
Kaku 2009	83	1,68 (0,2401)	86	-0,47 (0,2959)	2,15 [1,40;2,90]
		Metformin+Dapagliflozin	Metformin+Placebo		
MB102014	133	-2,9 (0,3)	136	-0,9 (0,3)	-2,00 [-2,83;-1,17]
D1690C00012	89	-2,96 (0,2766)	91	-0,88 (0,2746)	-2,08 [-2,84;-1,32]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					-2,04 [-2,61;-1,48] I ² =0%; p=0,890
		Metformin+GLP-1-Analoga	Metformin+Placebo		
DeFronzo 2005*	110	-2,21 (0,46)	113	-0,3 (0,3)	-1,91 [-2,98;-0,84]
Nauck 2009*	240	-2,7 (0,2)	121	-1,5 (0,3)	-1,20 [-1,91;-0,49]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					-1,43 [-2,09;-0,78] I ² =14,3%, p=0,280
	Metformin+TZD		Metformin+DPP-4-Inhibitoren		
Bolli 2008	246	1,9 (0,2)	264	0,3 (0,2)	1,60 [1,05;2,15]
Bergental 2010b	165	2,8 (0,3)	166	-0,8 (0,4)	3,60 [2,62;4,58]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					2,56 [0,60;4,52] I ² =91,8%; p=0,0001
	Metformin+DPP-4-Inhibitoren		Metformin+GLP-1-Analoga		
Pratley 2010*	221	-3,12 (0,28)	219	-0,96 (0,28)	-2,16 [-2,93;-1,38]

*2 verschiedene Dosierungen jeweils von GLP-1-Analoga (die Mittelwerte der beiden Dosierungsgruppen wurden gebildet)

[°]anhand Modell mit zufälligen Effekten; KI: Konfidenzintervall, SE: Standardfehler, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Tabelle 4-110: Ergebnisse für Gewichtsveränderung aus RCT für indirekte Vergleiche (nach 52 Wochen)

Studien	N	Veränderung des Gewichts Mittelwert (SE) kg	N	Veränderung des Gewichts Mittelwert (SE) kg	Gruppenunterschied kg Mittelwert [95%-KI]
		Metformin+DPP-4- Inhibitoren		Metformin+Sulfonylharnstoff	
Matthews 2010	1118	-0,23 (0,11)	1072	1,56 (0,12)	-1,79 [-2,11;-1,47]
Nauck 2007	588	-1,5 (0,3)	584	1,1 (0,3)	-2,60 [-3,43;-1,77]
Göke 2010	424	-1,1 (0,17)	426	1,1 (0,17)	-2,20 [-2,67;-1,73]
Filozof 2010	386	0,08 (0,1605)	393	1,36 (0,1634)	-1,28 [-1,73;-0,83]
Gesamtschätzer (mit Heterogenitätsmaß)^o					-1,90 [-2,36;-1,43] I ² =73,8% p=0,009
		Metformin+Dapagliflozin		Metformin+Sulfonylharnstoff	
D1690C00004	400	-3,22 (0,18)	401	1,44 (0,18)	-4,66 [-5,16;-4,16]

^o anhand Modell mit zufälligen Effekten

KI: Konfidenzintervall, SE: Standardfehler, DPP-4: Dipeptidyl-Peptidase-4

Signifikante Ergebnisse zeigen sich im Vergleich mit Metformin+Dapagliflozin versus Metformin+Placebo nach 24 Wochen und bei Metformin+Dapagliflozin versus Metformin+Sulfonylharnstoff nach 52 Wochen jeweils zugunsten von Metformin+Dapagliflozin. Der Vergleich Metformin+DPP-4-Inhibitoren versus Metformin+Sulfonylharnstoff zeigt nach 52 Wochen ein signifikantes Ergebnis zugunsten Metformin+DPP-4-Inhibitoren. In der Studie von Matthews 2005 wurde der Endpunkt berichtet, jedoch nicht in die Analyse eingeschlossen, da die Standardabweichung / Standardfehler der Veränderung des Gewichts nicht berichtet wurde.

In der aktualisierten Literaturrecherche wurden zusätzlich weitere Publikationen (Yang 2011, Pan 2012, Derosa 2011a und Derosa 2011b) gefunden, die den Endpunkt „Gewichtsveränderung“ beschreiben. Yang 2011 beschreibt den Vergleich Metformin+DPP-4-Inhibitor versus Metformin+Placebo nach 24 Wochen und liefert ähnliche Ergebnisse; Derosa 2011a beschreibt jeweils nach 52 Wochen den Vergleich Metformin+TZD versus Metformin+Sulfonylharnstoff und Derosa 2011b beschreibt den Vergleich Metformin+GLP-1-Analoga versus Metformin+Sulfonylharnstoff.

Tabelle 4-111: Ergebnisse für Gewichtsveränderung aus RCT für paarweise indirekte Vergleiche –aktualisierte Literaturrecherche

Studien	N	Veränderung des Gewichts kg Mittelwert (SE) / [95%-KI]	N	Veränderung des Gewichts kg Mittelwert (SE) / [95%-KI]	Gruppenunterschied Mittelwert kg [95%-KI]
Nach 24 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Placebo		
Yang 2011	278	-1,05 (0,1)	281	-0,97 (0,1)	-0,08 [-0,36;0,20]
Pan 2012	145	-0,7 (n. b.)	144	-1,0 (n. b.)	0,3 [n. b.]
Yang 2012	197	0,0 [-0,3;0,4]	198	-0,5 [-0,9;-0,1]	0,5 [0,1;0,9]
	Metformin+TZD		Metformin+Placebo		
DeFonzo 2012	388	1,5 (0,2)	129	-0,7 (n.b.)	Nicht berechenbar
Nach 52 Wochen					
	Metformin+TZD		Metformin+Sulfonylharnstoff		
Derosa 2011a **	102	0,8 (1,3)	99	1,2 (1,3)	-0,40 [-4,00;3,20]
	Metformin+GLP-1-Analoga		Metformin+Sulfonylharnstoff		
Derosa 2011b ** °	57	-5,1 (1,3)	54	-0,9 (1,5)	4,20 [-4,00;3,20]
Nach 104 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff		
Gallwitz 2012	733	-1,4 (0,2)	722	1,3 (0,2)	-2,5[-2,8;-2,2]

** werden nicht in die primäre Analyse nicht miteinbezogen → Sensitivitätsanalyse

KI: Konfidenzintervall, SE: Standardfehler, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4, n. b.: nicht berechenbar

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.

Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.

Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.

Tabelle 4-112: Ergebnisse für Gewichtsveränderung aus indirekten Vergleichen nach 24 und 52 Wochen - Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel

Zu bewertende Arzneimittel	Komparator	Mittelwertdifferenz kg (95%-Kredibilitätsintervall*)
Nach 24 Wochen		
Metformin+Dapagliflozin	Metformin+DPP-4-Inhibitoren	-2,37 [-2,98;-1,77]
Nach 52 Wochen		
Metformin+Dapagliflozin	Metformin+ DPP-4-Inhibitoren	-2,74 [-5,35;-0,10]

*basierend auf dem best angepassten Modell; DPP-4: Dipeptidyl-Peptidase-4

Die MTC zeigte sowohl nach 24 als auch nach 52 Wochen eine deutliche und statistisch signifikante Gewichtsreduktion für Metformin+Dapagliflozin im Vergleich zu Metformin+DPP-4-Inhibitoren.

Um die Annahme der Konsistenz zu überprüfen, wurden die Resultate der *Head-to-Head*-Vergleiche mit den Resultaten aus der MTC verglichen. Es wurden keine Inkonsistenzen festgestellt.

Nimmt man die Studie von Derosa 2010 in die Analyse der MTC mit auf, ergeben sich keine relevanten Unterschiede in den Effektschätzern (Sensitivitätsanalyse). Derosa 2011a und Derosa 2011b (dieser Endpunkt war im Abstract von Salvadeo 2010 nicht verfügbar) würden ebenfalls in die Sensitivitätsanalyse (aufgrund des intensiven Sport- und Diätprogramms) einfließen. Dadurch, dass es keine anderen Studien mit den direkten Vergleichen (Metformin+TZD versus Metformin+Sulfonylharnstoff und Metformin+GLP-1-Analoga versus Metformin+Sulfonylharnstoff) gab, konnte der Einfluss auf die MTC nicht bestimmt werden.

Würde die Netzwerk-Meta-Analyse nach 24 Wochen mit der Publikation von Yang 2011 und Yang 2012 aktualisiert werden (Pan 2012 berichtet keine Streuungsmaße, so dass keine Effektschätzer berichtet werden können), würden keine wesentlichen Veränderungen an der Gesamtinterpretation der MTC entstehen. Gallwitz 2012 berichtet die Veränderung des Gewichts nur nach 104 Wochen und könnte daher gar nicht in der MTC berücksichtigt werden.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.1.4 Blutdrucksenkung – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-113: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)

Anzahl Studien	Referenzen Studien	Metformin + Dapa-gliflozin	Metformin + Placebo	Metformin + TZD	Metformin+ DPP-4-Inhibitoren	Metformin + GLP-1-Analoga	Metformin+ Sulfonylharnstoff
2	MB102014 D1690C000 12	•	•				
1	Papathanassi ou 2009			•			•
1	Charpentier 2001		•				•
1	Bergental 2010b			•	•		
1	Pratley 2010*				•	•	
1	Nauck 2009*		•			•	•
2	DeFronzo 2009 Yang 2011 ^o		•		•		

*GLP-1-Analoga jeweils mit 2 verschiedenen Dosierungen

^oaktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

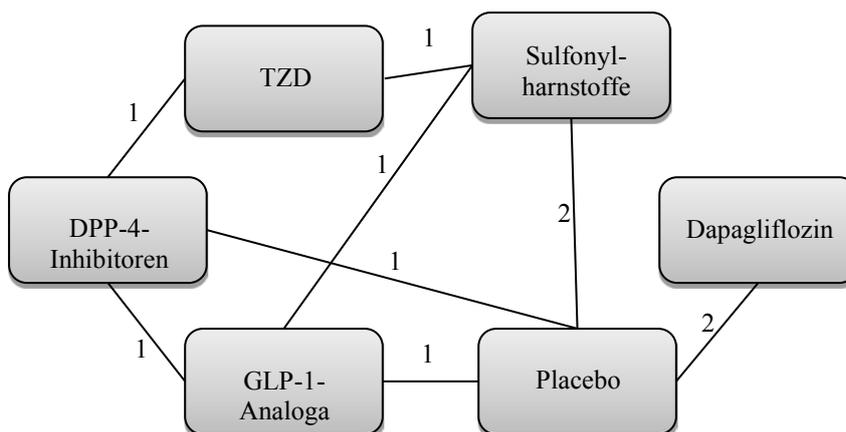
Tabelle 4-114: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin+ Sulfonylharnstoff	Metformin+ DPP-4-Inhibitoren	Metformin+ TZD	Metformin+ GLP-1-Analoga
1	D1690C0004	•	•			
1	Göke 2010		•	•		

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

A



B



Abbildung 15: Netzwerk der Studien für Endpunkt Blutdrucksenkung. A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk

* die Zahlen repräsentieren die Anzahl der RCT

Es wurden 2 Netzwerke beschrieben, die aufgrund der Verfügbarkeit der Endpunkte und Studiendauer unterschiedlich für 24 und 52 Wochen waren. In den 52-Wochen Sulfonylharnstoff-Studien gab es Titrationsphasen von 18 Wochen, daher wurden sie im 24 Wochen Netzwerk, mit Ausnahme des Endpunkts „Blutdrucksenkung“ (um für diesen Endpunkt die Evidenz der MTC zu erhöhen), ausgeschlossen. Im Netzwerk A wurden alle RCT mit einer Studiendauer von 24±6 Wochen, im Netzwerk B alle RCT mit einer Studiendauer von 52±6 Wochen dargestellt. Es wurde hier nur die systolische Blutdrucksenkung beschrieben. 8 RCT berichteten diesen Endpunkt nach 24 Wochen inklusive der RCT, die Metformin+Sulfonylharnstoffe versus einen Komparator beschrieben. Nach 52 Wochen berichteten nur 2 Studien diesen Endpunkt. Sie wurden nicht meta-analysiert.

Im Folgenden werden die Einzelergebnisse aller relevanten Studien tabellarisch beschrieben und gegebenenfalls, falls die Studien zusammenfassbar waren, ein Gesamtschätzer angegeben. Anschließend wird der indirekte Vergleich für Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren beschrieben.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-115: Operationalisierung von Blutdrucksenkung

Studie	Operationalisierung
MB102014	Anhand der Veränderung des systolischen Blutdrucks nach 24 Wochen
D1690C00012	Anhand der Veränderung des systolischen Blutdrucks nach 24 Wochen
Papathanassiou 2009	Anhand der Veränderung des systolischen Blutdrucks nach 26 Wochen
Charpentier 2001	Anhand der Veränderung des systolischen Blutdrucks nach 20 Wochen
Bergental 2010b	Anhand der Veränderung des systolischen Blutdrucks nach 24 Wochen
Pratley 2010	Anhand der Veränderung des systolischen Blutdrucks nach 26 Wochen
Nauck 2009	Anhand der Veränderung des systolischen Blutdrucks nach 26 Wochen
DeFronzo 2009	Anhand der Veränderung des systolischen Blutdrucks nach 24 Wochen
D1690C00004	Anhand der Veränderung des systolischen Blutdrucks nach 52 Wochen
Göke 2010	Anhand der Veränderung des systolischen Blutdrucks nach 52 Wochen
Yang 2011°	Anhand der Veränderung des systolischen Blutdrucks nach 24 Wochen

°aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-116: Bewertung des Verzerrungspotenzials für Blutdrucksenkung in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MB102014	niedrig	ja	ja	nein	nein	niedrig
D1690C00012	niedrig	ja	ja	nein	nein	niedrig
Papathanassiou 2009	hoch	nein	ja	unklar	ja	hoch
Charpentier 2001	hoch	unklar	ja	unklar	nein	hoch
Bergental 2010b	niedrig	ja	ja	nein	nein	niedrig
Pratley 2010	niedrig	nein	ja	nein	nein	niedrig
Nauck 2009	niedrig	unklar	ja	nein	nein	niedrig
D1690C00004	niedrig	ja	ja	nein	nein	niedrig
Matthews 2005	niedrig	ja	ja	nein	nein	niedrig
Göke 2010	niedrig	ja	ja	nein	nein	niedrig
Yang 2011 ^o	niedrig	ja	ja	nein	nein	niedrig

^oaktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

Begründen Sie für jede Studie die abschließende Einschätzung.

In den Studien MB102014, D1690C00012, Bergental 2010b, D1690C00004, Matthews 2005, Göke 2010 und und Yang 2011 (Yang 2011 wurde später durch die aktualisierte Literaturrecherche identifiziert) ist das Verzerrungspotenzial des Endpunktes als niedrig anzusehen. Der Endpunkterheber war verblindet und das ITT-Prinzip wurde adäquat umgesetzt. Es wurden weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gefunden.

In der Studie von Nauck 2009 lag keine Information über die Verblindung des Endpunkterhebers vor, es wurde aber von einem niedrigen Verzerrungspotenzial ausgegangen. Bei Pratley 2010 war der Endpunkterheber nicht verblindet, es gab aber keine weiteren Hinweise auf eine Verzerrung, so dass dieser Endpunkt ebenfalls als niedrig verzerrt eingestuft werden konnte.

In den Publikationen von Papathanassiou 2009 und Charpentier 2001 handelte es sich bereits um hoch verzerrte Studien. Daher wurde die endpunktspezifische Verzerrung ebenfalls als hoch angesehen.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-117: Ergebnisse für systolische Blutdrucksenkung aus RCT für indirekte Vergleiche (nach 24 Wochen)

Studien	N	Veränderung des Blutdruckes Mittelwert (SD) mmHg	N	Veränderung des Blutdruckes Mittelwert (SD) mmHg	Gruppenunterschied Mittelwert mmHg [95%-KI]
		Metformin+DPP-4-Inhibitoren			Metformin+Placebo
DeFronzo 2009	141	-3,8 (1,4)	105	-3,7 (1,3)	-0,10 [-3,84;3,64]
		Metformin+Sulfonylharnstoff			Metformin+Placebo
Charpentier 2001	147	-0,14 (1,089)	75	-0,65 (1,411)	0,51 [-2,98;4,00]
Nauck 2009 ^{oo}	242	0,4 (0,891)	121	-1,8 (1,125)	2,20 [-1,52;5,92]
Gesamtschätzer (mit Heterogenitätsmaß)^o					1,30 [-1,24;3,85] I ² =0% ; p=0,516
		Metformin+Dapagliflozin			Metformin+Placebo
MB102014	122	-5,1 (1,3)	119	-0,2 (1,2)	-4,90 [-8,37;-1,43]
D1690C00012	88	-2,7 (1,088)	91	0,1 (1,071)	-2,80 [-5,79;0,19]
Gesamtschätzer (mit Heterogenitätsmaß)^o					-3,70 [-5,96;-1,43] I ² =0% ; p=0,369
		Metformin+GLP-1-Analoga			Metformin+Placebo
Nauck 2009* ^{oo}	240	-2,55 (0,891)	121	-1,8 (1,125)	-0,75 [-3,70;2,20]
		Metformin+TZD			Metformin+DPP-4-Inhibitoren
Bergental 2010	165	-1,58 (0,94)	166	0,22 (0,93)	-1,80 [-4,39;0,79]
		Metformin+GLP-1-Analoga			Metformin+DPP-4-Inhibitoren
Pratley 2010*	221	-0,634 (0,89)	219	-0,94 (0,89)	0,31 [-2,16;2,77]
		Metformin+GLP-1-Analoga			Metformin+Sulfonylharnstoff
Nauck 2009* ^{oo}	240	-2,55 (0,891)	242	0,4 (0,892)	-2,95 [-6,30;0,40]
		Metformin+TZD			Metformin+Sulfonylharnstoff
Papathanassiou 2009	14	-8,67 (5,067)	14	-11,08 (5,340)	2,41 [-12,02;16,84]

*2 verschiedene Dosierungen jeweils von GLP-1-Analoga (die Mittelwerte der beiden Dosierungsgruppen wurden gebildet)

^{oo}Nauck 2009 hat 3 Behandlungsoptionen

^oanhand Modell mit zufälligen Effekten

KI: Konfidenzintervall, SD: Standardabweichung, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Tabelle 4-118: Ergebnisse für systolische Blutdrucksenkung aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoff)

Studien	N	Veränderung des Blutdruckes Mittelwert (SD) mmHg	N	Veränderung des Blutdruckes Mittelwert (SD) mmHg	Gruppenunterschied Mittelwert mmHg [95%-KI]
		Metformin+Dapagliflozin	Metformin+Sulfonylharnstoff		
D1690 C00004	399	-4,3 (0,59)	396	0,8 (0,59)	-5,1 [-6,74;-3,46]
		Metformin+DPP-4-Inhibitoren	Metformin+Sulfonylharnstoff		
Göke 2010	293	-4,1 (0,82)	293	-1,2 (0,63)	-2,90 [-6,23;-1,94]

KI: Konfidenzintervall, SD: Standardabweichung, DPP-4: Dipeptidyl-Peptidase-4

In der aktualisierten Literaturrecherche wurde zusätzlich eine weitere Publikation (Yang 2011) gefunden, die den Endpunkt „Blutdrucksenkung“ beschreibt. Diese Publikation beschreibt den Vergleich Metformin+DPP-4-Inhibitor versus Metformin+Placebo nach 24 Wochen und liefert ähnliche Ergebnisse:

Tabelle 4-119: Ergebnisse für „Blutdrucksenkung“ aus RCT für indirekte Vergleiche – aktualisierte Literaturrecherche

Studien	N	Veränderung des Blutdrucks Mittelwert (SE) mmHg	N	Veränderung des Blutdrucks Mittelwert (SE) mmHg	Gruppenunterschied Mittelwert [95%-KI]
Nach 24 Wochen					
		Metformin+DPP-4-Inhibitoren	Metformin+Placebo		
Yang 2011	283	-5,22 (n.b)	287	3,10 (n. b)	-2,12 [nicht berechenbar]

n. b.: nicht berichtet, KI: Konfidenzintervall, SE: Standardfehler, DPP-4: Dipeptidyl-Peptidase-4

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.

Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.

Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.

Tabelle 4-120: Ergebnisse für systolische Blutdrucksenkung aus indirekten Vergleichen nach 24 Wochen- Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel

Zu bewertende Arzneimittel	Komparator	Mittelwertdifferenz mmHg (95%- Kredibilitätsintervall*)
Nach 24 Wochen		
Metformin+Dapagliflozin	Metformin+DPP-4-Inhibitoren	-3,13 [-7,19;0,80]

*basierend auf dem best angepassten Modell

DPP-4: Dipeptidyl-Peptidase-4

Im Vergleich zu Metformin+DPP-4-Inhibitoren zeigte sich kein signifikanter Unterschied für Metformin+Dapagliflozin.

Um die Annahme der Konsistenz zu überprüfen, wurden die Resultate der *Head-to-Head*-Vergleiche mit den Resultaten aus den MTC verglichen. Es wurden keine Inkonsistenzen festgestellt.

Es würde keinen Unterschied in der Interpretation des Ergebnisses machen, falls die Studie von Yang 2011 in die MTC eingeschlossen wäre.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.1.5 Unerwünschte Ereignisse – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-121: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 24 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin+ Placebo	Metformin+ TZD	Metformin+ DPP-4-Inhibitoren	Metformin+ GLP-1-Analoga
2	MB102014 D1690C00012	•	•			
1	Kaku 2009		•	•		
9	Charbonnel 2006 DeFronzo 2009 Raz 2008 Bosi 2007 Taskinen 2011 Scott 2008 Yang 2011° Pan 2012° Yang 2012		•		•	
2	DeFronzo 2005* Nauck 2009		•			•
1	Bolli 2008			•	•	
1	Bergental 2010b			•	•	
1	Pratley 2010*				•	•

*GLP-1-Analoga jeweils mit 2 verschiedenen Dosierungen

°aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Tabelle 4-122: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden (nach 52 Wochen)

Anzahl Studien	Referenzen Studien	Metformin+ Dapagliflozin	Metformin+ Sulfonylharnstoff	Metformin+ DPP-4-Inhibitoren	Metformin+ TZD	Metformin+ GLP-1-Analoga
1	D1690C00004	•	•			
2	Matthews 2005 Derosa 2011a ^o *		•		•	
4	Nauck 2007 Göke 2010 Filozof 2010 Gallwitz 2012		•	•		
1	Derosa 2010* Derosa 2011b ^o * ^		•			•

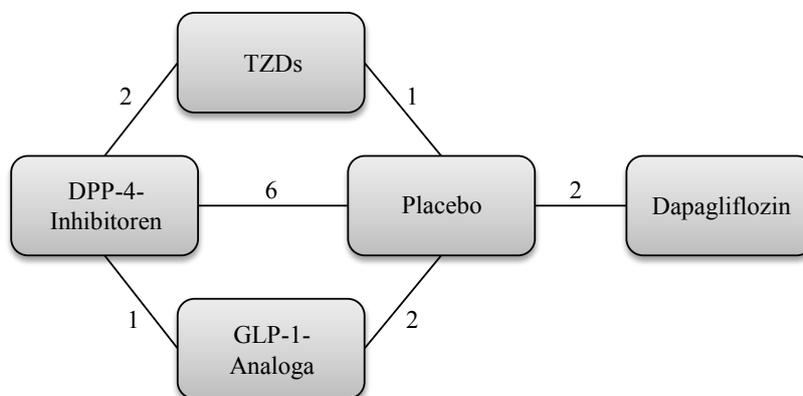
* wird in der primären Analyse nicht berücksichtigt, da ein intensives Diätprogramm Bestandteil der Studien waren → Sensitivitätsanalyse

^oaktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen) [^]Publikation von Salvadeo 2010

TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

A



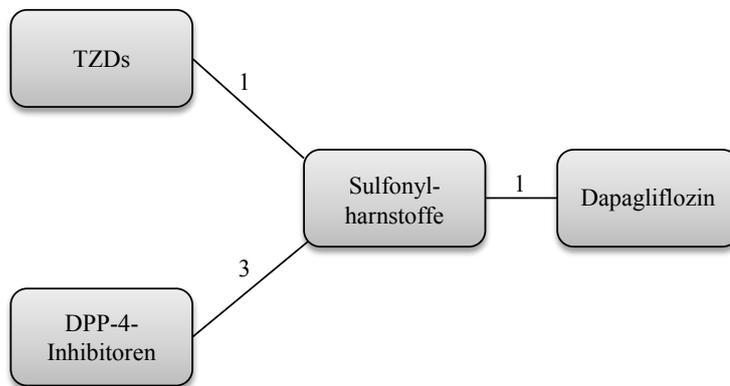
B

Abbildung 16: Netzwerk der Studien für Endpunkt Unerwünschte Ereignisse. A: 24 Wochen Netzwerk. B: 52 Wochen Netzwerk

Es wurden 2 Netzwerke beschrieben, die aufgrund der Verfügbarkeit der Endpunkte und Studiendauer unterschiedlich für 24 und 52 Wochen waren. Die Endpunkte „Gesamtrate unerwünschter Ereignisse“ sowie „schwerwiegende unerwünschte Ereignisse“ und „Studienabbrüche aufgrund unerwünschter Ereignisse“ wurden in unterschiedlich vielen RCT beschrieben. 11 RCT berichteten die Gesamtrate unerwünschter Ereignisse nach 24 Wochen, und 5 RCT nach 52 Wochen. Schwerwiegende unerwünschter Ereignisse wurden in 13 RCT nach 24 Wochen und in 4 RCT nach 52 Wochen beschrieben. Die Studienabbrüche aufgrund unerwünschter Ereignisse wurden in 13 und 5 RCT nach 24 und 52 Wochen berichtet.

Im Folgenden werden die Einzelergebnisse aller relevanten Studien tabellarisch beschrieben und gegebenenfalls, falls die Studien zusammenfassbar waren, ein Gesamtschätzer angegeben. Anschließend wird der indirekte Vergleich für Metformin+Dapagliflozin versus Metformin+DPP-4-Inhibitoren beschrieben (jeweils nach 24 und nach 52 Wochen).

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-123: Operationalisierung von unerwünschte Ereignisse

Studie	Operationalisierung
MB102014	Gesamtrate unerwünschter Ereignisse (exklusive Hypoglykämien, mit UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
D1690C00012	Gesamtrate unerwünschter Ereignisse (exklusive Hypoglykämien, mit UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Kaku 2009	Gesamtrate unerwünschter Ereignisse (inklusive Hypoglykämien, mit UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse (vom Autor nachgefordert) Studienabbrüche aufgrund unerwünschter Ereignisse (nicht als Sicherheitsendpunkt berichtet)
Charbonnel 2006	Gesamtrate unerwünschter Ereignisse (unklar, ob mit oder ohne Hypoglykämien; ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
DeFronzo 2009	Gesamtrate unerwünschter Ereignisse (exklusive Hypoglykämien, mit UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Raz 2008	Gesamtrate unerwünschter Ereignisse (inklusive Hypoglykämien, ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Bosi 2007	Gesamtrate unerwünschter Ereignisse (unklar, ob mit oder ohne Hypoglykämien; unklar, ob mit oder ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Taskinen 2011	Gesamtrate unerwünschter Ereignisse (inklusive Hypoglykämien, ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Scott 2008	Gesamtrate unerwünschter Ereignisse (inklusive Hypoglykämien, ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
DeFronzo 2005	Gesamtrate unerwünschter Ereignisse: nicht berichtet Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse (nicht als Sicherheitsendpunkt berichtet)

Studie	Operationalisierung
Bolli 2008	Gesamtrate unerwünschter Ereignisse (unklar, ob mit oder ohne Hypoglykämien; ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Bergental 2010b	Gesamtrate unerwünschter Ereignisse: nicht berichtet Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Nauck 2005	Gesamtrate unerwünschter Ereignisse: nicht berichtet Gesamtrate schwerwiegender unerwünschter Ereignisse: nicht berichtet Studienabbrüche aufgrund unerwünschter Ereignisse
Pratley 2010	Gesamtrate unerwünschter Ereignisse (exklusive Hypoglykämien, ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse (nicht als Sicherheitsendpunkt berichtet)
D1690C00004	Gesamtrate unerwünschter Ereignisse (exklusive Hypoglykämien, mit UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Matthews 2005	Gesamtrate unerwünschter Ereignisse (inklusive Hypoglykämien, ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Nauck 2007	Gesamtrate unerwünschter Ereignisse (inklusive Hypoglykämien, ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Göke 2010	Gesamtrate unerwünschter Ereignisse (exklusive Hypoglykämien, mit UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse Studienabbrüche aufgrund unerwünschter Ereignisse
Filozof 2010	Gesamtrate unerwünschter Ereignisse (unklar, ob mit oder ohne Hypoglykämien; ohne UEs, die sich auf Laborparameter beziehen) Gesamtrate schwerwiegender unerwünschter Ereignisse: nicht berichtet Studienabbrüche aufgrund unerwünschter Ereignisse
Derosa 2010	Gesamtrate unerwünschter Ereignisse: nicht berichtet Gesamtrate schwerwiegender unerwünschter Ereignisse: nicht berichtet Studienabbrüche aufgrund unerwünschter Ereignisse
Yang 2011*	Gesamtrate unerwünschter Ereignisse: nicht berichtet Gesamtrate schwerwiegender unerwünschter Ereignisse: nicht berichtet Studienabbrüche aufgrund unerwünschter Ereignisse

Studie	Operationalisierung
Derosa 2011a*	Gesamtrate unerwünschter Ereignisse: nicht berichtet Gesamtrate schwerwiegender unerwünschter Ereignisse: nicht berichtet Studienabbrüche aufgrund unerwünschter Ereignisse
Derosa 2011b° *	Gesamtrate unerwünschter Ereignisse: nicht berichtet Gesamtrate schwerwiegender unerwünschter Ereignisse: nicht berichtet Studienabbrüche aufgrund unerwünschter Ereignisse
Pan 2012*	Gesamtrate unerwünschter Ereignisse: nicht berichtet Gesamtrate schwerwiegender unerwünschter Ereignisse: nicht berichtet Studienabbrüche aufgrund unerwünschter Ereignisse
DeFronzo 2012*	Gesamtrate unerwünschter Ereignisse: Gesamtrate schwerwiegender unerwünschter Ereignisse: Studienabbrüche aufgrund unerwünschter Ereignisse
Gallwitz 2012*	Gesamtrate unerwünschter Ereignisse: Gesamtrate schwerwiegender unerwünschter Ereignisse: Studienabbrüche aufgrund unerwünschter Ereignisse
Yang 2012*	Gesamtrate unerwünschter Ereignisse: Gesamtrate schwerwiegender unerwünschter Ereignisse: Studienabbrüche aufgrund unerwünschter Ereignisse

UE: unerwünschtes Ereignis

*aktualisierte Literaturrecherche (wurden nicht in die MTC miteingeschlossen)

°Publikation zum Abstract von Salvadeo 2010

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-124: Bewertung des Verzerrungspotenzials für Gesamtrate unerwünschter Ereignisse in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MB102014	niedrig	ja	ja	nein	nein	niedrig
D1690C00012	niedrig	ja	ja	nein	nein	niedrig
Kaku 2009	hoch	unklar	ja	nein	nein	Hoch
Charbonnel 2006	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2009	niedrig	unklar	ja	nein	nein	niedrig
Raz 2008	niedrig	unklar	ja	nein	nein	niedrig
Bosi 2007	hoch	ja	ja	nein	nein	hoch
Taskinen 2011	niedrig	unklar	ja	nein	nein	niedrig
Scott 2008	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2005	niedrig	unklar	ja	nein	nein	niedrig
Bolli 2008	niedrig	ja	ja	nein	nein	niedrig
Bergental 2010b	niedrig	ja	ja	nein	nein	niedrig
Pratley 2010	niedrig	nein	Ja	nein	nein	hoch
Nauck 2009	niedrig	unklar	ja	nein	nein	niedrig
D1690C00004	niedrig	ja	ja	nein	nein	niedrig
Matthews 2005	niedrig	ja	ja	nein	nein	niedrig
Matthews 2010	niedrig	ja	ja	nein	nein	niedrig
Nauck 2007	niedrig	unklar	ja	nein	nein	niedrig
Göke 2010	niedrig	ja	ja	nein	nein	niedrig
Filozof 2010	niedrig	ja	ja	nein	nein	niedrig
Yang 2011°	niedrig	unklar	ja	nein	nein	niedrig
Pan 2012°	niedrig	unklar	ja	nein	nein	niedrig
DeFronzo 2012°	niedrig	unklar	ja	nein	nein	niedrig
Gallwitz 2012°	niedrig	ja	ja	nein	nein	niedrig
Yang 2012°	niedrig	unklar	ja	nein	nein	niedrig

° aktualisierte Literaturrecherche

Tabelle 4-125: Bewertung des Verzerrungspotenzials für Gesamtrate schwerwiegender unerwünschter Ereignisse in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MB102014	niedrig	ja	ja	nein	nein	niedrig
D1690C00012	niedrig	ja	ja	nein	nein	niedrig
Kaku 2009	hoch	unklar	ja	nein	nein	hoch
Charbonnel 2006	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2009	niedrig	unklar	ja	nein	nein	niedrig
Raz 2008	niedrig	unklar	ja	nein	nein	niedrig
Bosi 2007	hoch	ja	ja	nein	nein	hoch
Taskinen 2011	niedrig	unklar	ja	nein	nein	niedrig
Scott 2008	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2005	niedrig	unklar	ja	nein	nein	niedrig
Bolli 2008	niedrig	ja	ja	nein	nein	niedrig
Bergental 2010b	niedrig	ja	ja	nein	nein	niedrig
Pratley 2010	niedrig	nein	ja	nein	nein	niedrig
Nauck 2009	niedrig	unklar	ja	nein	nein	niedrig
D1690C00004	niedrig	ja	ja	nein	nein	niedrig
Matthews 2005	niedrig	ja	ja	nein	nein	niedrig
Matthews 2010	niedrig	ja	ja	nein	nein	niedrig
Nauck 2007	niedrig	unklar	ja	nein	nein	niedrig
Göke 2010	niedrig	ja	ja	nein	nein	niedrig
Yang 2011 ^o	niedrig	unklar	ja	nein	nein	niedrig
Pan 2012 ^o	niedrig	unklar	ja	nein	nein	niedrig
DeFronzo 2012 ^o	niedrig	unklar	ja	nein	nein	niedrig
Gallwitz 2012 ^o	niedrig	ja	ja	nein	nein	niedrig
Yang 2012 ^o	niedrig	unklar	ja	nein	nein	niedrig

^o aktualisierte Literaturrecherche

Tabelle 4-126: Bewertung des Verzerrungspotenzials für Studienabbrüche aufgrund unerwünschter Ereignisse in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
MB102014	niedrig	ja	ja	nein	nein	niedrig
D1690C00012	niedrig	ja	ja	nein	nein	niedrig
Kaku 2009	hoch	unklar	ja	nein	nein	hoch
Charbonnel 2006	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2009	niedrig	unklar	ja	nein	nein	niedrig
Raz 2008	niedrig	unklar	ja	nein	nein	niedrig
Bosi 2007	hoch	ja	ja	nein	nein	hoch
Taskinen 2011	niedrig	unklar	ja	nein	nein	niedrig
Scott 2008	niedrig	ja	ja	nein	nein	niedrig
DeFronzo 2005	niedrig	unklar	ja	nein	nein	niedrig
Bolli 2008	niedrig	ja	ja	nein	nein	niedrig
Bergental 2010b	niedrig	ja	ja	nein	nein	niedrig
Pratley 2010	niedrig	nein	ja	ja	ja	hoch
Nauck 2009	niedrig	unklar	ja	nein	nein	niedrig
D1690C00004	niedrig	ja	ja	nein	nein	niedrig
Matthews 2005	niedrig	ja	ja	nein	nein	niedrig
Matthews 2010	niedrig	ja	ja	nein	nein	niedrig
Nauck 2007	niedrig	unklar	ja	nein	nein	niedrig
Göke 2010	niedrig	ja	ja	nein	nein	niedrig
Filozof 2010	niedrig	ja	ja	nein	nein	niedrig
Yang 2011°	niedrig	unklar	ja	nein	nein	niedrig
Derosa 2011a° *	niedrig	ja	ja	nein	nein	niedrig
Derosa 2011b° *	niedrig	nein	ja	nein	nein	niedrig
Pan 2012°	niedrig	unklar	ja	nein	nein	niedrig
DeFronzo 2012°	niedrig	unklar	ja	nein	nein	niedrig
Gallwitz 2012°	niedrig	ja	ja	nein	nein	niedrig
Yang 2012°	niedrig	unklar	ja	nein	nein	niedrig

° aktualisierte Literaturrecherche

* werden nicht berücksichtigt, da ein intensives Diätprogramm Bestandteil der Studien waren → Sensitivitätsanalyse

Begründen Sie für jede Studie die abschließende Einschätzung.

Die **Gesamtrate der unerwünschten Ereignisse** wurde unterschiedlich operationalisiert. In den Studien MB102014, D1690C00012, Charbonnel 2006, Scott 2008, Bolli 2008, Bergenstal 2010b, D1690C00004, Matthews 2005, Matthews 2010, Göke 2010, Filozof 2010 und Gallwitz 2012 (Gallwitz 2012 wurde in der aktualisierten Literaturrecherche identifiziert) wurde das Verzerrungspotenzial des Endpunktes als niedrig angesehen. Der Endpunkterheber war verblindet, das ITT-Prinzip wurde adäquat umgesetzt, es wurden weder Hinweise auf ergebnisgesteuerte Berichterstattung noch auf sonstige das Verzerrungspotenzial beeinflussende Punkte gefunden.

Kaku 2009 war bereits auf Studienebene hoch verzerrt; es lagen keine Informationen über die Verblindung des Endpunkterheber vor und die Ergebnisse des Endpunktes waren möglicherweise aufgrund des unterschiedlichen Alters und der unterschiedlichen Diabetesdauer der Patienten in den Behandlungsgruppen verzerrt.

In den Studien von DeFronzo 2009, Raz 2008, Taskinen 2011, DeFronzo 2005, Nauck 2009 Nauck 2007, Yang 2011, Pan 2012 und Yang 2012 (Yang 2011, Pan 2012 und Yang 2012 wurden in der aktualisierten Literaturrecherche identifiziert) lagen keine Informationen über die Verblindung des Endpunkterhebers vor, es wurde von einem niedrigen Verzerrungspotenzial ausgegangen. Es gab aber keine weiteren Hinweise auf eine Verzerrung, so dass dieser Endpunkt als niedrig verzerrt eingestuft werden konnte. In der Studie von Pratley 2010 war der Endpunkterheber nicht verblindet. Es bestand daher ein Verzerrungspotenzial bei der Berichterstattung von Nebenwirkungen (z. B. Übelkeit), da es sich um eine offene Studie handelte und GLP-1-Analoga (z. B. Liraglutide) bekannt für das Erzeugen von Übelkeit sind.

Die Publikation von Bosi 2007 war bereits auf Studienebene hoch verzerrt, die Verblindung des Endpunkterhebers war gewährleistet, jedoch wurde das ITT-Prinzip unzureichend dargestellt, so dass dieser Endpunkt als hoch verzerrt anzusehen war.

Die **Gesamtrate schwerwiegender unerwünschter Ereignisse** wurde gleich operationalisiert. In den Studien MB102014, D1690C00012, Charbonnel 2006, Scott 2008, Bolli 2008, Bergenstal 2010, D1690C00004, Matthews 2005, Matthews 2010, und Göke 2010 war das Verzerrungspotenzial des Endpunkts als niedrig anzusehen. Der Endpunkterheber war verblindet, das ITT-Prinzip wurde adäquat umgesetzt, es wurden weder Hinweise auf ergebnisgesteuerte Berichterstattung noch auf sonstige das Verzerrungspotenzial beeinflussende Punkte gefunden.

Die Publikation von Kaku 2009 war bereits auf Studienebene hoch verzerrt; es lagen keine Informationen über die Verblindung des Endpunkterheber vor und die Ergebnisse des Endpunktes wurden möglicherweise aufgrund des unterschiedlichen Alters und der unterschiedlichen Diabetesdauer der Patienten in den Behandlungsgruppen verzerrt.

In den Studien von DeFronzo 2009, Raz 2008, Taskinen 2011, DeFronzo 2005, Nauck 2009 Nauck 2007, Yang 2011, Pan 2012 und Yang 2012 (Yang 2011, Pan 2012 und Yang 2012 wurden in der aktualisierten Literaturrecherche identifiziert) lagen keine Informationen über die Verblindung des Endpunkterhebers vor, es wurde aber von einem niedrigen Verzerrungspotenzial ausgegangen. Bei der Studie von Pratley 2010 war der Endpunkterheber nicht verblindet, es gab aber keine weiteren Hinweise auf eine Verzerrung, so dass dieser Endpunkt als niedrig verzerrt eingestuft werden konnte.

Die Publikation von Bosi 2007 war bereits auf Studienebene hoch verzerrt, die Verblindung des Endpunkterhebers war gewährleistet, jedoch wurde das ITT-Prinzip unzureichend dargestellt, so dass dieser Endpunkt als hoch verzerrt anzusehen war.

Die **Studienabbrüche aufgrund unerwünschter Ereignisse** wurden gleich operationalisiert. In den Studien von MB102014, D1690C00012, Charbonnel 2006, Scott 2008, Bolli 2008, Bergenstal 2010, D1690C00004, Matthews 2005, Matthews 2010, Göke 2010, Filozof 2010, Derosa 2011a und Gallwitz 2012 (Derosa 2011a und Gallwitz 2012 sind aus der aktualisierten Literaturrecherche) war das Verzerrungspotenzial des Endpunktes als niedrig anzusehen. Der Endpunkterheber war verblindet, das ITT-Prinzip wurde adäquat umgesetzt, es wurden weder Hinweise auf ergebnisgesteuerte Berichterstattung noch auf sonstige das Verzerrungspotenzial beeinflussende Punkte gefunden.

Die Publikation von Kaku 2009 war bereits auf Studienebene hoch verzerrt; es lagen keine Informationen über die Verblindung des Endpunkterheber vor und die Ergebnisse des Endpunktes wurden möglicherweise aufgrund des unterschiedlichen Alters und der unterschiedlichen Diabetesdauer der Patienten in den Behandlungsgruppen verzerrt.

In den Studien von DeFronzo 2009, Raz 2008, Taskinen 2011, DeFronzo 2005, Nauck 2009 Nauck 2007, Yang 2012, Pan 2012 und Yang 2012 (Yang 2011, Pan 2012 und Yang 2012 wurden in der aktualisierten Literaturrecherche identifiziert) lagen keine Informationen über die Verblindung des Endpunkterhebers vor, es wurde aber von einem niedrigen Verzerrungspotenzial ausgegangen. Es gab aber keine weiteren Hinweise auf eine Verzerrung, so dass dieser Endpunkt als niedrig verzerrt eingestuft werden konnte. Bei Derosa 2011b (aus der aktualisierten Literaturrecherche) war der Endpunkterheber nicht verblindet, es gab aber keine weiteren verzerrungsbeeinflussende Punkte, so dass von einem niedrigen Verzerrungspotential ausgegangen werden kann. Bei Pratley 2010 war der Endpunkterheber nicht verblindet. Es bestand daher ein Verzerrungspotenzial bei der Berichterstattung von Nebenwirkungen (z. B. Übelkeit), da es sich um eine offene Studie handelt und GLP-1-Analoga (z.B. Liraglutide) bekannt für das Erzeugen von Übelkeit sind.

Die Studie von Bosi 2007 war bereits auf Studienebene hoch verzerrt, die Verblindung des Endpunkterhebers war gewährleistet, jedoch wurde das ITT-Prinzip unzureichend dargestellt, so dass dieser Endpunkt als hoch verzerrt anzusehen war.

Bei DeFronzo 2012 wurde die Gesamtrate unerwünschter Ereignisse, die Gesamtrate schwerwiegender unerwünschter Ereignisse und die Studienabbrüche aufgrund unerwünschter

Ereignisse nur für die gepoolte Pioglitazon-Gruppe und die Kombination von Alogliptin+Pioglitazon berichtet. Dieser Vergleich ist für die Nutzenbewertung allerdings nicht relevant und wurde nicht dargestellt.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-127: Ergebnisse für Gesamtrate unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 24 Wochen)

Studien	N	n (%)	N	n (%)	Odds Ratio [95% KI]
		Metformin+DPP-4-Inhibitoren		Metformin+Placebo	
Charbonnel 2006	464	262 (56,5)	237	128 (54,0)	1,10 [0,81;1,51]
DeFronzo 2009	191	133 (69,6)	179	116 (92,7)	1,25 [0,81;1,92]
Raz 2008	96	55 (57,3)	94	56 (59,6)	0,91 [0,51;1,62]
Bosi 2007	183	119 (65,0)	181	115 (65,5)	1,07 [0,69;1,64]
Taskinen 2011	523	276 (53,3)	177	98 (83,8)	0,90 [0,64;1,27]
Scott 2008	94	37 (39,4)	92	27 (29,3)	1,54 [0,84;2,84]
		Metformin+TZD		Metformin+Placebo	
Kaku 2009	83	71 (85,5)	86	64 (74,4)	2,03 [0,93;4,44]
		Metformin+Dapagliflozin		Metformin+Placebo	
MB102014	135	98 (72,6)	137	88 (64,2)	1,47 [0,88;2,47]
D1690C00012	91	39 (42,9)	91	36 (39,6)	1,15 [0,63;2,07]
		Metformin+TZD		Metformin+DPP-4 Inhibitoren	
Bolli 2008	280	158 (56,4)	295	177 (60,0)	0,86 [0,62;1,20]
		Metformin+GLP-1-Analoga		Metformin+DPP-4 Inhibitoren	
Pratley 2010*	439	305 (69,5)	219	127 (58,0)	1,65 [1,18;2,31]

*2 verschiedene Dosierungen von GLP-1-Analoga (werden addiert)

KI: Konfidenzintervall, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Aufgrund der unterschiedlichen Operationalisierungen / klinischen Heterogenität wurden keine Gesamtschätzer und damit auch keine MTC berechnet. Nur ein signifikantes Einzelergebnis zeigte sich bei Pratley 2010, da es eine signifikant erhöhte Rate der unerwünschten Ereignisse bei der Behandlungsgruppe Metformin+DPP-4-Inhibitoren im Gegensatz zu Metformin+GLP-1-Analoga gab.

Tabelle 4-128: Ergebnisse für Gesamtrate unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoffe)

Studien	N	n (%)	N	n (%)	Odds Ratio
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff		
Filozof 2010	510	315 (61,8)	493	302 (63,3)	1,02 [0,79;1,32]
Nauck 2007	588	419 (71,3)	584	444 (76,0)	0,78 [0,60;1,01]
Göke 2010	428	257 (60,0)	430	244 (56,7)	1,15 [0,87;1,15]
	Metformin+TZD		Metformin+Sulfonylharnstoff		
Matthews 2005	317	176 (55,5)	313	182 (58,1)	0,90 [0,66;1,23]
	Metformin+Dapagliflozin		Metformin+Sulfonylharnstoff		
D1690C00004	406	318 (78,3)	408	318 (77,9)	1,02 [0,73;1,43]

TZD: Thiazolidinedione, DPP-4: Dipeptidyl-Peptidase-4

Auch für die Gesamtrate unerwünschter Ereignisse der einzelnen Studien nach 52 Wochen wurde aufgrund der klinischen Heterogenität kein Gesamtschätzer und damit auch keine MTC berechnet. Es ergaben sich keine signifikanten Unterschiede der Einzelergebnisse.

Auch in den Publikationen von Yang 2011, Pan 2012 und Yang 2012, die in der aktualisierten Literaturrecherche gefunden wurden, ergaben sich keine statistisch signifikanten Ergebnisse. Gallwitz 2012 berichtete die Gesamtrate der unerwünschten Ereignisse nach 104 Wochen. Hierbei zeigte sich ein signifikanter Unterschied zugunsten von Metformin+DPP-4-Inhibitoren gegenüber Metformin+Sulfonylharnstoff (was aber wahrscheinlich an der höheren Hypoglykämierate der Metformin+Sulfonylharnstoff-Gruppe lag).

Tabelle 4-129: Ergebnisse für Gesamtrate unerwünschter Ereignisse aus RCT für indirekte Vergleiche –aktualisierte Literaturrecherche

Studien	N	n (%)	N	n (%)	Odds Ratio [95%-KI]
Nach 24 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Placebo		
Yang 2011	283	124 (43,8)	287	119 (41,5)	1,10 [0,79;1,53]
Pan 2012	146	50 (34,2)	144	54 (37,5)	0,87 [0,54;1,40]
Yang 2012	198	56 (28,4)	197	60 (30,3)	0,91 [0,59;1,41]
Nach 104 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff		
Gallwitz 2012	776	663 (85)	775	706 (91)	0,57 [0,42;0,79]

KI: Konfidenzintervall, DPP-4: Dipeptidyl-Peptidase-4

Wegen der Heterogenität und der unterschiedlichen Operationalisierung wurde für die Gesamtrate unerwünschter Ereignisse keine MTC durchgeführt, da durch das unterschiedliche Erfassen von Hypoglykämien bzw. Laborveränderungen in den unerwünschten Ereignissen die Ergebnisse verzerrt waren.

Tabelle 4-130: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 24 Wochen)

Studien	N	n (%)	N	n (%)	Odds Ratio [95%-KI]
		Metformin+DPP-4-Inhibitoren		Metformin+Placebo	
Charbonnel 2006	464	13 (2,8)	237	7 (3,0)	0,95 [0,37;2,41]
DeFronzo 2009	191	8 (4,2)	179	5 (2,8)	1,52 [0,49;4,74]
Raz 2008	96	0 (0,0)	94	5 (5,3)	0,08 [0,00;1,55]
Bosi 2007	183	5 (2,7)	181	8 (4,4)	0,61 [0,19;1,89]
Taskinen 20011	523	18 (3,4)	177	4 (2,3)	1,54 [0,51;4,62]
Scott 2008	94	5 (5,3)	92	5 (5,4)	0,97 [0,27;3,46]
Gesamtschätzer (mit Heterogenitätsmaß)^o					0,98 [0,61;1,59] I ² =0% ;p=0,451
		Metformin+TZD		Metformin+Placebo	
Kaku 2009	83	5 (6,0)	86	2 (2,3)	2,69 [0,51;14,28]
		Metformin+Dapagliflozin		Metformin+Placebo	
MB102014	135	4 (3,0)	137	5 (3,6)	0,81 [0,21;3,07]
D1690C00012	91	6 (6,6)	91	1 (1,1)	6,35 [0,75;53,87]
Gesamtschätzer (mit Heterogenitätsmaß)^o					1,91 [0,25;14,52] I ² =62,4%;p=0,103
		Metformin+GLP-1-Analoga		Metformin+Placebo	
DeFronzo 2005	223	8 (3,6)	113	4 (3,5)	1,01 [0,30;3,44]
		Metformin+TZD		Metformin+DPP-4-Inhibitoren	
Bolli 2008	280	13 (4,6)	295	6 (2,0)	2,35 [0,88;6,26]
Bergental 2010	165	10 (6,1)	166	5 (3,0)	2,08 [0,69;6,22]
Gesamtschätzer (mit Heterogenitätsmaß)^o					2,22 [1,07;4,62] I ² =0%; p=872
		Metformin+GLP-1-Analoga		Metformin+DPP-4-Inhibitoren	
Pratley 2010*	439	12 (2,7)	219	4 (1,8)	1,51 [0,48;4,74]

*2 verschiedene Dosierungen von GLP-1-Analoga (werden addiert)

^oanhand Modell mit zufälligen Effekten

KI: Konfidenzintervall, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Tabelle 4-131: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoffe)

Studien	N	n (%)	N	n (%)	Odds Ratio [95%-KI]
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoffe		
Nauck 2007	588	43 (7,3)	584	44 (7,5)	0,97 [0,63;1,50]
Göke 2010	428	39 (9,1)	430	32 (7,4)	1,25 [0,77;2,03]
Gesamtschätzer (mit Heterogenitätsmaß)^o					1,08 [0,78;1,50] I ² =0% ; p=0,449
	Metformin+TZD		Metformin+Sulfonylharnstoffe		
Matthews 2005	317	15 (4,7)	313	20 (6,4)	0,73 [0,37;1,45]
	Metformin+Dapagliflozin		Metformin+Sulfonylharnstoffe		
D1690C00004	406	35 (8,6)	408	46 (11,3)	0,74 [0,47;1,18]

^oanhand Modell mit zufälligen Effekten

KI: Konfidenzintervall, TZD: Thiazolidinedione, DPP-4: Dipeptidyl-Peptidase-4

Es gab keine signifikanten Unterschiede in der Gesamtrate der schwerwiegenden unerwünschten Ereignisse bei den direkten paarweisen Vergleichen.

Auch in den Publikationen von Yang 2011, Pan 2012 und Yang 2012, die in der aktualisierten Literaturrecherche gefunden wurden, ergaben sich keine statistisch signifikanten Ergebnisse. Gallwitz 2012 berichtete die Gesamtrate der schwerwiegenden unerwünschten Ereignisse nach 104 Wochen. Hierbei zeigte sich ebenfalls kein signifikanter Unterschied zugunsten von Metformin+DPP-4-Inhibitoren oder Metformin+Sulfonylharnstoff.

Tabelle 4-132: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT für indirekte Vergleiche –aktualisierte Literaturrecherche

Studien	N	n (%)	N	n (%)	Odds Ratio [95%-KI]
Nach 24 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Placebo		
Yang 2011	283	8 (2,8)	287	3 (1,0)	2,75 [0,72;10,49]
Pan 2012	146	0 (0,0)	144	1 (0,7)	0,33 [0,01;8,08]
Yang 2012	197	7 (3,6)	198	5 (2,5)	1,42 [0,44; 4,56]
Nach 104 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff		
Gallwitz 2012	776	135 (17)	775	162 (21)	0,80 [0,62;1,03]

KI: Konfidenzintervall, DPP-4: Dipeptidyl-Peptidase-4

Tabelle 4-133: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 24 Wochen)

Studien	N	n (%)	N	n (%)	Odds Ratio [95% KI]
	Metformin+DDP-4-Inhibitoren		Metformin+Placebo		
Charbonnel 2006	464	11 (2,4)	237	7 (2,9)	0,80 [0,31;2,09]
DeFronzo 2009	191	6 (3,1)	179	2 (11,1)	2,87 [0,57;14,41]
Raz 2008	96	0 (0,0)	94	2 (2,1)	0,19 [0,01;4,05]
Bosi 2007	183	5 (2,7)	181	8 (4,4)	0,61 [0,19;1,89]
Taskinen 20011	523	8 (1,5)	177	3 (1,7)	0,90 [0,24;3,43,]
Scott 2008	94	2 (2,1)	92	1 (1,1)	1,96 [0,17;21,96]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					1,19 [0,67;2,12] I ² =0% ;p=0,515
	Metformin+TZD		Metformin+Placebo		
Kaku 2009	83	5 (6,0)	86	2 (2,3)	2,69 [0,51;14,28]
	Metformin+Dapagliflozin		Metformin+Placebo		
MB102014	135	4 (3,0)	137	5 (3,6)	0,81 [0,21;3,07]
D1690C00012	91	4 (4,3)	91	0 (0,0)	9,41 [0,50;177,39]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					1,97 [0,18;21,65] I ² =58,2% p=0,122
	Metformin+GLP-1-Analoga		Metformin+Placebo		
DeFronzo 2005	223	12 (5,4)	113	1 (0,9)	6,37 [0,28;49,62]
Nauck 2009	482	52 (10,8)	121	2 (1,7)	7,20 [1,73;29,97]
Gesamtschätzer (mit Heterogenitätsmaß)[°]					6,92 [2,14;22,32] I ² =0%;p=0,924
	Metformin+TZD		Metformin+DPP-4 Inhibitoren		
Bolli 2008	280	9 (3,2)	295	9 (3,1)	1,06 [0,41;2,70]
Bergental 2010	165	6 (3,6)	166	5 (3,0)	1,22 [0,36;4,06]
					1,11 [0,53;2,33]
	Metformin+GLP-1-Analoga		Metformin+DPP-4- Inhibitoren		
Pratley 2010*	439	29 (6,6)	219	4 (1,8)	3,80 [1,32;10,96]

[°]anhand Modell mit zufälligen Effekten

*2 verschiedene Dosierungen von GLP-1 Analoga (werden addiert)

KI: Konfidenzintervall, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Tabelle 4-134: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche (nach 52 Wochen im Vergleich zu Metformin+Sulfonylharnstoffe)

Studien	N	n (%)	N	n (%)	Odds Ratio [95% KI]
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff		
Nauck 2007	588	16 (2,7)	584	21 (3,6)	0,97 [0,75;1,45]
Göke 2010	428	18 (4,2)	430	19 (4,4)	0,95 [0,49;1,84]
Filozof 2010	510	34 (6,7)	493	23 (4,7)	1,46 [0,85;2,52]
Gesamtschätzer (mit Heterogenitätsmaß)^o					1,05 [0,71;1,57] I ² =19,6% ; p=0,288
	Metformin+TZD		Metformin+Sulfonylharnstoff		
Matthews 2005	317	13 (4,1)	313	14 (4,5)	0,91 [0,42;1,98]
	Metformin+Dapagliflozin		Metformin+Sulfonylharnstoff		
D1690C00004	406	37 (9,1)	408	24 (5,9)	1,60 [0,94;2,73]
	Metformin+GLP1-Analoga		Metformin+Sulfonylharnstoff		
Derosa 2010*	63	4 (6,3)	65	4 (6,1)	1,03 [0,25;4,33]

^o anhand Modell mit zufälligen Effekten

* wird nicht berücksichtigt, da ein intensives Diätprogramm Bestandteil der Studien waren → Sensitivitätsanalyse

KI: Konfidenzintervall, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Es gab keine signifikanten Unterschiede in der Anzahl der Studienabbrüche mit Ausnahme von Metformin+GLP-1-Analoga versus Metformin+DPP-4-Inhibitoren und Metformin+GLP-1-Analoga versus Metformin+Placebo nach 24 Wochen. Es gab jeweils höhere Raten in der Metformin+GLP1-Analoga-Behandlungsgruppe.

In den Publikationen von Yang 2011, Pan 2012 und Yang 2012, die in der aktualisierten Literaturrecherche gefunden wurden, ergaben sich keine statistisch signifikanten Ergebnisse. Gallwitz 2012 berichtete Studienabbrüche aufgrund unerwünschter Ereignisse nach 104 Wochen. Hierbei zeigte sich ein signifikanter Unterschied zugunsten von Metformin+DPP-4-Inhibitoren gegenüber Metformin+Sulfonylharnstoff.

Tabelle 4-135: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche –aktualisierte Literaturrecherche

Studien	N	n (%)	N	n (%)	Odds Ratio [95%-KI]
Nach 24 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Placebo		
Yang 2011	283	6 (2,1)	287	3 (1,0)	2,05 [0,51;8,28]
Pan 2012	146	2 (1,4)	144	2 (1,4)	0,99 [0,14;7,10]
Yang 2012	197	8 (4,1)	198	4 (2,0)	2,05 [0,60;6,93]
Nach 52 Wochen					
	Metformin+TZD		Metformin+Sulfonylharnstoff		
Derosa 2011a *	102	2 (2,0)	99	3 (3,0)	0,64 [0,10;3,91]
	Metformin+GLP-1-Analoga		Metformin+Sulfonylharnstoff		
Derosa 2011b * °	57	4 (7,0)	54	4 (7,4)	0,94 [0,22;3,98]
Nach 104 Wochen					
	Metformin+DPP-4-Inhibitoren		Metformin+Sulfonylharnstoff		
Gallwitz 2012	776	60 (8)	775	85 (11)	0,68 [0,48;0,96]

* werden nicht in die primäre Analyse nicht miteinbezogen → Sensitivitätsanalyse ° Publikation von Salvadeo 2010

KI: Konfidenzintervall, TZD: Thiazolidinedione, GLP-1: Glucagon-like peptide-1, DPP-4: Dipeptidyl-Peptidase-4

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.

Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.

Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.

Tabelle 4-136: Ergebnisse für schwerwiegende unerwünschte Ereignisse aus indirekten Vergleichen nach 24 und 52 Wochen - Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel

Zu bewertende Arzneimittel	Komparator	Odds Ratio (95%- Kredititätsintervall*)
Nach 24 Wochen		
Metformin+Dapagliflozin	Metformin+DPP-4-Inhibitoren	2,80 [0,53;9,53]
Nach 52 Wochen		
Metformin+Dapagliflozin	Metformin+ DPP-4-Inhibitoren	1,02 [0,11;4,09]

*basierend auf dem best angepassten Modell

DPP-4: Dipeptidyl-Peptidase-4

Für die Gesamtrate schwerwiegender unerwünschter Ereignisse zeigte sich weder nach 24 Wochen noch nach 52 Wochen ein statistisch signifikanter Unterschied zwischen Metformin+Dapagliflozin und Metformin+DPP-4-Inhibitoren. Auch die Publikationen von Yang 2011, Pan 2012 und Yang 2012 würden das Ergebnis der MTC nach 24 Wochen nicht ändern (aktualisierte Literaturrecherche).

Um die Annahme der Konsistenz zu überprüfen, wurden die Resultate der *Head-to-Head*-Vergleiche mit den Resultaten aus den MTC verglichen. Es wurden keine Inkonsistenzen festgestellt.

Tabelle 4-137: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse aus indirekten Vergleichen nach 24 und 52 Wochen - Metformin+DPP-4-Inhibitoren mit dem zu bewertenden Arzneimittel

Zu bewertende Arzneimittel	Komparator	Odds Ratio (95%- Kredititätsintervall*)
Nach 24 Wochen		
Metformin+Dapagliflozin	Metformin+DPP-4-Inhibitoren	2,10 [0,31;7,88]
Nach 52 Wochen		
Metformin+Dapagliflozin	Metformin+ DPP-4-Inhibitoren	2,25 [0,30;8,24]

*basierend auf dem best angepassten Modell

DPP-4: Dipeptidyl-Peptidase-4

Für die Anzahl der Studienabbrüche aufgrund unerwünschter Ereignisse zeigte sich weder nach 24 Wochen noch nach 52 Wochen ein statistisch signifikanter Unterschied zwischen Metformin+Dapagliflozin und Metformin+DPP-4-Inhibitoren. Auch die Publikationen von Yang 2011, Pan 2012 und Yang 2012 würden das Ergebnis der MTC nach 24 Wochen nicht ändern (aktualisierte Literaturrecherche).

In den Studien Derosa 2011a und Derosa 2011b (aktualisierte Literaturrecherche), die neben Derosa 2010 in die Sensitivitätsanalyse einfließen würden, ergaben sich keine signifikanten Unterschiede zugunsten einer der Behandlungsoptionen. Daher würde man keinen Unterschied in der Gesamtinterpretation der MTC erwarten.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.2.

Wegen unzureichender Datenverfügbarkeit wurden für die indirekten Vergleiche keine Subgruppenanalysen durchgeführt.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Es wurde keine Suche nach nicht-randomisierten Studien durchgeführt.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-138: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Andere Aspekte, die Verzerrungen verursachen können
			Patient	Behandler		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-139: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4-140: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver- zerrungspotenzial be- einflussende Punkte
<Studie 1>	ja / nein / unklar	ja / nein / unklar	ja / nein / unklar	ja / nein

1) Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar.

Es wurden keine nicht randomisierten Vergleiche durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.2.

Es wurden keine nicht randomisierten Vergleiche durchgeführt.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

- Studien des pharmazeutischen Unternehmers

- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungspotenzial.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Es wurden keine weiteren Untersuchungen zur Bewertung des Zusatznutzen eingesetzt.

4.3.2.3.3 < Ergebnisse aus weiteren Untersuchungen

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-141: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Es wurden keine nicht randomisierten Vergleiche durchgeführt.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung.

Es wurden keine nicht randomisierten Vergleiche durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.2.

Es wurden keine nicht randomisierten Vergleiche durchgeführt.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

In Bezug auf Metformin+DPP-4-Inhibitoren, welche, wie in Abschnitt 4.2.2 erläutert, die zweckmäßige Vergleichstherapie für diejenigen Patienten darstellt, die nicht mit den Sulfonylharnstoffen behandelt werden können, kann man anhand der Ergebnisse der MTC zusammenfassend sagen, dass:

Metformin+Dapagliflozin und Metformin+DPP-4-Inhibitoren vergleichbare Ergebnisse zeigten bzgl. der **Blutzuckerkontrolle** nach 24 Wochen (MWD [95%-Kredibilitätsintervall]: 0,16 [-0,17;0,49]) und nach 52 Wochen (MWD [95%-Kredibilitätsintervall]: -0,08 [-0,25;0,10]).

Metformin+Dapagliflozin und Metformin+DPP-4-Inhibitoren vergleichbare Ergebnisse zeigten bzgl. der **Anzahl der Hypoglykämien** nach 24 Wochen (OR [95%-Kredibilitätsintervall]: 1,75 [0,31;5,73] und nach 52 Wochen (OR [95%-Kredibilitätsintervall]: 0,81 [0,18;2,59]).

Metformin+Dapagliflozin das **Gewicht** sowohl nach 24 als auch nach 52 Wochen statistisch signifikant stärker senkte (MWD [95%-Kredibilitätsintervall]: -2,37 [-2,98;-1,77] / -2,74 [-5,35;-0,10]).

Metformin+Dapagliflozin und Metformin+DPP-4-Inhibitoren vergleichbare Ergebnisse zeigten bzgl. der **systolischen Blutdrucksenkung** nach 24 Wochen (MWD [95%-Kredibilitätsintervall]: -3,13[-7,19;0,80]).

Metformin+Dapagliflozin und Metformin+DPP-4-Inhibitoren vergleichbare Ergebnisse zeigten beim Auftreten von **schwerwiegenden unerwünschten Ereignissen** (OR [95%-Kredibilitätsintervall]: 2,80 [0,53;9,53] / 1,02 [0,11;4,09]) und **Studienabbrüchen aufgrund von unerwünschten Ereignissen** (OR [95%-Kredibilitätsintervall]: 2,10 [0,31;7,88] / 2,25 [0,30;8,24]) jeweils nach 24 und nach 52 Wochen. Der Vergleich bezüglich der Gesamtrate **unerwünschter Ereignisse** konnte aufgrund der unterschiedlichen Operationalisierungen in den einzelnen Studien nicht durchgeführt werden.

Es wurde noch eine aktualisierte Literaturrecherche durchgeführt, in der weitere Publikationen gefunden wurden. Diese konnten aber nicht mehr in die MTC einfließen; die Ergebnisse wurden dennoch dargestellt und bestätigten die Ergebnisse der in die MTC einfließenden Studien. Es würde daher kein Einfluss auf die Gesamtinterpretation der MTC geben.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Zur Beantwortung der primären Fragestellung lag eine relevante Studie vor. Die Studie ist eine randomisierte, kontrollierte Studie und entspricht demnach der Evidenzstufe Ib. Die Studie beinhaltet einen direkten Vergleich mit der zweckmäßigen Vergleichstherapie. Nach Prüfung aller möglichen Verzerrungsaspekte ist sowohl auf Studienebene als auch auf Endpunktebene von einem niedrigen Verzerrungspotenzial auszugehen. Die betrachteten Endpunkte sind in ihren Operationalisierungen valide und stellen, wie in Abschnitt 4.2.5.2 begründet, patientenrelevante Endpunkte dar. Darüber hinaus wurden Änderungen im systolischen und diastolischen Blutdruck als Surrogatparameter für das kardiovaskuläre Risiko berücksichtigt (siehe dazu auch 4.5.4). Die Validität der herangezogenen Endpunkte ist somit insgesamt als hoch einzuschätzen.

Im Hinblick auf die Evidenzstufe, Studienqualität und Validität der Endpunkte besitzen die erbrachten Nachweise zum Nutzen von Dapagliflozin im Vergleich zu einem Sulfonylharnstoff anhand einer randomisierten *Head-to-Head*-Studie eine hohe Validität und Aussagekraft.

Für alle diskutierten Endpunkte können deswegen, im Falle statistisch signifikanter Ergebnisse, Aussagen gemacht werden, die als Hinweis gewertet werden müssen.

Zur Erfüllung der Anforderungen an ein Dossier zur Nutzenbewertung gemäß G-BA und IQWiG-Vorgaben sind im Rahmen der vorliegenden Nutzenbewertung Post-hoc-Analysen durchgeführt worden.

Die in diesem Dossier zur Nutzenbewertung durchgeführten Post-hoc-Analysen und die daraus abgeleiteten Aussagen basieren methodisch auf dem Methodenpapier zur Nutzenbewertung des IQWiG (IQWiG, 2011b). Ein wesentlicher Bestandteil dieser Methodik ist die Darstellung des (Zusatz-) Nutzens anhand von Effektschätzern aus den Einzelstudien inkl. Angabe von 95% Konfidenzintervallen, nominellen p-Werten statistischer Testverfahren und „statistischer Signifikanz“ der Ergebnisse basierend auf den nominellen p-Werten ohne, dass eine Adjustierung der p-Werte für die Anzahl der durchgeführten statistischen Tests erfolgte. Auch wurden *post hoc* statistische Tests unabhängig von der Anzahl der in der

jeweiligen Analyse berücksichtigten Patienten / bzw. Ereignisse durchgeführt, was insbesondere bei einer kleinen Patienten- und/oder Fallzahl in der Interpretation der Ergebnisse berücksichtigt werden muss. Effektschätzer und Signifikanzen, die für die Verlängerungsphase einer Studie berechnet werden, müssen mit besonderer Vorsicht interpretiert werden, da der Anteil der Patienten, die die Studie abbrechen oder eine Notfall-Medikation erhalten in der Verlängerungsphase ansteigt.

Darüber hinaus wurde zur Beantwortung der nachgelagerten Fragestellung der Vergleich von Metformin+Dapagliflozin gegen Metformin+DPP-4-Inhibitoren anhand einer MTC vorgenommen. In diese sind ausschließlich RCT eingeflossen, da sie mit der geringsten Ergebnisunsicherheit behaftet sind. Darüber hinaus waren die Studien größtenteils doppelt verblindet und die Charakteristika der Studienpopulationen weitgehend homogen.

Die durchgeführte MTC erfüllt damit den momentanen Stand der Wissenschaft hinsichtlich der Anforderungen an eine solche Analyse.

Einschränkungen ergeben sich allerdings hinsichtlich der Operationalisierung einzelner Endpunkte. Für HbA1c und Gewichtsveränderungen war die Operationalisierung konsistent. Beim Blutdruck wurde teils nur der systolische Blutdruck berichtet. Es fehlten zudem die Angaben, in welcher Körperlage der Blutdruck gemessen wurde. Die Operationalisierung für den Endpunkt „Hypoglykämien“ war heterogen. Diese Heterogenität hatte jedoch keinen Einfluss auf die relative Effektgröße zwischen den einzelnen Komparatoren. Auch die Operationalisierung der Gesamtrate der unerwünschten Ereignisse war heterogen. Das führte in diesem Fall dazu, dass dieser Endpunkt nicht für die MTC berücksichtigt werden konnte.

Vor diesem Hintergrund können anhand der Ergebnisse der MTC Aussagen getroffen werden, die als Anhaltspunkt gewertet werden müssen.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*

- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Beschreibung des Zusatznutzens

Die hier betrachtete Patientenpopulation schließt erwachsene Patienten mit manifestem Diabetes mellitus Typ 2 ein, die unter einer Metformin-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen.

Gemäß den evidenzbasierten Therapieleitlinien wird bei den Patienten, die ihren Blutzucker nicht mehr ausreichend kontrollieren können, die Kombinationstherapie mit zwei verschiedenen OAD als Zweitlinientherapie bzw. OAD+GLP-1-Agonisten empfohlen. Als OAD werden Alpha-Glucosidasehemmer (Acarbose), DPP-4-Inhibitoren (Gliptine), Glitazone, Sulfonylharnstoffe und Sulfonylharnstoff-Analoga (Glinide), jeweils in Kombination mit Metformin, genannt (Matthaei et al., 2009).

Die meisten für diese Patienten in Frage kommenden Behandlungsoptionen, die das Ziel haben den Blutzuckerspiegel zu senken, fokussieren sich auf die Stimulation der Insulinsekretion (z. B. Sulfonylharnstoffe, Glinide, DPP-4-Inhibitoren) und sind damit abhängig von der Insulinverfügbarkeit. Da sich die Insulinresistenz und die Betazellfunktion mit zunehmender Erkrankungsdauer verschlechtern, können einige dieser Substanzen mit der Zeit ihre Wirksamkeit verlieren (UK Prospective Diabetes Study Group, 1998c; Viberti et al., 2002). Für die DPP-4-Inhibitoren konnte eine Langzeitwirksamkeit nachgewiesen werden (Göke et al., 2010), wohingegen insbesondere die Wirksamkeit der Sulfonylharnstoffe in der Regel im Behandlungsverlauf schneller nachlässt als die der anderen OAD (Kahn et al., 2006).

Darüber hinaus bewirken einige Antidiabetika (z. B. Insulin, Sulfonylharnstoffe, Glinide) eine Gewichtszunahme und das vermehrte Auftreten von Hypoglykämien (AkdÄ, 2009). Glitazone sind aufgrund eines erhöhten Risikos für Frakturen und kardiovaskuläre Ereignisse in Deutschland nicht mehr zu Lasten der Gesetzlichen Krankenversicherung verordnungsfähig (G-BA, 2010). Alpha-Glukosidasehemmer rufen aufgrund ihrer Wirkungsweise gastrointestinale Nebenwirkungen hervor, die beim Patienten zu einer verminderten Therapie-Adhärenz führen können und spielen hinsichtlich der verordneten Tagesdosen heute nur noch eine untergeordnete Rolle (AkdÄ, 2009; Mengel K., 2011). Schließlich ist die Kombination von Sulfonylharnstoff und Metformin wegen des unklaren Sicherheitsrisikos bezüglich kardiovaskulärer Endpunkte kritisch zu sehen (AkdÄ, 2009). Dadurch besteht ein therapeutischer Bedarf für neue Behandlungsstrategien für Typ-2-

Diabetiker, welche wirksam und sicher sind und auch bei einer Insulinresistenz bzw. geringer Insulinverfügbarkeit den Blutzucker noch effektiv senken können

Dapagliflozin ist solch eine neue Behandlungsoption und ist bisher das erste zugelassene Medikament seiner Wirkstoffklasse zur Verbesserung der Blutzuckerkontrolle. Dapagliflozin ist ein Vertreter der so genannten SGLT-2-Inhibitoren. Bei SGLT-2 handelt es sich um ein Carrier-Protein, das in der menschlichen Niere im S1 Segment des proximalen Tubulus 90% der Glukose aus dem Primärharn in das Blut rückresorbiert. Die restlichen 10% werden über den SGLT-1 Transporter rückresorbiert, welcher im S3 Segment des proximalen Tubulus lokalisiert ist und eine sehr hohe Glukoseaffinität besitzt (Brooks und Thacker, 2009). SGLT-2-Inhibitoren hemmen den renalen natriumabhängigen Glukosetransport in den Nierentubuli selektiv, führen damit zu einer Reduzierung der Glukoserückresorption und fördern so die konzentrationsabhängige Ausscheidung von Glukose über den Urin. Dies führt sowohl zu einer Senkung des Blutzuckers, als auch zu einem direkten Abtransport von Kalorien aus dem Körper und damit zu einem möglichen Gewichtsverlust. Dapagliflozin hat keinen Einfluss auf die endogene Glukose-Produktion, sein Wirkmechanismus ist unabhängig von der Insulinproduktion durch die Betazellen des Pankreas. Da die Menge der Glukose, die von den Nieren ausgeschieden wird, von der Blutglukosekonzentration abhängig ist und die Aktivität des SGLT-1-Transporters von der Hemmung unbeeinflusst bleibt, ist das Hypoglykämie-Risiko gering (Bristol-Myers Squibb und AstraZeneca GmbH, 2012a).

Im Folgenden werden die Ergebnisse zum medizinischen Zusatznutzen betrachtet. Sie basieren auf einer direkt vergleichenden RCT mit der freien Kombination der Wirkstoffe Metformin+Dapagliflozin gegen Metformin+Sulfonylharnstoff bei Typ-2-Diabetikern, die unter einer Metformin-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen. Patienten mit einer eGFR <60 mL/min/1,73 m² bzw. Kreatininclearance <60 mL/min, Patienten, die ≥75 Jahre alt sind und solchen, die mit einem Schleifendiuretikum behandelt werden, wurden ausgeschlossen (Zielpopulation gemäß Zulassung).

Gemeinsame Betrachtung der Blutzuckerkontrolle und des Auftretens von Hypoglykämien

Eine optimierte Blutzuckerkontrolle (und damit die HbA1c-Einstellung), wird von der DDG empfohlen (Zielvorgabe HbA1c <6,5% ohne das vermehrte Auftreten von Hypoglykämien und Gewichtszunahmen). Kann dieser Zielwert unter Monotherapie mit einem OAD nicht erreicht werden, wird die Kombinationstherapie mit einem anderen OAD bzw. einem GLP-1-Agonisten empfohlen (Matthaei et al., 2009). Der aktuelle Entwurf der Nationalen VersorgungsLeitlinie (NVL) zur Therapieplanung bei Typ-2-Diabetes empfiehlt einen HbA1c-Korridor von 6,5% bis 7,5% unter Berücksichtigung der individualisierten Therapieziele (Bundesärztekammer (BÄK) et al., 2012). Ähnliche Ausführungen zur Individualisierung der Therapieziele finden sich in einem aktuellen Positionspapier der Europäischen und Amerikanischen Diabetesgesellschaften (Consensus statement) (Inzucchi et al., 2012). In der großen klinischen Studie UKPDS wurde gezeigt, dass sich bereits bei einer Senkung des HbA1c um 1% das Risiko mikrovaskulärer Erkrankungen um 37% und das

Risiko makrovaskulärer Erkrankungen um 12-16% verringert (Stratton et al., 2000). In der ADVANCE-Studie bewirkte die intensivierete antihyperglykämische Therapie eine signifikante Reduktion des Neuauftretens bzw. der Progression der Nephropathie und des kombinierten Endpunktes „mikro- und makrovaskuläre Erkrankungen“ (Patel et al., 2008). Eine optimale Blutzuckereinstellung stellt somit ein wichtiges Therapieziel dar, insbesondere wenn sie unter Vermeidung von Nebenwirkungen wie z. B. Hypoglykämien erreicht werden kann.

In aktuellen Meta-Analysen prospektiv-randomisierter Studien bei Patienten mit Typ 2 Diabetes fand sich ferner eine signifikante Reduktion nicht-tödlicher Myokardinfarkte durch eine intensive Blutzucker-Senkung (Ray et al., 2009). Schließlich legt die Follow-up Untersuchung der Patienten der UKPD Studie die Schlußfolgerung nahe, dass eine frühe normnahe Blutzucker-Einstellung auf lange Sicht zu einer Reduktion mikro- und makrovaskulärer Endpunkte sowie einer Reduktion der Sterblichkeit führen kann (Holman et al., 2008).

Hypoglykämien stellen vor allem bei Patienten mit lange bestehendem Diabetes und makrovaskulären Veränderungen einen erheblichen Risikofaktor für kardiovaskuläre Episoden dar. So kann eine Hypoglykämie zu Störungen des myokardialen Blutflusses und zu myokardialen Ischämien führen und damit myokardiale Schäden verursachen. Niedrige Blutzuckerwerte in der Nacht können Veränderungen im EKG, insbesondere QTc-Veränderungen hervorrufen und zu Herzrhythmusstörungen und Herzfrequenzstörungen führen. Besonders Patienten mit kardiovaskulären Vorschädigungen haben ein hohes Risiko für kardiovaskuläre Episoden nach einer Hypoglykämie, die auch zum Tod führen können (Wiesner, 2010; Frier et al., 2011). Auch bestimmte Berufsgruppen wie z. B. LKW-Fahrer oder Personen, die in großer Höhe arbeiten, sowie allein lebende Senioren sind speziell hypoglykämiegefährdet. Die Gefahr von Hypoglykämien beeinträchtigt die Fahrtüchtigkeit und kann sogar zum Entzug der Fahrerlaubnis führen. Für die Erlaubnis zum Führen eines Fahrzeuges wird aus juristischer Perspektive eine stabile Diabeteseinstellung ohne Hypoglykämien vorausgesetzt (Ehlers, 2011). Aus diesem Grunde kann die Therapie mit einem Sulfonylharnstoff regelhaft zur Notwendigkeit von Fahrverboten mit entsprechenden Konsequenzen für Mobilität und Berufstätigkeit der Patienten führen. Aufgrund des gesteigerten Hypoglykämierisikos kann ferner eine regelmäßige Blutzucker-Selbstkontrolle angezeigt sein, was mit einer entsprechenden Belastung für den Patienten und auch einem Anstieg der Therapie assoziierten Kosten verbunden sein kann.

Erschwerend kommt hinzu, dass Todesfälle durch Hypoglykämien vermutlich nicht in vollem Umfang in der Literatur berichtet werden, so dass deren tatsächliche Zahl systematisch unterschätzt wird und dieses Problem in der Realität noch größer ist als allgemein bekannt (Inzucchi et al., 2012). Eine blutzuckersenkende Therapie, die das Auftreten von Hypoglykämien weitgehend vermeidet, stellt somit ein wichtiges Ziel der Diabetestherapie dar. Nicht zuletzt führen Hypoglykämien regelhaft zu Hospitalisierung und zusätzlicher Beanspruchung von Ressourcen des Gesundheitssystems (Holstein et al., 2003; Holstein et al., 2012).

In Hinblick auf die Interpretation der direkten Vergleichsstudie zwischen Dapagliflozin und Glipizid muss somit berücksichtigt werden, dass Probleme der mangelnden Adhärenz in kontrollierten klinischen Studien in der Regel nicht zum Tragen kommen, was möglicherweise zu einer systematischen Nutzen-Überbewertung von Sulfonylharnstoffen in klinischen Studien im Vergleich zur praktischen Therapie führt (Alvarez Guisasola et al., 2008).

Die Blutzuckerkontrolle, gemessen am HbA1c, unterschied sich nach 52 Wochen zwischen den beiden Behandlungsgruppen Metformin+Dapagliflozin versus Metformin+Sulfonylharnstoff kaum. Beide Behandlungen konnten den HbA1c-Wert effektiv und klinisch relevant ($>0,5\%$ Reduktion des HbA1c im Vergleich zum Ausgangswert) senken. Allerdings war der blutzuckersenkende Effekt im Vergleich zum Ausgangswert nach 104 Wochen Studiendauer bei der mit Metformin+Dapagliflozin behandelten Gruppe signifikant stärker ausgeprägt als bei den Patienten, die mit Metformin+Sulfonylharnstoff behandelt wurden (MWD [95%-KI]: $-0,19\%$ [$-0,34; -0,03$]). Studienabbrüche aufgrund einer unzureichenden glykämischen Kontrolle traten signifikant häufiger in der Metformin+Sulfonylharnstoff-Gruppe auf (OR [95%-KI]: $0,08$ [$0,01; 0,59$] nach 52 Wochen und $0,65$ [$0,44; 0,97$] nach 104 Wochen).

In der direkten Vergleichsstudie von Metformin+Dapagliflozin mit Metformin+Sulfonylharnstoff konnte sowohl nach 52 als auch nach 104 Wochen gezeigt werden, dass bei signifikant weniger Patienten, die mit Metformin+Dapagliflozin behandelt wurden, patientenberichtete Hypoglykämien, d. h. vom Patienten mittels Tagebuch dokumentierte Hypoglykämien, auftraten als in der Gruppe der mit Metformin+Sulfonylharnstoff behandelten Patienten (OR [95%-KI], $0,05$ [$0,03; 0,09$]) nach 52 Wochen / OR [95%-KI], $0,05$ [$0,03; 0,09$]; ARR: $44,2\%$ nach 104 Wochen). Von diesen hatten 3 ($0,8\%$) Patienten, die mit Metformin+Sulfonylharnstoff behandelt wurden, eine schwere Hypoglykämie. Bei den mit Metformin+Dapagliflozin behandelten Patienten trat dieses Ereignis nicht auf. Auch Studienabbrüche aufgrund einer Hypoglykämie traten bei Patienten, die mit Metformin+Dapagliflozin behandelt wurden, nicht auf, wohingegen 6 ($1,7\%$) Patienten, die mit der zweckmäßigen Vergleichstherapie behandelt wurden, die Studie nach 104 Wochen abbrachen.

Die adäquate Langzeitblutzuckerkontrolle (über zwei Jahre) bei gleichzeitiger Hypoglykämievermeidung ist eine bedeutsame Vermeidung von Nebenwirkungen durch die Behandlung mit Metformin+Dapagliflozin, sodass sich ein Hinweis auf einen beträchtlichen Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie (Metformin+Sulfonylharnstoff) ergibt.

Der Einfluss von Hypoglykämien auf die gesundheitsbezogene Lebensqualität von Typ-2-Diabetikern wurde in mehreren Studien untersucht (Davis et al., 2005; Lundkvist et al., 2005; Vexiau et al., 2008; Pettersson et al., 2011). Es konnte gezeigt werden, dass das Auftreten sowie die Schwere einer Hypoglykämie die gesundheitsbezogene Lebensqualität mindert (Davis et al., 2005; Vexiau et al., 2008; Barnett et al., 2010; Pettersson et al., 2011). Durch die ständige Angst eine Hypoglykämie zu erleiden, kann es zu Beeinträchtigungen z. B. beim

Autofahren, bei der Freizeitgestaltung und damit im Sozialleben kommen (Barnett et al., 2010). Die von den Patienten berichtete Angst vor einer hypoglykämischen Episode kann so groß sein, dass sie zu verminderter Therapieadhärenz führen kann (siehe Modul 3B, Abschnitt 3.2.2). Dazu kommt, dass die Patienten durch häufigeres Essen versuchen, den Hypoglykämien entgegenzuwirken, was wiederum zur Gewichtszunahme führt. Durch die oben erwähnte bedeutsame Vermeidung der Hypoglykämien könnte damit wiederum die Lebensqualität der Patienten verbessert werden.

Betrachtung der Gewichtsveränderung

Adipositas (Fettleibigkeit), im Besonderen die viszerale/abdominelle Adipositas, ist der wichtigste Risikofaktor zur Entwicklung eines Diabetes mellitus Typ 2 (Freemantle et al., 2008). Adipositas führt zusammen mit der Bewegungsarmut und Dyslipoproteinämie zu einer Insulinresistenz (unzureichende Insulinwirksamkeit), welche essentiell zur Entstehung eines Diabetes mellitus Typ 2 beiträgt. Eine Verminderung des Körpergewichts stellt somit zusammen mit einer Ernährungsumstellung und vermehrter körperlicher Aktivität die Grundlage einer Diabetestherapie dar und spielt für die Patienten eine ähnlich wichtige Rolle wie die Blutzuckereinstellung (Porzolt et al., 2010). Einige Therapieoptionen haben diesbezüglich günstige Wirkungen, wie die OAD Metformin oder DPP-4-Inhibitoren. Einige andere Wirkstoffklassen sind ungünstig, wie etwa Sulfonylharnstoffe, weshalb auch in diesem Punkt ein klarer therapeutischer Bedarf besteht. Das metabolische Syndrom¹⁴, dessen wichtigster Einzelfaktor die Adipositas darstellt, ist weiterhin wesentlich für das hohe kardiovaskuläre bzw. Makroangiopathierisiko verantwortlich (Perk et al., 2012), welches wiederum die wichtigste Ursache für die Morbidität und Sterblichkeit bei Typ-2-Diabetikern darstellt (AkdÄ, 2009). Insbesondere der Anteil der Körperfettmasse korreliert positiv mit den wichtigsten kardiovaskulären Risikofaktoren und es scheint, dass eine Verringerung der Fettmasse vor diesen Risikofaktoren schützen kann (Vega et al., 2006).

Bezogen auf die Gewichtsveränderung zeigte sich in der direkten Vergleichsstudie ein signifikanter Unterschied zwischen den Behandlungsgruppen Metformin+Dapagliflozin gegen Metformin+Sulfonylharnstoff (MWD [95%-KI]: -5,02 kg [-5,57;-4,48] nach 52 Wochen und -5,38 kg [-6,07;-4,68] nach 104 Wochen). Die mit Metformin+Dapagliflozin behandelten Patienten nahmen im Mittel um 3,49 kg / 3,82 kg ab, während die mit Metformin+Sulfonylharnstoff behandelten Patienten an Gewicht zunahmen (1,54 kg / 1,56 kg). Diese mittlere Gewichts Differenz von gut 5 kg ist bei einem mittleren Körpergewicht der Studienpopulation von etwa 88 kg nicht nur statistisch signifikant, sondern auch klinisch relevant. Eine Gewichtsreduktion von 5-10% wird bei einem BMI von 25-35 kg/m² empfohlen (siehe Modul 3B, Abschnitt 3.2.2). Der Anteil der Patienten, die mehr als 5% an Gewicht verloren hatten, war in der Metformin+Dapagliflozin-Gruppe signifikant höher. Während nach 52 Wochen nur 2,6% der Patienten aus der

¹⁴ Zusammenschluss mehrerer Merkmale bzw. Risikofaktoren wie abdominale Adipositas (Bauchumfang: Männer > 94 cm, Frauen > 80 cm), Insulinresistenz, Hyperinsulinämie, gestörte Glukosetoleranz, Dyslipoproteinämie, Albuminurie, Hypertonie, die zur Entstehung des klinischen Krankheitsbildes beitragen (AkdÄ 2009).

Metformin+Sulfonylharnstoff-Gruppe mindestens 5% abgenommen hatten, lag der Anteil an Patienten in der Metformin+Dapagliflozin-Gruppe bei 36,6% (OR [95%-KI]: 22,01 [10,93;44,33]). Nach 104 Wochen hatten 2,6% der Patienten aus der Metformin+Sulfonylharnstoff-Gruppe und 25,5 % in der Metformin+Dapagliflozin-Gruppe mindestens 5% abgenommen (OR [95%-KI]:13,10 [6,45;26,60]). Dies bedeutet, dass Patienten, die mit Metformin+Dapagliflozin behandelt wurden, eine etwa 22-fach höhere Chance auf eine Gewichtsreduktion von >5% hatten.

Die gewichtsreduzierende Eigenschaft von Metformin+Dapagliflozin wird auch in einer weiteren von Bristol-Myers Squibb und AstraZeneca durchgeführten Studie gezeigt, die speziell darauf ausgelegt war, den Einfluss auf das Gewicht zu untersuchen (Bolinder et al., 2012). Die Ergebnisse zeigen, dass durch die Behandlung mit Metformin+Dapagliflozin nicht nur das Körpergewicht abnahm, sondern auch der Taillenumfang reduziert werden konnte. Es zeigte sich, dass der Gewichtsverlust hauptsächlich auf einer Reduktion der Körperfettmasse, des viszeralen Fettgewebes und des subkutanen Fettgewebes beruht. Der Verlust an Fettmasse war in dieser Studie im Vergleich zu Placebo signifikant und klinisch relevant.

Die anhaltende Gewichtsreduktion unter Metformin+Dapagliflozin stellt eine bedeutsame Vermeidung von Nebenwirkungen dar und trägt zur erfolgreichen Diabetestherapie bei. Vor allem die Tatsache, dass größtenteils die Körperfettmasse reduziert wird, ist ein wichtiger Vorteil dieser neuen Behandlungsoption. Dapagliflozin hat durch seinen einzigartigen Wirkmechanismus und den dadurch bedingten ständigen Verlust von Kalorien über den Urin, das Potenzial einen anhaltenden Gewichtsverlust, bedingt durch die Abnahme der Körperfettmasse zu erreichen. Durch das bedeutsam höhere Auftreten relevanter Gewichtsreduktionen (>5%) ergibt sich für Metformin+Dapagliflozin ein Hinweis für einen beträchtlichen Zusatznutzen.

Betrachtung der Therapiezufriedenheit

Ein patientenrelevantes Ziel einer Diabetestherapie ist eine möglichst geringe Beeinträchtigung der Lebensqualität durch die Erkrankung und Folgeerkrankungen, aber auch durch die Therapie selbst zu erreichen. Viele Diabetiker leiden unter einer eingeschränkten Lebensqualität, welche unter anderem durch körperliche Beschwerden, psychische Belastung durch die Therapie und Therapieüberwachung und durch das Bewusstsein des Morbiditäts- und Mortalitätsrisikos sowie durch soziale Diskriminierung bedingt sein kann (AkdÄ, 2009).

Die mittleren Werte der Therapiezufriedenheit basierend auf den DTSQc-Fragebögen waren in beiden Behandlungsgruppen hoch (Metformin+Dapagliflozin: 14,1 (SE: 0,317) und Metformin+Sulfonylharnstoff: 13,6 (SE: 0,312)). Das zeigt eine Verbesserung der Therapiezufriedenheit in beiden Behandlungsgruppen. Patienten, die mit Metformin+Dapagliflozin behandelt wurden, erreichten dabei einen etwas höheren Wert. Der Unterschied zwischen den Behandlungsgruppen war nicht signifikant.

Da jedoch bereits die Ergebnisse aus den DTSQs-Fragebögen hoch waren (Ausgangswert), war eine signifikante Verbesserung nicht zu erwarten.

Damit ergibt sich kein Hinweis auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen in Bezug auf die Therapiezufriedenheit, gemessen mit dem DTSQc-Fragebogen.

Für Dapagliflozin gilt es zu berücksichtigen, dass seine insulinunabhängige Wirkungsweise und seine einmal tägliche Gabe (unabhängig von den Mahlzeiten) zu einer geringeren Komplexität der Therapie beitragen. Entsprechend sind einfach zu handhabende Therapieregimes wie z. B. Dapagliflozin - bei vergleichbarer Wirksamkeit und Sicherheit unter Studienbedingungen - unter Alltagsbedingungen vorteilhaft, weil damit auch die Therapie-Adhärenz verbessert werden kann (Saini et al., 2009). Die Anwenderfreundlichkeit von Dapagliflozin ergibt sich zudem daraus, dass die Dosis im Gegensatz zu den Sulfonylharnstoffen normalerweise nicht fortlaufend während der Behandlung angepasst werden muss (Adler et al., 2000). Für die Dosis-Titrierung von Sulfonylharnstoffen sind häufigere Blut- und Harnzuckermessungen sowie Arztbesuche notwendig, die bei einer Therapie mit Dapagliflozin nicht notwendig sind. Darüber hinaus leiden die meisten Typ-2-Diabetiker unter vielen Komorbiditäten und müssen deshalb viele Begleitmedikamente einnehmen. Für die meisten in der Diabetestherapie gängigen Begleitmedikamente konnten keine klinisch relevanten Wechselwirkungen mit Dapagliflozin nachgewiesen werden, wodurch Dapagliflozin mit einer Vielfalt von Medikamenten kombiniert werden kann (Bristol-Myers Squibb und AstraZeneca GmbH, 2010).

Einfluss auf den Blutdruck

Wie die Adipositas, so ist auch die essentielle Hypertonie mit einer Insulinresistenz assoziiert und begünstigt die Entwicklung der Mikro- und Makroangiopathie (AkdÄ, 2009). Die Senkung eines erhöhten Blutdrucks kann die kardio- und zerebrovaskuläre Morbidität und Mortalität innerhalb weniger Jahre reduzieren (IQWiG, 2011a). Darüber hinaus existiert ein linearer Zusammenhang zwischen der Höhe des Blutdruckes und dem Auftreten kardialer sowie zerebrovaskulärer Ereignisse (Mancia, 2007). Weiterhin bestehen direkte, signifikante Beziehungen zwischen der Höhe des Blutdrucks und der Inzidenz und Progression von Nephropathie (Adler et al., 2000) und Retinopathie (UK Prospective Diabetes Study Group, 1998a; Adler et al., 2000). Vor allem ein hoher systolischer Blutdruck ist mit der Entwicklung einer Mikroalbuminurie und einer Proteinurie assoziiert. Ein begleitender Hypertonus ist in >80% der Fälle bei Typ-2-Diabetes in Deutschland zu erwarten (Bramlage et al., 2010). Die Therapie einer Hypertonie wird daher in den medizinischen Leitlinien empfohlen (AkdÄ, 2009; Matthaer et al., 2009).

Patienten mit einer Metformin+Dapagliflozin-Behandlung hatten nach 52 Wochen eine höhere Blutdrucksenkung (systolisch: -4,4 mmHg / diastolisch: -1,6 mmHg) als Patienten mit einer Metformin+Sulfonylharnstoff-Behandlung (systolisch: +1,0 mmHg / diastolisch: -0,5 mmHg). Der Unterschied zwischen den beiden Behandlungsgruppen war für den systolischen Blutdruck statistisch signifikant (MWD [95%-KI]: -5,3 mmHg [-7,1;-3,6]) und hielt auch nach 104 Wochen an (MWD [95%-KI]: -4,46 mmHg; [-6,67;-2,26]). Der Unterschied in der Reduktion des diastolischen Blutdrucks mit Metformin+Dapagliflozin

gegenüber Metformin+Sulfonylharnstoff wies ebenfalls diesen Trend auf, war aber im Ausmaß geringer und erreichte keine statistische Signifikanz.

Vor dem Hintergrund, dass zum Erreichen des Zielblutdruckes oft eine Kombination von zwei oder mehr Antihypertensiva benötigt werden (Deutsche Hochdruck Liga e.V., 2011), kann die durch Metformin+Dapagliflozin erzielte Senkung des Blutdrucks und hier im Besonderen des systolischen Blutdrucks einen wertvollen Beitrag leisten. Darüber hinaus kann die erzielte Blutdrucksenkung zur Prävention einer Mikroangiopathie und von kardiovaskulären Ereignissen beitragen. Aufgrund der unter Metformin+Dapagliflozin erzielten stärkeren Senkung des systolischen Blutdrucks wird das Auftreten von nicht schwerwiegenden Symptome verringert, sodass sich ein Hinweis auf einen geringen Zusatznutzen ergibt.

Betrachtung der unerwünschten Ereignisse und der kardialen unerwünschten Ereignisse

Die Gesamtrate der Patienten, die ein unerwünschtes Ereignis, ein schwerwiegendes unerwünschtes Ereignis oder einen Studienabbruch aufgrund eines unerwünschten Ereignisses hatten, war in beiden Gruppen nach 52 sowie nach 104 Wochen ähnlich.

Signifikante Unterschiede zugunsten Metformin+Sulfonylharnstoff im Vergleich zu Metformin+Dapagliflozin ergaben sich bei der Betrachtung der Gesamtraten der Patienten mit Anzeichen für eine Genitalinfektion (OR [95%-KI]: 5,92 [2,83;12,37]) oder Anzeichen für eine Harnwegsinfektion (OR [95%-KI]: 2,02 [1,16;3,50]) sowohl nach 52 Wochen als auch nach 104 Wochen (OR [95%-KI], 6,67 [3,32;13,37] und 1,76 [1,10;2,81]).

Die aufgetretenen Anzeichen für Genital- und Harnwegsinfektionen waren allerdings meist nur mild oder moderat in ihrer Intensität und mit den gängigen Behandlungsmethoden einfach zu behandeln. Das bestätigen auch die Ergebnisse einer Untersuchung zur Häufigkeit von Urogenitalinfektionen bei Typ-2-Diabetikern. Mittels eines Fragebogens, der an 1194 diabetologisch tätige Hausärzte im gesamten Bundesgebiet verschickt wurde, wurde dabei nach der Prävalenz, Geschlechterverteilung sowie Schweregrad und Verlauf von Urogenitalinfektionen gefragt. Die Auswertung zeigte, dass Urogenitalinfektionen eine häufige Komplikation des Diabetes mellitus Typ 2 sind und die Mehrzahl der Infektionen in der Regel einfach zu behandeln sind und nicht zu schweren Verläufen führt (Weissenbacher E et al., 2012). Studienabbrüche aufgrund von Anzeichen für Genital- und Harnwegsinfektionen traten sehr selten auf. Da sich die Gesamtrate der unerwünschten Ereignisse in den beiden Behandlungsgruppen nicht voneinander unterschied, ist nicht von einem Schaden auszugehen, denn Patienten unter einer Behandlung mit Metformin+Sulfonylharnstoff erleiden andere unerwünschte Ereignisse häufiger als unter Metformin+Dapagliflozin. Hier sind vor allem Schwindel (9,3% der mit Metformin+Sulfonylharnstoff behandelten Patienten im Vergleich zu 5,2% der mit Metformin+Dapagliflozin behandelten Patienten), vermehrtes Schwitzen (7,8% der Patienten in der Metformin+Sulfonylharnstoff-Gruppe im Vergleich zu 1,5% derer in der Metformin+Dapagliflozin-Gruppe) sowie Zittern (7,8% der mit

Metformin+Sulfonylharnstoff behandelten Patienten gegenüber 0,5% der mit Metformin+Dapagliflozin behandelten Patienten) zu nennen (AstraZeneca, 2010b).

Weiterhin wurde das Auftreten von kardialen Ereignissen betrachtet. Die Anzahl der im Rahmen der Sicherheitsevaluation aufgetretenen kardialen Ereignisse war in beiden Gruppen sehr gering und ließ keinen signifikanten Unterschied zwischen den beiden Gruppen erkennen. Darüber hinaus plant Bristol-Myers Squibb / AstraZeneca nach der Zulassung eine große Langzeitstudie, um den kardiovaskulären Nutzen sowie das kardiovaskuläre Sicherheitsprofil von Dapagliflozin zu untersuchen.

Patienten, bei denen die zweckmäßige Vergleichstherapie nicht angewendet werden kann

Wie bereits im Rahmen des G-BA-Gespräches angemerkt, gibt es Patienten, bei denen die zweckmäßige Vergleichstherapie nicht angewendet werden kann (Hein, 2012). Das ist darin begründet, dass es unter der Behandlung mit Sulfonylharnstoffen vermehrt zum Auftreten von Hypoglykämien und zu Gewichtszunahmen kommt (UK Prospective Diabetes Study Group, 1998b; Patel et al., 2008; AkdÄ, 2009). Vor allem das häufige Auftreten von Hypoglykämien kann die Behandlung mit Sulfonylharnstoffen für bestimmte Patienten (z. B. Bus- oder LKW-Fahrer, Patienten, die in großer Höhe oder mit gefährlichen Maschinen arbeiten, allein lebende Senioren, Patienten mit kardiovaskulärer Vorbelastung) ungeeignet machen (National Collaborating Centre for Chronic Conditions, 2008; AkdÄ, 2009; NICE, 2009). Auch eine therapiebedingte zusätzliche Gewichtszunahme stellt eine Belastung für die meist adipösen Typ-2-Diabetiker dar, da diese einer erfolgreichen Diabetestherapie, welche eine Gewichtsreduktion vorsieht, im Wege steht.

Wie in Abschnitt 4.2.2 erläutert sieht Bristol-Myers Squibb / AstraZeneca bei diesen Patienten die Behandlung mit Metformin+DPP-4-Inhibitor als zweckmäßige Vergleichstherapie an. Die Ergebnisse der MTC zeigen gegenüber Metformin+DPP-4-Inhibitoren, dass Metformin+Dapagliflozin sowohl nach 24 als auch nach 52 Wochen eine signifikant stärkere Gewichtsreduktion bewirkt (MWD [95%-Kredibilitätsintervall]: -2,37 kg [-2,98;-1,77] nach 24 Wochen / -2,74 kg [-5,35;-0,10] nach 52 Wochen). Während die Behandlung mit Metformin+Dapagliflozin zu einer Gewichtsreduktion von etwa 3,5 kg führt, kann die Behandlung mit Metformin+DPP-4-Inhibitoren das Gewicht nur um 0,091 kg bis maximal 1,5 kg senken. Bezüglich den anderen betrachteten Endpunkten (Blutzuckerkontrolle [HbA1c]), Hypoglykämien, Blutdrucksenkung, unerwünschte Ereignisse) gab es keinen Anhaltspunkt auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen. Durch die signifikant stärkere Gewichtsreduktion ergibt sich ein Anhaltspunkt auf einen geringen Zusatznutzen gegenüber Metformin+DPP-4-Inhibitoren.

Gesamtbetrachtung des Zusatznutzens

Aus den Ergebnissen dieser Nutzenbewertung zeigt sich mit der Evidenzstufe 1b und einer hohen Ergebnissicherheit (niedriges Verzerrungspotenzial, hohe methodische Qualität), dass die Therapie mit Dapagliflozin in Kombination mit Metformin einige klinisch relevante Vorteile im Vergleich zu den Sulfonylharnstoffen in Kombination mit Metformin bietet.

Der Wirkmechanismus von Dapagliflozin, der unabhängig von Insulin sowie von der Funktion der Pankreaszellen ist, senkt konzentrationsabhängig den Blutzucker, ohne von einer Insulinresistenz beeinträchtigt zu werden. Dieser neuartige Wirkmechanismus bietet gegenüber den Sulfonylharnstoffen mehrere potentielle Vorteile für den Patienten, die durch die hier dargestellten Ergebnisse bestätigt werden:

- Signifikante Vermeidung von Hypoglykämien bei effektiver Blutzuckerkontrolle
- Langanhaltende HbA1c-Senkung
- Signifikante und klinisch relevante Gewichtsreduktion (v. a. Reduktion von Körperfettmasse, viszeralem Fettgewebe und subkutanem Fettgewebe)
- Signifikant stärkere Senkung des systolischen Blutdruckes

Damit weist Dapagliflozin ein Wirkprofil auf, welches das kardiovaskuläre Risiko senken sollte. Wie im EPAR beschrieben, zeigt sich in einer Meta-Analyse kardiovaskulärer Ereignisse (kombinierter Endpunkt – kardiovaskulärer Tod, Herzinfarkt oder Schlaganfall) aller Phase IIb und III Studien, welche verblindet unabhängig adjudiziert wurden, ein positiver Trend für Dapagliflozin. Eine große Phase III Outcome-Studie ist geplant und sollte in Kürze beginnen.

Die Gesamtanzahl unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse, unerwünschter Ereignisse, welche zum Abbruch der Therapie führten und kardialer Ereignisse unterschied sich in den beiden Behandlungsarmen nicht. Die unter Metformin+Dapagliflozin vermehrt auftretenden Anzeichen für Genital- und Harnwegsinfektionen waren größtenteils mild und moderat und mit den gängigen Behandlungsmethoden einfach zu behandeln.

Die Behandlung mit Metformin+Dapagliflozin stellt damit eine wirksame, langanhaltende (über 2 Jahre) und im Allgemeinen gut verträgliche blutzuckersenkende Therapie mit günstigem Nutzen-Risiko-Verhältnis und geringem Risiko für die überaus belastenden und potenziell lebensgefährlichen Hypoglykämien dar. Die im Gegensatz zu Metformin+Sulfonylharnstoff auch nach 2 Jahren noch anhaltende effektive Blutzuckerkontrolle lässt vermuten, dass Metformin+Dapagliflozin unabhängig von der Diabetesdauer und damit unabhängig vom fortschreitenden Funktionsverlust der pankreatischen Inselzellen wirken kann.

Aus der Gesamtbetrachtung der Ergebnisse ergibt sich damit ein Hinweis auf einen beträchtlichen Zusatznutzen von Dapagliflozin in Kombination mit Metformin gegenüber Metformin+Sulfonylharnstoff.

Für Patienten, für die eine Behandlung mit Sulfonylharnstoffen nicht in Frage kommt, ist die Behandlung mit Metformin+DPP-4-Inhibitoren als zweckmäßige Vergleichstherapie anzusehen. Die Ergebnisse der MTC zeigten für Metformin+Dapagliflozin eine signifikant stärkere Gewichtsreduktion, wodurch ein Anhaltspunkt für einen geringen Zusatznutzen besteht.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-142: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit Diabetes mellitus Typ 2, die unter einer Metformin-Monotherapie zusammen mit einer Diät und Bewegung einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen, unter Ausschluss von Patienten mit einer eGFR <60 mL/min/1,73 m ² bzw. Kreatinin-clearance <60 mL/min und von Patienten, die ≥75 Jahre sind und Patienten, die mit Schleifendiuretika behandelt werden.	Beträchtlicher Zusatznutzen
Erwachsene Patienten mit Diabetes mellitus Typ 2, die unter einer Metformin-Monotherapie zusammen mit einer Diät und Bewegung einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen, unter Ausschluss von Patienten mit einer eGFR <60 mL/min/1,73 m ² bzw. Kreatinin-clearance <60 mL/min und von Patienten, die ≥75 Jahre sind und Patienten, die mit Schleifendiuretika behandelt werden und für die Sulfonylharnstoffe nicht geeignet sind , und bei denen die Anwendung von Insulin noch nicht angezeigt ist (Teilpopulation).*	Geringer Zusatznutzen

* Gegenüber der alternativen Vergleichstherapie: Metformin+DPP-4-Inhibitor

4.4.4 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens

Für Arzneimittel, die zur Behandlung eines seltenen Leidens zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt.

Für diese Arzneimittel sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, im vorliegenden Abschnitt einzureichen. Dabei soll zunächst das Ausmaß des Zusatznutzens (in der Definition gemäß AM-NutzenV) begründet beschrieben werden. Unter Berücksichtigung des Ausmaßes des Zusatznutzens sollen

Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, benannt werden. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

Beschreiben Sie begründet das Ausmaß des Zusatznutzens unter Verwendung der Kategorisierung gemäß AM-NutzenV. Berücksichtigen Sie bei den Aussagen zum Zusatznutzen nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen. Verweisen Sie, wo notwendig, auf weitere Abschnitte von Modul 4.

Nicht zutreffend.

Geben Sie auf Basis des Ausmaßes des Zusatznutzens in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-143: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Anders als für den Vergleich von Metformin+Dapagliflozin mit Metformin+Sulfonylharnstoff liegt für den Vergleich mit Metformin+DPP-4-Inhibitoren keine direkt vergleichende Studie vor. Aus diesem Grund wurde ein indirekter Vergleich vorgenommen. Hierfür wurde eine MTC durchgeführt, welche eine wichtige Weiterentwicklung der üblichen meta-analytischen Verfahren darstellt (IQWiG, 2011b). Die in die MTC eingeschlossenen Studien waren ausschließlich RCT, da diese mit der geringsten Ergebnisunsicherheit behaftet sind. Darüber hinaus waren die Studien größtenteils doppelt verblindet und die Charakteristika der Studienpopulationen weitgehend ähnlich.

Die durchgeführte Netzwerk-Meta-Analyse erfüllt damit den momentanen Stand der Wissenschaft hinsichtlich der Anforderungen an eine solche Analyse. Der Netzwerkvergleich ist daher aussagekräftig.

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nichtrandomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Es wurden keine nicht randomisierten vergleichenden Studien herangezogen.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹⁵, Molenberghs 2010¹⁶). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁷) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und

¹⁵ Burzykowski T (Ed.: *The evaluation of surrogate endpoints*. New York: Springer; 2005.

¹⁶ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: *A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials*. *Stat Methods Med Res* 2010; 19(3): 205-236.

¹⁷ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006¹⁸) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Für die aktuelle Nutzenbewertung ist der patientenrelevante Endpunkt kardiovaskuläre Morbidität und Mortalität nicht direkt verfügbar. Die für die Fragestellung relevanten Studien sollten gemäß Einschlusskriterien eine Beobachtungsdauer von mindestens 24 Wochen aufweisen. Diese Zeit ist angemessen, um patientenrelevante Endpunkte, wie z. B. Hypoglykämien zu beurteilen. Um valide Aussagen bezüglich des patientenrelevanten Endpunktes kardiovaskuläre Morbidität und Mortalität zu erlauben, erscheint eine Beobachtungsdauer von 24 bzw. 52 Wochen jedoch zu kurz. Daher wird die Senkung des Blutdrucks als Surrogat für das kardiovaskuläre Risiko betrachtet. Die Validität dieses Surrogats wird im Folgenden begründet. Grundsätzlich ist anzumerken, dass die Senkung des Blutdrucks kein primäres Therapieziel der Behandlung mit Dapagliflozin darstellt. Die Blutdrucksenkung und die daraus abgeleiteten Effekte auf das kardiale Risiko der Patienten sind als Nebeneffekte der Behandlung mit Dapagliflozin zu sehen.

Validität des Surrogats im Indikationsgebiet Diabetes:

Hypertonie gilt als einer der stärksten prognostischen Faktoren für zerebro- und kardiovaskuläre Morbidität und Mortalität. Zwischen der Höhe des Blutdrucks und dem

¹⁸ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Stat Med* 2006; 25(2): 183-203.

Auftreten kardialer sowie zerebrovaskulärer Ereignisse existiert eine lineare Relation (Mancia, 2007). Eine Korrelation existiert auch zwischen Bluthochdruck und mikrovaskulären Komplikationen, der Hypertonus gilt als Progressionsfaktor für Retinopathie und Nephropathie (Adler et al., 2000; Standl et al., 2000).

Bluthochdruck und Diabetes mellitus Typ 2 sind häufig assoziiert, die Prävalenz des gemeinsamen Auftretens der beiden Erkrankungen wird in Deutschland mit bis zu 80% angegeben (Standl et al., 2000; Wittchen et al., 2007; Bramlage et al., 2010). Hypertensive Diabetiker haben im Vergleich zu normotensiven Nicht-Diabetikern ein 4-fach erhöhtes Risiko für kardiovaskuläre Morbidität und Mortalität (Standl et al., 2000) und werden deshalb Hoch-Risikogruppen zugeordnet, bei denen mit einer 20-30%igen Wahrscheinlichkeit nach 10 Jahren mit einem kardiovaskulären Ereignis zu rechnen ist. Das Risiko ist bei Patienten mit zusätzlicher Nephropathie und insbesondere Mikroalbuminurie, welche als Zeichen einer Gefäßschädigung angesehen wird, dabei nochmals erhöht (Bundesärztekammer et al., 2011).

Die Makroangiopathie ist nach Anlage 1 der Verordnung zur Änderung der Risikostruktur-Ausgleichsverordnung (RSA-ÄndV) in Form der koronaren Herzkrankheit das Hauptproblem der Patientinnen / Patienten mit Diabetes mellitus Typ 2. Die Senkung eines erhöhten Blutdrucks kann zur Reduktion der kardio- und zerebrovaskulären Morbidität und Mortalität innerhalb weniger Jahre beitragen. Das *Disease Management Programm* (DMP) empfiehlt daher eine geeignete Risikoabschätzung hinsichtlich makroangiopathischer Komplikationen (IQWiG, 2011a).

Das IQWiG stellt in der Zusammenfassung der deutschen, europäischen und außereuropäischen Leitlinien die Bedeutung der Risikoreduktion speziell in Bezug auf den Bluthochdruck bei Patienten mit Diabetes dar. Die Leitlinien heben die Wichtigkeit der Erfassung, routinemäßigen Kontrolle des Blutdrucks als auch die langfristige Blutdrucksenkung bei Diabetes Patienten hervor. Das IQWiG sieht bzgl. der Zielwerte für die Blutdrucksenkung Klärungsbedarf, stellt jedoch die allgemeine Bedeutung der Blutdruckkontrolle nicht in Frage.

4.6 Liste der eingeschlossenen Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Datenquellen auf.

Studie für den direkten Vergleich

Studie	Titel	Im Dossier verwendete Datenquellen
D1690C00004	<i>A 52-Week International, Multi-centre, Randomized, Parallel-group, Double-blind, Active-controlled, Phase III study with a 52-Week Extension Period to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Metformin compared with Sulphonylurea in Combination with Metformin in Adult Patients with Type 2 Diabetes who have Inadequate Glycemic Control on Metformin Therapy Alone</i>	(AstraZeneca, 2010b) (AstraZeneca, 2011b)

Studien für die MTC:

Studie	Titel	Im Dossier verwendete Datenquellen
MB102014	<i>Effect of dapagliflozin in patients with type 2 diabetes who have inadequate glycaemic control with metformin: a randomised, double-blind, placebo-controlled trial.</i>	(Bristol-Myers Squibb Company, 2009) (Bristol-Myers Squibb Company, 2010a)
D1690C00012	<i>A 24-week, multi-centre, international, double-blind, randomized, parallel-group, placebo-controlled, phase III study with a 78-week extension period to evaluate the effect of dapagliflozin in combination with metformin on body weight in subjects with type 2 diabetes mellitus who have inadequate glycaemic control on metformin alone. Report for the 24-week short-term treatment period</i>	(AstraZeneca, 2010a) (AstraZeneca, 2011a) (AstraZeneca, 2012)
Scott 2008	<i>Efficacy and safety of sitagliptin when added to ongoing metformin therapy in patients with type 2 diabetes</i>	(Scott et al., 2008)
Charbonnel 2006	<i>Efficacy and safety of the dipeptidyl peptidase-4 inhibitor sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes inadequately controlled with metformin alone</i>	(Charbonnel et al., 2006)
DeFronzo 2009	<i>The efficacy and safety of saxagliptin when added to metformin therapy in patients with inadequately controlled type 2 diabetes with metformin alone</i>	(DeFronzo et al., 2009)
Bosi 2007	<i>Effects of vildagliptin on glucose control over 24 weeks in patients with type 2 diabetes inadequately controlled with metformin</i>	(Bosi et al., 2007)
Taskinen 2011	<i>Safety and efficacy of linagliptin as add-on therapy to metformin in patients with type 2 diabetes: A randomized, double-blind, placebo-controlled study</i>	(Taskinen et al., 2011)
Bergenstal 2010a	<i>Once Weekly Taspoglutide, a Human GLP-1 Analog, Is Superior to Sitagliptin in Improving Glycemic Control and Weight Loss in Patients with Type 2 Diabetes (T2D): Results from the T-emerge 4 Trial</i>	(Bergenstal et al., 2010a)

Studie	Titel	Im Dossier verwendete Datenquellen
Raz 2008	<i>Efficacy and safety of sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes</i>	(Raz et al., 2008)
Bolli 2008	<i>Efficacy and tolerability of vildagliptin vs. pioglitazone when added to metformin: A 24-week, randomized, double-blind study</i>	(Bolli et al., 2008)
Bergental 2010b ¹	<i>Efficacy and safety of exenatide once weekly versus sitagliptin or pioglitazone as an adjunct to metformin for treatment of type 2 diabetes (DURATION-2): a randomised trial</i>	(Bergental et al., 2010b)
DeFronzo 2005	<i>Effects of exenatide (exendin-4) on glycemic control and weight over 30 weeks in metformin-treated patients with type 2 diabetes</i>	(DeFronzo et al., 2005)
Nauck 2009;	<i>Efficacy and safety comparison of liraglutide, glimepiride, and placebo, all in combination with metformin, in type 2 diabetes</i>	(Nauck et al., 2009)
Pratley 2010	<i>Liraglutide versus sitagliptin for patients with type 2 diabetes who did not have adequate glycaemic control with metformin: a 26-week, randomised, parallel-group, open-label trial</i>	(Pratley et al., 2010)
Kaku 2009	<i>Efficacy and safety of therapy with metformin plus pioglitazone in the treatment of patients with type 2 diabetes: A double-blind, placebo-controlled, clinical trial</i>	(Kaku, 2009)
Charpentier 2001	<i>Improved glycaemic control by addition of glimepiride to metformin monotherapy in Type 2 diabetic patients</i>	(Charpentier et al., 2001)
Papathanassiou 2009	<i>Pioglitazone vs glimepiride: differential effects on vascular endothelial function in patients with type 2 diabetes</i>	(Papathanassiou et al., 2009)
D1690C00004	<i>Dapagliflozin vs glipizide in patients with type 2 diabetes mellitus inadequately controlled on metformin: 52-week results of a double-blind, randomised, controlled trial</i>	(AstraZeneca, 2010b; Nauck et al., 2010; Nauck et al., 2011)
Matthews 2005	<i>Long-term therapy with addition of pioglitazone to metformin compared with the addition of gliclazide to metformin in patients with type 2 diabetes: a randomized, comparative study</i>	(Matthews et al., 2005)
Nauck 2007	<i>Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonylurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: A randomized, double-blind, non-inferiority trial</i>	(Nauck et al., 2007)
Göke 2010	<i>Saxagliptin is non-inferior to glipizide in patients with type 2 diabetes mellitus inadequately controlled on metformin alone: A 52-week randomised controlled trial</i>	(Göke et al., 2010)
Filozof 2010	<i>A comparison of efficacy and safety of vildagliptin and gliclazide in combination with metformin in patients with type 2 diabetes inadequately controlled with metformin alone: A 52-week, randomized study</i>	(Filozof und Gautier, 2010)
Salvadeo 2010	<i>Comparison between exenatide and glimepiride on metabolic control and on insulin resistance in type 2 diabetic patients with metformin therapy</i>	(Salvadeo et al., 2010)
Derosa 2010	<i>Exenatide Versus Glibenclamide in Patients with Diabetes</i>	(Derosa et al., 2010)

Studie	Titel	Im Dossier verwendete Datenquellen
Matthews 2010	<i>Vildagliptin add-on to metformin produces similar efficacy and reduced hypoglycaemic risk compared with glimepiride, with no weight gain: Results from a 2-year study</i>	(Matthews et al., 2010)
Yang, 2011*	<i>Efficacy and safety of saxagliptin added to metformin in Asian people with type 2 diabetes mellitus: a randomized controlled trial.</i>	(Yang et al., 2011)
Derosa, 2011a*	<i>Pioglitazone compared to glibenclamide on lipid profile and inflammation markers in type 2 diabetic patients during an oral fat load.</i>	(Derosa et al., 2011a)
Derosa, 2011b* ²	<i>Exenatide or glimepiride added to metformin on metabolic control and on insulin resistance in type 2 diabetic patients.</i>	(Derosa et al., 2011b)
Pan 2012*	<i>Efficacy and tolerability of vildagliptin as add-on therapy to metformin in Chinese patients with type 2 diabetes mellitus.</i>	(Pan et al., 2012)
Yang 2012*	<i>The addition of sitagliptin to ongoing metformin therapy significantly improves glycemic control in Chinese patients with type 2 diabetes</i>	(Yang et al., 2012)
Gallwitz 2012*	<i>2-year efficacy and safety of linagliptin compared with glimepiride in patients with type 2 diabetes inadequately controlled on metformin: a randomised, double-blind, non-inferiority trial</i>	(Gallwitz et al., 2012)
DeFronzo 2012*	<i>Efficacy and Tolerability of the DPP-4 Inhibitor Alogliptin Combined with Pioglitazone, in Metformin-Treated Patients with Type 2 Diabetes</i>	(DeFronzo et al., 2012)

¹ der Studienarm Exenatide+Metformin wurde in der Meta-Analyse nicht berücksichtigt, da die Dosierung (2 mg Exenatide [einmal wöchentlich] +Metformin 1,5 g) zum Zeitpunkt der Durchführung der Meta-Analyse nicht zugelassen war.

² entspricht der Publikation zum Konferenz-Abstract von Salvadeo, 2010

* Studien, die durch aktualisierte Literaturrecherche bzw. Registersuche gefunden wurden

4.7 Referenzliste

Benennen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie in Modul 4 angegeben haben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard).

- [1] Adler, A. I., Stratton, I. M., Neil, H. A., et al. 2000. Association of systolic blood pressure with macrovascular and microvascular complications of type 2 diabetes (UKPDS 36): prospective observational study. *BMJ*, 321, 412-9.
- [2] AkdÄ 2009. Empfehlungen zur antihyperglykämischen Therapie des Diabetes mellitus Typ 2. *Arzneiverordnung in der Praxis, Band 36, Sonderheft 1 (Therapieempfehlungen)*.
- [3] Alvarez Guisasola, F., Tofe Povedano, S., Krishnarajah, G., et al. 2008. Hypoglycaemic symptoms, treatment satisfaction, adherence and their associations

- with glycaemic goal in patients with type 2 diabetes mellitus: findings from the Real-Life Effectiveness and Care Patterns of Diabetes Management (RECAP-DM) Study. *Diabetes, obesity & metabolism*, 10 Suppl 1, 25-32.
- [4] AstraZeneca. 2010a. A 24-week, multi-centre, international, double-blind, randomized, parallel-group, placebo-controlled, phase III study with a 78-week extension period to evaluate the effect of dapagliflozin in combination with metformin on body weight in subjects with type 2 diabetes mellitus who have inadequate glycaemic control on metformin alone. Report for the 24-week short-term treatment period.
- [5] AstraZeneca. 2010b. A 52-Week International, Multi-centre, Randomized, Parallel-group, Double-blind, Active-controlled, Phase III study with a 52-Week Extension Period to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Metformin compared with Sulphonylurea in Combination with Metformin in Adult Patients with Type 2 Diabetes who have Inadequate Glycemic Control on Metformin Therapy Alone - Report for the 52-week short-term treatment period.
- [6] AstraZeneca. 2011a. A 24-week, Multi-centre, International, Double-blind, Randomized, Parallel-group, Placebo-controlled, Phase III Study with a 78-week Extension Period to Evaluate the Effect of Dapagliflozin in Combination with Metformin on Body Weight in Subjects with Type 2 Diabetes Mellitus Who Have Inadequate Glycaemic Control on Metformin Alone Report for the first 50 weeks of treatment
- [7] AstraZeneca. 2011b. A 52-Week International, Multi-centre, Randomized, Parallel-group, Double-blind, Active-controlled, Phase III study with a 156-Week Extension Period to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Metformin compared with Sulphonylurea in Combination with Metformin in Adult Patients with Type 2 Diabetes who have Inadequate Glycemic Control on Metformin Therapy Alone Report for the 52-week short-term treatment period plus the 52-week long-term extension period I
- [8] AstraZeneca. 2012. A 24-week, Multi-centre, International, Double-blind, Randomized, Parallel-group, Placebo-controlled, Phase III Study with a 78-week Extension Period to Evaluate the Effect of Dapagliflozin in Combination with Metformin on Body Weight in Subjects with Type 2 Diabetes Mellitus Who Have Inadequate Glycaemic Control on Metformin Alone Report for the 24-week short-term treatment period plus 26-week longterm extension period I and 52-week week long-term extension period II.
- [9] Bailey, C. J., Gross, J. L., Pieters, A., et al. 2010. Effect of dapagliflozin in patients with type 2 diabetes who have inadequate glycaemic control with metformin: a randomised, double-blind, placebo-controlled trial. *Lancet*, 375, 2223-33.
- [10] Barnett, A. H., Craddock, S., Fisher, M., et al. 2010. Key considerations around the risks and consequences of hypoglycaemia in people with type 2 diabetes. *International journal of clinical practice*, 64, 1121-9.
- [11] Bergenstal, R. M., Forti, A., Chiasson, J.-L., et al. 2010a. Once Weekly Taspoglutide, a Human GLP-1 Analog, Is Superior to Sitagliptin in Improving Glycemic Control and Weight Loss in Patients with Type 2 Diabetes (T2D): Results from the T-emerge 4 Trial.

- [12] Bergenstal, R. M., Wysham, C., Macconell, L., et al. 2010b. Efficacy and safety of exenatide once weekly versus sitagliptin or pioglitazone as an adjunct to metformin for treatment of type 2 diabetes (DURATION-2): a randomised trial. *Lancet*, 376, 431-9.
- [13] Bolinder, J., Ljunggren, O., Kullberg, J., et al. 2012. Effects of Dapagliflozin on Body Weight, Total Fat Mass, and Regional Adipose Tissue Distribution in Patients with Type 2 Diabetes Mellitus with Inadequate Glycemic Control on Metformin. *The Journal of clinical endocrinology and metabolism*.
- [14] Bolli, G., Dotta, F., Rochotte, E., et al. 2008. Efficacy and tolerability of vildagliptin vs. pioglitazone when added to metformin: a 24-week, randomized, double-blind study. *Diabetes, obesity & metabolism*, 10, 82-90.
- [15] Bosi, E., Camisasca, R. P., Collober, C., et al. 2007. Effects of vildagliptin on glucose control over 24 weeks in patients with type 2 diabetes inadequately controlled with metformin. *Diabetes care*, 30, 890-5.
- [16] Bradley C. 1994. The Diabetes Treatment Satisfaction Questionnaire: DTSQ. In: BRADLEY, C. (Hrsg.) *Handbook of psychology and diabetes: a guide to psychological measurement in diabetes research and practice*. Chur, Switzerland: Harwood Academic Publishers.
- [17] Bramlage, P., Binz, C., Gitt, A. K., et al. 2010. Diabetes treatment patterns and goal achievement in primary diabetes care (DiaRegis) - study protocol and patient characteristics at baseline. *Cardiovascular diabetology*, 9, 53.
- [18] Bristol-Myers Squibb Company. 2009. A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel Group, Phase 3 Trial to Evaluate the Safety and Efficacy of Dapagliflozin in Combination with Metformin in Subjects with Type 2 Diabetes who have Inadequate Glycemic Control on Metformin Alone.
- [19] Bristol-Myers Squibb Company. 2010a. A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel Group, Phase 3 Trial to Evaluate the Safety and Efficacy of Dapagliflozin in Combination with Metformin in Subjects with Type 2 Diabetes who have Inadequate Glycemic Control on Metformin Alone. 102-Week Short Term + Long Term Final Clinical Study Report MB102014.
- [20] Bristol-Myers Squibb Company. 2010b. Characterization of the Kinetics of Renal Glucose Reabsorption in Response to Dapagliflozin in Healthy Subjects and in Subjects with Type 2 Diabetes Mellitus.
- [21] Bristol-Myers Squibb und AstraZeneca GmbH. 2010. Clinical Overview: Dapagliflozin in the treatment of patients with type 2 diabetes.
- [22] Bristol-Myers Squibb und AstraZeneca GmbH. 2012a. Fachinformation Forxiga (Dapagliflozin).
- [23] Bristol-Myers Squibb und AstraZeneca GmbH 2012b. Post-hoc Analysen der einzelnen Endpunkte und Subgruppen für die Zielpopulation von Dapagliflozin (Studie D1690C00004).
- [24] Brooks, A. M. & Thacker, S. M. 2009. Dapagliflozin for the treatment of type 2 diabetes. *The Annals of pharmacotherapy*, 43, 1286-93.
- [25] Bundesärztekammer, KBV & AWMF. 2011. Nationale Versorgungsleitlinie: Nierenerkrankungen bei Diabetes im Erwachsenenalter – Version 1.3: Langfassung.

- [26] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV) & Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) 2012. Nationale VersorgungsLeitlinie Therapieplanung bei Typ-2-Diabetes. 1.0. Konsultationsfassung.
- [27] Charbonnel, B., Karasik, A., Liu, J., et al. 2006. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes inadequately controlled with metformin alone. *Diabetes care*, 29, 2638-43.
- [28] Charpentier, G., Fleury, F., Kabir, M., et al. 2001. Improved glycaemic control by addition of glimepiride to metformin monotherapy in type 2 diabetic patients. *Diabetic medicine : a journal of the British Diabetic Association*, 18, 828-34.
- [29] Clinicaltrials.gov. 2012a. *Efficacy and Safety of Dapagliflozin in Combination With Metformin in Type 2 Diabetes Patients* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT00660907> [Aufgerufen am 30.10.2012].
- [30] Clinicaltrials.gov. 2012b. *A Study of BMS-512148 (Dapagliflozin) in Patients With Type 2 Diabetes With Inadequately Controlled Hypertension on an ACEI or ARB and an Additional Antihypertensive Medication* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT01195662> [Aufgerufen am 30.10.2012].
- [31] Clinicaltrials.gov. 2012c. *Renal Mechanism of Action/Splay vs. TmG (MOA)* [Online]. Adresse: <http://clinicaltrials.gov/show/NCT00726505> [Aufgerufen am 12.11.2012].
- [32] Clinicaltrials.gov. 2012d. *Safety and Efficacy of Combination Saxagliptin & Dapagliflozin Added to Metformin to Treat Subjects With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/show/NCT01606007> [Aufgerufen am 6.11.2012].
- [33] Clinicaltrials.gov. 2012e. *Efficacy and Safety of B I1356 (Linagliptin) vs. Placebo Added to Metformin Background Therapy in Patients With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00601250> [Aufgerufen am 31.10.2012].
- [34] Clinicaltrials.gov. 2012f. *A Phase III Study of BMS-512148 (Dapagliflozin) in Asian Patients With Type 2 Diabetes Who Are Not Well Controlled With Diet and Exercise* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT01095653> [Aufgerufen am 15.11.2012].
- [35] Clinicaltrials.gov. 2012g. *Evaluation of the Effect on Glucose Control of AC2993 in Patients With Type 2 Diabetes Mellitus Treated With Metformin* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00039013> [Aufgerufen am 31.10.2012].
- [36] Clinicaltrials.gov. 2012h. *Efficacy and Safety in Patients With Type 2 Diabetes Mellitus and Cardiovascular Disease* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT01042977> [Aufgerufen am 30.10.2012].
- [37] Clinicaltrials.gov. 2012i. *A Study to Compare the Glycemic Effects, Safety, and Tolerability of Exenatide Once Weekly to Those of Sitagliptin and a Thiazolidinedione in Subjects With Type 2 Diabetes Treated With Metformin (DURATION - 2)* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00637273> [Aufgerufen am 31.10.2012].
- [38] clinicaltrials.gov. 2012j. *Safety and Efficacy of Saxagliptin in Triple Therapy to Treat Subjects With Type 2 Diabetes* [Online]. Adresse:

- <http://clinicaltrials.gov/ct2/show/NCT01619059?term=NCT01619059&rank=1>
[Aufgerufen am 12.11.2012].
- [39] Clinicaltrials.gov. 2012k. *Vildagliptin Compared to Pioglitazone in Combination With Metformin in Patients With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00237237> [Aufgerufen am 31.10.2012].
- [40] Clinicaltrials.gov. 2012l. *Sitagliptin Metformin Add-on Study in Patients With Type 2 Diabetes Mellitus* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00337610> [Aufgerufen am 31.10.2012].
- [41] Clinicaltrials.gov. 2012m. *Efficacy and Safety of Vildagliptin in Combination With Metformin in Patients With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00099892> [Aufgerufen am 31.10.2012].
- [42] Clinicaltrials.gov. 2012n. *52-week add-on to Metformin Comparison of Saxagliptin and Sulphonylurea, With a 52-week Extension Period* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00575588> [Aufgerufen am 31.10.2012].
- [43] Clinicaltrials.gov. 2012o. *Study Assessing Saxagliptin Treatment In Type 2 Diabetic Subjects Who Are Not Controlled With Metformin Alone* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00121667> [Aufgerufen am 31.10.2012].
- [44] Clinicaltrials.gov. 2012p. *To Compare the Effect of Liraglutide When Given Together With Metformin With the Effect of Metformin Given Alone and With the Effect of Glimpiride and Metformin Given Together (LEAD-2)* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00318461> [Aufgerufen am 31.10.2012].
- [45] Clinicaltrials.gov. 2012q. *Metformin Add-on Study in Patients With Type 2 Diabetes Mellitus (0431-020) (completed)* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00086515> [Aufgerufen am 31.10.2012].
- [46] Clinicaltrials.gov. 2012r. *Efficacy and Safety in Patients With Type 2 Diabetes Mellitus, Cardiovascular Disease and Hypertension* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT01031680> [Aufgerufen am 30.10.2012].
- [47] Clinicaltrials.gov. 2012s. *Safety/Efficacy of Sitagliptin in Patient w/ Type 2 Diabetes (0431-801)* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00541775> [Aufgerufen am 31.10.2012].
- [48] Clinicaltrials.gov. 2012t. *Efficacy of add-on Therapy With Vildagliptin in Chinese Patients With Type 2 Diabetes Inadequately Controlled With Prior Metformin Monotherapy* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00822211> [Aufgerufen am 30.10.2012].
- [49] Clinicaltrials.gov. 2012u. *Exenatide Versus Glimpiride in Patients With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00359762> [Aufgerufen am 7.11.2012].
- [50] Clinicaltrials.gov. 2012v. *An Investigational Drug Study in Patients With Type 2 Diabetes Mellitus* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00094770> [Aufgerufen am 31.10.2012].
- [51] Clinicaltrials.gov. 2012w. *A Phase III Study of BMS-512148 (Dapagliflozin) in Patients With Type 2 Diabetes Who Are Not Well Controlled on Metformin Alone* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00528879> [Aufgerufen am 30.10.2012].

- [52] Clinicaltrials.gov. 2012x. *Evaluation of the Effect of Dapagliflozin in Combination With Metformin on Body Weight in Subjects With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00855166> [Aufgerufen am 30.10.2012].
- [53] Clinicaltrials.gov. 2012y. *Vildagliptin Compared to Gliclazide in Combination With Metformin in Patients With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00102466> [Aufgerufen am 31.10.2012].
- [54] Clinicaltrials.gov. 2012z. *A Study of Taspoglutide Versus Sitagliptin for the Treatment of Patients With Type 2 Diabetes Mellitus Inadequately Controlled With Metformin* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00754988> [Aufgerufen am 30.10.2012].
- [55] Clinicaltrials.gov. 2012|. *The Effect of Liraglutide Compared to Sitagliptin, Both in Combination With Metformin on Glycaemic Control in Subjects With Type 2 Diabetes Mellitus* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00700817> [Aufgerufen am 30.10.2012].
- [56] Clinicaltrials.gov. 2012|. *A Study to Test the Safety and Efficacy of Adding Sitagliptin in Patients With Type 2 Diabetes Mellitus (MK0431-074)* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00813995> [Aufgerufen am 7.11.2012].
- [57] Clinicaltrials.gov. 2012}. *A Phase III Study of BMS-512148 (Dapagliflozin) in Asian Patients With Type 2 Diabetes Who Are Not Well Controlled on Metformin Alone* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT01095666> [Aufgerufen am 31.10.2012].
- [58] Clinicaltrials.gov. 2012~. *A Study of BMS-512148 (Dapagliflozin) in Patients With Type 2 Diabetes With Inadequately Controlled Hypertension on an Angiotensin-Converting Enzyme Inhibitor (ACEI) or Angiotensin Receptor Blocker (ARB)* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT01137474> [Aufgerufen am 30.10.2012].
- [59] Cooper, N. J., Sutton, A. J., Morris, D., et al. 2009. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat Med*, 28, 1861-81.
- [60] Davis, R. E., Morrissey, M., Peters, J. R., et al. 2005. Impact of hypoglycaemia on quality of life and productivity in type 1 and type 2 diabetes. *Current medical research and opinion*, 21, 1477-83.
- [61] DeFronzo, R. A., Ratner, R. E., Han, J., et al. 2005. Effects of exenatide (exendin-4) on glycemic control and weight over 30 weeks in metformin-treated patients with type 2 diabetes. *Diabetes care*, 28, 1092-100.
- [62] DeFronzo, R. A., Hissa, M. N., Garber, A. J., et al. 2009. The efficacy and safety of saxagliptin when added to metformin therapy in patients with inadequately controlled type 2 diabetes with metformin alone. *Diabetes care*, 32, 1649-55.
- [63] DeFronzo, R. A., Burant, C. F., Fleck, P., et al. 2012. Efficacy and tolerability of the DPP-4 inhibitor alogliptin combined with pioglitazone, in metformin-treated patients with type 2 diabetes. *The Journal of clinical endocrinology and metabolism*, 97, 1615-22.

- [64] Derosa, G., Maffioli, P., Salvadeo, S. A., et al. 2010. Exenatide versus glibenclamide in patients with diabetes. *Diabetes technology & therapeutics*, 12, 233-40.
- [65] Derosa, G., Cicero, A. F. G., Fogari, E., et al. 2011a. Pioglitazone compared to glibenclamide on lipid profile and inflammation markers in type 2 diabetic patients during oral fat load. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme*, 43, 7.
- [66] Derosa, G., Putignano, P., Bossi, A. C., et al. 2011b. Exenatide or glimepiride added to metformin on metabolic control and insulin resistance in type 2 diabetic patients. *European Journal of Pharmacology*, 666, 5.
- [67] Deutsche Hochdruck Liga e.V. 2011. Neue Entwicklungen in der Hochdrucktherapie: Eine Bewertung durch die Deutsche Hochdruckliga e.V. DHL® Deutsche Gesellschaft für Hypertonie und Prävention.
- [68] Dias, S., Welton, N. J., Caldwell, D. M., et al. 2010. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med*, 29, 932-44.
- [69] Ehlers, P. F. 2011. Hypoglykämien unter antidiabetischer Therapie. *kliniker*, 40 (6+7), 296-299.
- [70] FDA 2008. Guidance for Industry: Diabetes Mellitus: Developing Drugs and Therapeutic Biologics for Treatment and Prevention (DRAFT GUIDANCE).
- [71] Ferrannini, E., Fonseca, V., Zinman, B., et al. 2009. Fifty-two-week efficacy and safety of vildagliptin vs. glimepiride in patients with type 2 diabetes mellitus inadequately controlled on metformin monotherapy. *Diabetes, Obesity and Metabolism*. 11 (2) (pp 157-166), 2009.
- [72] Filozof, C. & Gautier, J. F. 2010. A comparison of efficacy and safety of vildagliptin and gliclazide in combination with metformin in patients with Type 2 diabetes inadequately controlled with metformin alone: a 52-week, randomized study. *Diabetic medicine : a journal of the British Diabetic Association*, 27, 318-26.
- [73] Freemantle, N., Holmes, J., Hockey, A., et al. 2008. How strong is the association between abdominal obesity and the incidence of type 2 diabetes? *International journal of clinical practice*, 62, 1391-6.
- [74] Frier, B. M., Schernthaner, G. & Heller, S. R. 2011. Hypoglycemia and cardiovascular risks. *Diabetes care*, 34 Suppl 2, S132-7.
- [75] G-BA. 2010. Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage III – Übersicht der Verordnungseinschränkungen und -ausschlüsse Glitazone zur Behandlung des Diabetes mellitus Typ 2.
- [76] Gallwitz, B., Rosenstock, J., Rauch, T., et al. 2012. 2-year efficacy and safety of linagliptin compared with glimepiride in patients with type 2 diabetes inadequately controlled on metformin: a randomised, double-blind, non-inferiority trial. *Lancet*, 380, 475-83.
- [77] Göke, B., Gallwitz, B., Eriksson, J., et al. 2010. Saxagliptin is non-inferior to glipizide in patients with type 2 diabetes mellitus inadequately controlled on metformin alone: a 52-week randomised controlled trial. *International journal of clinical practice*, 64, 1619-31.

- [78] Hader C, Beischer W, Braun A, et al. 2004. Evidenzbasierte Diabetes-Leitlinie der Deutschen Diabetes-Gesellschaft (DDG) - Diagnostik, Therapie und Verlaufskontrolle des Diabetes mellitus im Alter. *Diabetes und Stoffwechsel*, 13, 31-56.
- [79] Hein 2012. Niederschrift (finale Fassung) zum Beratungsgespräch gemäß § 8 AM-NutzenV, Beratungsanforderung 2011-B-35, Dapagliflozin zur Behandlung des Diabetes mellitus Typ 2
- [80] Higgins, J. P., Thompson, S. G., Deeks, J. J., et al. 2003. Measuring inconsistency in meta-analyses. *BMJ*, 327, 557-60.
- [81] Higgins, J. P. & Green, S. 2008. *Cochrane Handbook for Systematic Reviews of Interventions*. , The Cochrane Collaboration.
- [82] Holman, R. R., Paul, S. K., Bethel, M. A., et al. 2008. 10-year follow-up of intensive glucose control in type 2 diabetes. *The New England journal of medicine*, 359, 1577-89.
- [83] Holstein, A., Plaschke, A. & Egberts, E. H. 2003. Clinical characterisation of severe hypoglycaemia--a prospective population-based study. *Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association*, 111, 364-9.
- [84] Holstein, A., Patzer, O. M., Machalke, K., et al. 2012. Substantial increase in incidence of severe hypoglycemia between 1997-2000 and 2007-2010: a German longitudinal population-based study. *Diabetes care*, 35, 972-5.
- [85] ICTRP Search Portal. 2012a. *A 52-Week, International, Multi-centre, Randomised, Parallel-group, Double-blind, Active-controlled, Phase III study with a 156-Week Extension Period to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Metformin compared with Sulphonylurea in Combination with Metformin in Adult Patients with Type 2 Diabetes who have Inadequate Glycaemic Control on Metformin Therapy Alone* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2007-005220-33-GB> [Aufgerufen am 30.10.2012].
- [86] ICTRP Search Portal. 2012b. *A 24-week,Multi-centre,Int.,Double-blind,Rand.,Parallel-group,Plac.-Controlled,Phase III Study With a 78-week Ext.Per. to Evaluate the Effect of Dapagliflozin in Combination With Metformin on Body Weight in Subjects With Type2 Diabetes Mellitus Who Have Inadequate Glycemic Control on Metformin Alone* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00855166> [Aufgerufen am 30.10.2012].
- [87] ICTRP Search Portal. 2012c. *A 52-Week International, Multi-centre, Randomised, Parallel-group, Double-blind, Active-controlled, Phase III Study With a 156-Week Extension Period to Evaluate the Efficacy and Safety of Dapagliflozin in Combination With Metformin Compared With Sulphonylurea in Combination With Metformin in Adult Patients With Type 2 Diabetes Who Have Inadequate Glycaemic Control on Metformin Therapy Alone* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00660907> [Aufgerufen am 30.10.2012].
- [88] ICTRP Search Portal. 2012d. *A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase III Trial to Evaluate the Efficacy and Safety of BMS-477118 in*

- Combination With Metformin in Subjects With Type 2 Diabetes Who Have Inadequate Glycemic Control on Metformin Alone* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00121667> [Aufgerufen am 31.10.2012].
- [89] ICTRP Search Portal. 2012e. *Liraglutide Effect and Action in Diabetes (LEAD-2): Effect on Glycaemic Control After Once Daily Administration of Liraglutide in Combination With Metformin Versus Metformin Monotherapy Versus Metformin and Glimpiride Combination Therapy in Subjects With Type 2 Diabetes* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00318461> [Aufgerufen am 31.10.2012].
- [90] ICTRP Search Portal. 2012f. *The Effect of Liraglutide Compared to Sitagliptin, Both in Combination With Metformin in Subjects With Type 2 Diabetes. A 26-week, Randomised, Open-label, Active Comparator, Three-armed, Parallel-group, Multi-centre, Multinational Trial With a 52-week Extension* [Online]. Adresse: <http://apps.who.int/trialsearch/trial.aspx?trialid=NCT00700817> [Aufgerufen am 30.10.2012].
- [91] ICTRP Search Portal. 2012g. *A 24-Week, International, Multi-centre, Randomised, Parallel-group, Double-blind, Active-controlled, Phase III study to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Metformin* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2008-004916-12-DE> [Aufgerufen am 31.10.2012].
- [92] ICTRP Search Portal. 2012h. *Efficacy and Safety in Patients With Type 2 Diabetes Mellitus and Cardiovascular Disease* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01042977> [Aufgerufen am 30.10.2012].
- [93] ICTRP Search Portal. 2012i. *Efficacy and Safety in Patients With Type 2 Diabetes Mellitus, Cardiovascular Disease and Hypertension* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01031680> [Aufgerufen am 30.10.2012].
- [94] ICTRP Search Portal. 2012j. *Renal Mechanism of Action/Splay vs. TmG MOA* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00726505> [Aufgerufen am 30.10.2012].
- [95] Inzucchi, S. E., Bergenstal, R. M., Buse, J. B., et al. 2012. Management of Hyperglycemia in Type 2 Diabetes: A Patient-Centered Approach: Position Statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes care*, 35, 1364-79.
- [96] IQWiG 2009a. Glinide zur Behandlung des Diabetes mellitus Typ 2 - Abschlussbericht A05-05C. Version 1.0.
- [97] IQWiG. 2009b. Urin- und Blutzuckerselbstmessung bei Diabetes mellitus Typ 2
- [98] IQWiG 2011a. Systematische Leitlinienrecherche und -bewertung sowie Extraktion neuer und relevanter Empfehlungen für das DMP Diabetes mellitus Typ 2: Vorbericht.
- [99] IQWiG. 2011b. Allgemeine Methoden Version 4.0 vom 23.09.2011. Verfügbar: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf.

- [100] Ismail-Beigi, F., Craven, T., Banerji, M. A., et al. 2010. Effect of intensive treatment of hyperglycaemia on microvascular outcomes in type 2 diabetes: an analysis of the ACCORD randomised trial. *Lancet*, 376, 419-30.
- [101] Kahn, S. E., Haffner, S. M., Heise, M. A., et al. 2006. Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *The New England journal of medicine*, 355, 2427-43.
- [102] Kaku, K. 2009. Efficacy and safety of therapy with metformin plus pioglitazone in the treatment of patients with type 2 diabetes: a double-blind, placebo-controlled, clinical trial. *Current medical research and opinion*, 25, 1111-9.
- [103] Kerner, W., Brückel, J. & Böhm, B. O. 2004. Definition, Klassifikation und Diagnostik des Diabetes mellitus. *Evidenzbasierte Leitlinie DDG – Aktualisierung 10/2004* [Online]. Verfügbar: http://www.deutsche-diabetes-gesellschaft.de/leitlinien/EBL_Klassifikation_Update_2004.pdf (18.11.11).
- [104] Lundkvist, J., Berne, C., Bolinder, B., et al. 2005. The economic and quality of life impact of hypoglycemia. *The European journal of health economics : HEPAC : health economics in prevention and care*, 6, 197-202.
- [105] Mancina, G. 2007. Blood pressure reduction and cardiovascular outcomes: past, present, and future. *The American journal of cardiology*, 100, 3J-9J.
- [106] Matthaiei, S., Bierwirth, R., Fritsche, A., et al. 2009. Medikamentöse antihyperglykämische Therapie des Diabetes mellitus Typ 2. *Diabetologie* 2009, 4, 32-64.
- [107] Matthews, D. R., Charbonnel, B. H., Hanefeld, M., et al. 2005. Long-term therapy with addition of pioglitazone to metformin compared with the addition of gliclazide to metformin in patients with type 2 diabetes: a randomized, comparative study. *Diabetes/metabolism research and reviews*, 21, 167-74.
- [108] Matthews, D. R., Dejager, S., Ahren, B., et al. 2010. Vildagliptin add-on to metformin produces similar efficacy and reduced hypoglycaemic risk compared with glimepiride, with no weight gain: results from a 2-year study. *Diabetes, obesity & metabolism*, 12, 780-9.
- [109] Mengel K. 2011. Antidiabetika. In: SCHWABE U. & PFAFFRATH D. (Hrsg.) *Arzneiverordnungsreport 2011*. Springer.
- [110] National Collaborating Centre for Chronic Conditions 2008. Type 2 Diabetes: National Clinical Guideline for Management in Primary and Secondary Care (Update).
- [111] Nauck, M., Frid, A., Hermansen, K., et al. 2009. Efficacy and safety comparison of liraglutide, glimepiride, and placebo, all in combination with metformin, in type 2 diabetes: the LEAD (liraglutide effect and action in diabetes)-2 study. *Diabetes care*, 32, 84-90.
- [112] Nauck, M., Del Prato, S., Rohwedder, K., et al. 2010. Dapagliflozin vs glipizide in patients with type 2 diabetes mellitus inadequately controlled on metformin: 52-week results of a double-blind, randomised, controlled trial. *Diabetologia*.
- [113] Nauck, M. A., Meininger, G., Sheng, D., et al. 2007. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonyleurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: a

- randomized, double-blind, non-inferiority trial. *Diabetes, obesity & metabolism*, 9, 194-205.
- [114] Nauck, M. A., Del Prato, S., Meier, J. J., et al. 2011. Dapagliflozin versus glipizide as add-on therapy in patients with type 2 diabetes who have inadequate glycemic control with metformin: a randomized, 52-week, double-blind, active-controlled noninferiority trial. *Diabetes care*, 34, 2015-22.
- [115] NICE 2009. Type 2 diabetes: National clinical guideline 87 for the management of type 2 diabetes.
- [116] Oxford Outcomes Ltd. 2011a. Network Meta-Analysis of Anti-Diabetic Agents in Type 2 Diabetes Mellitus: Metformin add-on therapy.
- [117] Oxford Outcomes Ltd. 2011b. Systematic Review of Anti-Diabetic Agents in Type 2 Diabetes Mellitus: Metformin add-on therapy.
- [118] Oxford Outcomes Ltd. 2012a. Systematic Review of Anti-Diabetic Agents in Type 2 Diabetes Mellitus: Metformin add-on therapy Updated search results to 2012 (November 2012).
- [119] Oxford Outcomes Ltd. 2012b. Network Meta-Analysis of Anti-Diabetic Agents in Type 2 Diabetes Mellitus: Safety.
- [120] Oxford Outcomes Ltd. 2012c. Systematic Review of Anti-Diabetic Agents in Type 2 Diabetes Mellitus: Metformin add-on therapy Updated search results to 2012 (June 2012).
- [121] Pan, C., Xing, X., Han, P., et al. 2012. Efficacy and tolerability of vildagliptin as add-on therapy to metformin in Chinese patients with type 2 diabetes mellitus. *Diabetes, obesity & metabolism*.
- [122] Papathanassiou, K., Naka, K. K., Kazakos, N., et al. 2009. Pioglitazone vs glimepiride: Differential effects on vascular endothelial function in patients with type 2 diabetes. *Atherosclerosis*, 205, 221-6.
- [123] Patel, A., MacMahon, S., Chalmers, J., et al. 2008. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *The New England journal of medicine*, 358, 2560-72.
- [124] Perk, J., De Backer, G., Gohlke, H., et al. 2012. European Guidelines on cardiovascular disease prevention in clinical practice (version 2012): The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts) * Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *European heart journal*.
- [125] Pettersson, B., Rosenqvist, U., Deleskog, A., et al. 2011. Self-reported experience of hypoglycemia among adults with type 2 diabetes mellitus (Exhype). *Diabetes research and clinical practice*, 92, 19-25.
- [126] Porzolt, F., Clouth, J., Deutschmann, M., et al. 2010. Preferences of diabetes patients and physicians: a feasibility study to identify the key indicators for appraisal of health care values. *Health and quality of life outcomes*, 8, 125.

- [127] Pratley, R. E., Nauck, M., Bailey, T., et al. 2010. Liraglutide versus sitagliptin for patients with type 2 diabetes who did not have adequate glycaemic control with metformin: a 26-week, randomised, parallel-group, open-label trial. *Lancet*, 375, 1447-56.
- [128] Ray, K. K., Seshasai, S. R., Wijesuriya, S., et al. 2009. Effect of intensive control of glucose on cardiovascular outcomes and death in patients with diabetes mellitus: a meta-analysis of randomised controlled trials. *Lancet*, 373, 1765-72.
- [129] Raz, I., Chen, Y., Wu, M., et al. 2008. Efficacy and safety of sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes. *Current medical research and opinion*, 24, 537-50.
- [130] Saini, S. D., Schoenfeld, P., Kaulback, K., et al. 2009. Effect of medication dosing frequency on adherence in chronic diseases. *Am J Manag Care*, 15, e22-33.
- [131] Salvadeo, S. A. T., Maffioli, P., Ferrari, I., et al. 2010. Comparison between exenatide and glimepiride on metabolic control and on insulin resistance in type 2 diabetic patients with metformin therapy. In: SMITH, U. (Hrsg.) *Minutes of the 45th general assembly of the european association for the study of diabetes*.
- [132] Scott, R., Loeys, T., Davies, M. J., et al. 2008. Efficacy and safety of sitagliptin when added to ongoing metformin therapy in patients with type 2 diabetes. *Diabetes, obesity & metabolism*, 10, 959-69.
- [133] Skipka, G. & Bender, R. 2010. Intervention effects in the case of heterogeneity between three subgroups. Assessment within the framework of systematic reviews. *Methods of information in medicine*, 49, 613-7.
- [134] Standl, E., Fuchs, C. H., Parandeh-Shab, F., et al. 2000. Management der Hypertonie beim Patienten mit Diabetes mellitus.
- [135] Stratton, I. M., Adler, A. I., Neil, H. A., et al. 2000. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. *BMJ*, 321, 405-12.
- [136] Taskinen, M. R., Rosenstock, J., Tamminen, I., et al. 2011. Safety and efficacy of linagliptin as add-on therapy to metformin in patients with type 2 diabetes: a randomized, double-blind, placebo-controlled study. *Diabetes, obesity & metabolism*, 13, 65-74.
- [137] UK Prospective Diabetes Study Group 1998a. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. UK Prospective Diabetes Study Group. *BMJ*, 317, 703-13.
- [138] UK Prospective Diabetes Study Group 1998b. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). UK Prospective Diabetes Study (UKPDS) Group. *Lancet*, 352, 854-65.
- [139] UK Prospective Diabetes Study Group 1998c. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. *Lancet*, 352, 837-53.
- [140] Vega, G. L., Adams-Huet, B., Peshock, R., et al. 2006. Influence of body fat content and distribution on variation in metabolic risk. *The Journal of clinical endocrinology and metabolism*, 91, 4459-66.

- [141] Vexiau, P., Mavros, P., Krishnarajah, G., et al. 2008. Hypoglycaemia in patients with type 2 diabetes treated with a combination of metformin and sulphonylurea therapy in France. *Diabetes, obesity & metabolism*, 10 Suppl 1, 16-24.
- [142] Viberti, G., Kahn, S. E., Greene, D. A., et al. 2002. A diabetes outcome progression trial (ADOPT): an international multicenter study of the comparative efficacy of rosiglitazone, glyburide, and metformin in recently diagnosed type 2 diabetes. *Diabetes care*, 25, 1737-43.
- [143] Weissenbacher E, Autenrieth M, Bauer H, et al. 2012. Konsens für Therapie-Empfehlungen von Urogenitalinfektionen bei Diabetes mellitus Typ-2.
- [144] Wiesner, T. 2010. Ein aktueller Blick auf eine alte Bekannte: die Hypoglykämie.
- [145] Wittchen, H.-U., Pieper, L., Eichler, T., et al. 2007. Prävalenz und Versorgung von Diabetes mellitus und Herz-Kreislauf-Erkrankungen: DETECT - eine bundesweite Versorgungsstudie an über 55.000 Hausarztpatienten.
- [146] Yang, W., Pan, C. Y., Tou, C., et al. 2011. Efficacy and safety of saxagliptin added to metformin in Asian people with type 2 diabetes mellitus: a randomized controlled trial. *Diabetes Research & Clinical Practice*, 94, 217-224.
- [147] Yang, W., Guan, Y., Shentu, Y., et al. 2012. The addition of sitagliptin to ongoing metformin therapy significantly improves glycemic control in Chinese patients with type 2 diabetes. *J Diabetes*, 4, 227-37.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.2 beschriebene bibliografische Literaturrecherche an, und zwar einzeln für jede Datenbank. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁹] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

¹⁹ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

Für die Identifikation von Studien (RCT), zum Vergleich der Sicherheit und Wirksamkeit von Dapagliflozin mit anderen Antidiabetika wurde eine gemeinsame Literaturrecherche nach Studien mit Dapagliflozin und Studien mit anderen Arzneimitteln zur Behandlung von Patienten mit Typ-2-Diabetes durchgeführt. Anschließend wurde eine MTC durchgeführt. Die bibliografische Literaturrecherche wurde in den Datenbanken MEDLINE, EMBASE und Cochrane (*Cochrane Central Register of Controlled Trials*) durchgeführt. Die Suchstrategien sind in Tabelle 4-144, Tabelle 4-145, Tabelle 4-146 (Literaturrecherche am 11. Mai 2011) sowie Tabelle 4-147, Tabelle 4-148, Tabelle 4-149 (Aktualisierung der Literaturrecherche am 19. April 2012) und der Tabelle 4-150, Tabelle 4-151, Tabelle 4-152 (Aktualisierung am 09.10.2012) dargestellt. Die Suche wurde für Abschnitt 4.3.2.1 durchgeführt.

Tabelle 4-144: Ovid: Suchstrategie MEDLINE 1946 bis heute: am 11. Mai 2011

Datenbankname	Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid OLDMEDLINE(R)
Suchoberfläche	Ovid
Datum der Suche	11. Mai 2011
Zeitsegment	1946 bis 2011 Woche 19
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$).mp.	2758
2	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$).mp.	1449
3	dipeptidyl-peptidase IV inhibitors/	2390
4	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	2957
5	dapagliflozin.mp.	169
6	Glucagon-Like Peptide 1/	8103
7	(Glucagon-Like Peptide 1 or GLP-1).mp.	11637
8	(exenatide or liraglutide).mp.	2958
9	Metformin\$.mp.	29123
10	metformin/	26303
11	glyburide/	20891

12	(glibenclamide or glyburide).mp.	24931
13	gliclazide/	3973
14	gliclazide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4263
15	glimepiride.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	3828
16	glipizide/	4148
17	glipizide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4531
18	sulfonylurea compounds/ or sulfonylurea derivative/	10686
19	Thiazolidinediones/	15087
20	pioglitazone.mp.	11862
21	(meglitinide\$ or nateglinide\$ or repaglinide\$).mp.	4443
22	exp insulin/	331615
23	insulin\$.mp.	675750
24	or/1-23	716032
25	exp drug combinations/	200322
26	(drug therap\$ or drug combination\$).mp.	555340
27	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	442055
28	monotherap\$.mp.	84598
29	or/25-28	1021936
30	24 and 29	35107
31	exp Diabetes Mellitus, Type 2/	159588
32	(MODY or NIDDM or T2DM).mp.	22970
33	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	109525
34	((typ\$ 2 or typ\$ II) adj diabet\$).mp.	120726
35	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$).mp.	92402
36	or/31-35	277493
37	randomized controlled trial.pt.	306012
38	controlled clinical trial.pt.	82318
39	randomized.ab.	493774
40	placebo.ab.	279655
41	clinical trials as topic.sh.	153873
42	randomly.ab.	355061

43	trial.ti.	204270
44	or/37-43	1306271
45	exp animals/ not humans.sh.	5247601
46	99 not 100	1232093
47	30 and 36 and 46	3566
48	limit 47 to humans	3280

*a*Medline and EMBASE searches were executed jointly within the Ovid interface.

ab = abstract; **mp** = multipurpose; **sh** = subject heading; **ti** = title

Tabelle 4-145: Ovid: Suchstrategie EMBASE 1980 bis heute: am 11. Mai 2011

Datenbankname	Embase
Suchoberfläche	Ovid
Datum der Suche	11. Mai 2011
Zeitsegment	1980 bis 2011 Woche 19
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$).mp.	2758
2	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$).mp.	1449
3	dipeptidyl-peptidase IV inhibitor/	2390
4	sitagliptin/ or vildagliptin/ or saxagliptin/	2168
5	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	2957
6	dapagliflozin/	124
7	dapagliflozin.mp.	169
8	Glucagon-Like Peptide 1/	8103
9	(Glucagon-Like Peptide 1 or GLP-1).mp.	11637
10	exenatide/ or liraglutide/	3332
11	(exenatide or liraglutide).mp.	2958
12	Metformin\$.mp.	29123
13	metformin/	26303

14	glibenclamide/	20891
15	(glibenclamide or glyburide).mp.	24931
16	gliclazide/	3973
17	gliclazide.mp.	4263
18	glimepiride/	3099
19	glimepiride.mp.	3828
20	glipizide/	4148
21	glipizide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4531
22	sulfonylurea derivative/	6630
23	Thiazolidinediones/	15087
24	pioglitazone/	8896
25	pioglitazone.mp.	11862
26	meglitinide/	893
27	nateglinide/	1746
28	repaglinide/	2196
29	(nateglinide or repaglinide).mp.	3798
30	exp insulin derivative/	199424
31	insulin\$.mp.	675750
32	or/1-31	716102
33	exp drug combinations/	200322
34	(drug therap\$ or drug combination\$).mp.	555340
35	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	442055
36	monotherap\$.mp.	84598
37	or/33-36	1021936
38	32 and 37	35067
39	non insulin dependent diabetes mellitus/	159429
40	(MODY or NIDDM or T2DM).mp.	22970
41	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	109525
42	((typ\$ 2 or typ\$ II) adj diabet\$).mp.	120726
43	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$).mp.	92402
44	or/39-43	277338

45	Randomized controlled trial/	600436
46	controlled clinical trial/	254215
47	randomized.ab.	493774
48	placebo.ab.	279655
49	randomly.ab.	355061
50	trial.ti.	204270
51	or/45-50	1309559
52	exp animals/ not humans.sh.	5247601
53	51 not 52	1232221
54	38 and 44 and 53	3824
55	limit 54 to humans	3537

ab = abstract; *mp* = multipurpose; *sh* = subject heading; *ti* = title

Tabelle 4-146: Ovid: Suchstrategie Cochrane 1948 bis heute: am 11. Mai 2011

Datenbankname	EBM Reviews-Cochrane Central Register of controlled Trials
Suchoberfläche	Ovid
Datum der Suche	11. Mai 2011
Zeitsegment	1948 bis 2011, Woche 19
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	metformin/	1026
2	metformin\$.mp.	1471
3	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$).mp.	113
4	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$).mp.	42
5	dipeptidyl-peptidase IV inhibitors/	96
6	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	139
7	exp Glucagon-Like Peptide 1/	288
8	(Glucagon-Like Peptide 1 or GLP-1).mp.	404
9	(exenatide or liraglutide).mp.	110

10	dapagliflozin.mp.	6
11	glyburide/	445
12	glibenclamide.mp.	507
13	gliclazide/	134
14	gliclazide.mp.	194
15	glimepiride.mp.	180
16	glipizide/	121
17	glipizide.mp.	183
18	sulfonylurea compounds/ or sulfonylurea derivative/	439
19	exp insulin/	7214
20	insulin\$.mp.	15768
21	or/1-20	16714
22	drug therapy, combination/ 23 (drug therap\$ or drug combination\$).mp.	21069
24	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	60351
25	monotherap\$.mp.	18688
26	22 or 23 or 24 or 25	6387
27	21 and 26	70577
28	Diabetes Mellitus, Type 2/	2532
29	(MODY or NIDDM or T2DM).mp.	5934
30	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	1190
31	((typ\$ 2 or typ\$ II) adj diabet\$).mp	1880
32	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$).mp.	5960
33	or/28-32	5669
34	27 and 33	11092
35	limit 34 to medline records	1435

Die Literaturrecherche wurde am 04. Juni 2012 aktualisiert.

Tabelle 4-147: Ovid: Suchstrategie MEDLINE 1946 bis heute: am 04. Juni 2012

Datenbankname	Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid OLDMEDLINE(R)
Suchoberfläche	Ovid
Datum der Suche	04. Juni 2012
Zeitsegment	1946 bis 2012 Woche 23
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$.mp.	3723
2	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$.mp.	2222
3	dipeptidyl-peptidase IV inhibitors/	3224
4	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	4261
5	dapagliflozin.mp.	325
6	Glucagon-Like Peptide 1/	10205
7	(Glucagon-Like Peptide 1 or GLP-1).mp.	14792
8	(exenatide or liraglutide).mp.	4194
9	Metformin\$.mp.	34501
10	metformin/	30951
11	glyburide/	22397
12	(glibenclamide or glyburide).mp.	26863
13	gliclazide/	4370
14	gliclazide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4703
15	glimepiride.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4457
16	glipizide/	4507
17	glipizide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4938
18	sulfonylurea compounds/ or sulfonylurea derivative/	11354
19	Thiazolidinediones/	17033
20	pioglitazone.mp.	13896
21	(meglitinide\$ or nateglinide\$ or repaglinide\$.mp.	4974

22	exp insulin/	356844
23	insulin\$.mp.	742964
24	or/1-23	789316
25	exp drug combinations/	222004
26	(drug therap\$ or drug combination\$).mp.	600818
27	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	463613
28	monotherap\$.mp.	96114
29	or/25-28	1099706
30	24 and 29	39368
31	exp Diabetes Mellitus, Type 2/	181767
32	(MODY or NIDDM or T2DM).mp.	26897
33	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	125209
34	((typ\$ 2 or typ\$ II) adj diabet\$).mp.	145968
35	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$).mp.	103661
36	or/31-35	318731
37	randomized controlled trial.pt.	328617
38	controlled clinical trial.pt.	84173
39	randomized.ab.	555552
40	placebo.ab.	303899
41	clinical trials as topic.sh.	160292
42	randomly.ab.	395918
43	trial.ti.	235880
44	or/37-43	1440885
45	exp animals/ not humans.sh.	5499636
46	99 not 100	1358009
47	30 and 36 and 46	4124
48	limit 47 to humans	3818

^aMedline and EMBASE searches were executed jointly within the Ovid interface.

ab = abstract; **mp** = multipurpose; **sh** = subject heading; **ti** = title

Tabelle 4-148: Ovid: Suchstrategie EMBASE 1980 bis heute: am 04. Juni 2012

Datenbankname	Embase
Suchoberfläche	Ovid
Datum der Suche	04. Juni 2012
Zeitsegment	1980 bis 2012 Woche 23
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$).mp.	3723
2	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$).mp.	2222
3	dipeptidyl-peptidase IV inhibitor/	3224
4	sitagliptin/ or vildagliptin/ or saxagliptin/	3038
5	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	4261
6	dapagliflozin/	236
7	dapagliflozin.mp.	325
8	Glucagon-Like Peptide 1/	10205
9	(Glucagon-Like Peptide 1 or GLP-1).mp.	14792
10	exenatide/ or liraglutide/	4545
11	(exenatide or liraglutide).mp.	4194
12	Metformin\$.mp.	34501
13	metformin/	30951
14	glibenclamide/	22397
15	(glibenclamide or glyburide).mp.	26863
16	gliclazide/	4370
17	gliclazide.mp.	4703
18	glimepiride/	3599
19	glimepiride.mp.	4457
20	glipizide/	4507
21	glipizide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4938
22	sulfonylurea derivative/	7094

23	Thiazolidinediones/	17033
24	pioglitazone/	10441
25	pioglitazone.mp.	13896
26	meglitinide/	980
27	nateglinide/	1921
28	repaglinide/	2464
29	(nateglinide or repaglinide).mp.	4243
30	exp insulin derivative/	221313
31	insulin\$.mp.	742964
32	or/1-31	789531
33	exp drug combinations/	222004
34	(drug therap\$ or drug combination\$).mp.	600818
35	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	463613
36	monotherap\$.mp.	96114
37	or/33-36	1099706
38	32 and 37	39333
39	non insulin dependent diabetes mellitus/	181606
40	(MODY or NIDDM or T2DM).mp.	26897
41	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	125209
42	((typ\$ 2 or typ\$ II) adj diabet\$).mp.	145968
43	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$).mp.	103661
44	or/39-43	318574
45	Randomized controlled trial/	651186
46	controlled clinical trial/	473245
47	randomized.ab.	555552
48	placebo.ab.	303899
49	randomly.ab.	395918
50	trial.ti.	235880
51	or/45-50	1534545
52	exp animals/ not humans.sh.	5499636
53	51 not 52	1447857

54	38 and 44 and 53	4612
55	limit 54 to humans	4304

ab = abstract; *mp* = multipurpose; *sh* = subject heading; *ti* = title

Tabelle 4-149: Ovid: Suchstrategie Cochrane 1948 bis heute: am 04. Juni 2012

Datenbankname	EBM Reviews-Cochrane Central Register of controlled Trials
Suchoberfläche	Ovid
Datum der Suche	04. Juni 2012
Zeitsegment	1948 bis 2012, Woche 23
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	metformin/	1131
2	metformin\$.mp.	1643
3	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$).mp.	149
4	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$).mp.	56
5	dipeptidyl-peptidase IV inhibitors/	125
6	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	191
7	exp Glucagon-Like Peptide 1/	341
8	(Glucagon-Like Peptide 1 or GLP-1).mp.	470
9	(exenatide or liraglutide).mp.	139
10	dapagliflozin.mp.	11
11	glyburide/	457
12	glibenclamide.mp.	518
13	gliclazide/	139
14	gliclazide.mp.	200
15	glimepiride.mp.	204
16	glipizide/	125
17	glipizide.mp.	191
18	sulfonylurea compounds/ or sulfonylurea derivative/	462

19	exp insulin/	7482
20	insulin\$.mp.	16717
21	or/1-20	17783
22	drug therapy, combination/	21849
23	(drug therap\$ or drug combination\$.mp	63834
24	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	19522
25	monotherap\$.mp.	6803
26	22 or 23 or 24 or 25	74416
27	21 and 26	2759
28	Diabetes Mellitus, Type 2/	6488
29	(MODY or NIDDM or T2DM).mp.	1293
30	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	2027
31	((typ\$ 2 or typ\$ II) adj diabet\$.mp	6676
32	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$.mp.	6077
33	or/28-32	12067
34	27 and 33	1594

Die Literaturrecherche wurde am 09. Oktober 2012 erneut aktualisiert.

Tabelle 4-150: Ovid: Suchstrategie MEDLINE 1946 bis heute: am 09. Oktober 2012

Datenbankname	Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid OLDMEDLINE(R)
Suchoberfläche	Ovid
Datum der Suche	09. Oktober 2012
Zeitsegment	1946 bis Oktober 2012 Woche 41
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$.mp.	4090
2	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$.mp.	2479
3	dipeptidyl-peptidase IV inhibitors/	3520
4	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	4772
5	dapagliflozin.mp.	384
6	Glucagon-Like Peptide 1/	11111
7	(Glucagon-Like Peptide 1 or GLP-1).mp.	16053
8	(exenatide or liraglutide).mp.	4656
9	Metformin\$.mp.	36520
10	metformin/	32806
11	glyburide/	22919
12	(glibenclamide or glyburide).mp.	27476
13	gliclazide/	4509
14	gliclazide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4853
15	glimepiride.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	4678
16	glipizide/	4621
17	glipizide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	5067
18	sulfonylurea compounds/ or sulfonylurea derivative/	11567
19	Thiazolidinediones/	17759
20	pioglitazone.mp.	14661
21	(meglitinide\$ or nateglinide\$ or repaglinide\$.mp.	5127
22	exp insulin/	366877
23	insulin\$.mp.	768732
24	or/1-23	817318
25	exp drug combinations/	227427
26	(drug therap\$ or drug combination\$.mp.	613891
27	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	469895
28	monotherap\$.mp.	100274
29	or/25-28	1123269
30	24 and 29	40741

31	exp Diabetes Mellitus, Type 2/	191659
32	(MODY or NIDDM or T2DM).mp.	28539
33	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	131710
34	((typ\$ 2 or typ\$ II) adj diabet\$).mp.	155601
35	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$).mp.	116924
36	or/31-35	339137
37	randomized controlled trial.pt.	339434
38	controlled clinical trial.pt.	85392
39	randomized.ab.	580671
40	placebo.ab.	313449
41	clinical trials as topic.sh.	163003
42	randomly.ab.	411814
43	trial.ti.	245901
44	or/37-43	1492888
45	exp animals/ not humans.sh.	5593480
46	99 not 100	1406953
47	30 and 36 and 46	4405
48	limit 47 to humans	4084

aMedline and EMBASE searches were executed jointly within the Ovid interface.

ab = abstract; **mp** = multipurpose; **sh** = subject heading; **ti** = title

Tabelle 4-151: Ovid: Suchstrategie EMBASE 1980 bis heute: am 09. Oktober 2012

Datenbankname	Embase
Suchoberfläche	Ovid
Datum der Suche	09. Oktober 2012
Zeitsegment	1980 bis 2012 Woche 41
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$).mp.	4090

2	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$.mp.	2479
3	dipeptidyl-peptidase IV inhibitor/	3520
4	sitagliptin/ or vildagliptin/ or saxagliptin/	3365
5	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	4772
6	dapagliflozin/	282
7	dapagliflozin.mp.	384
8	Glucagon-Like Peptide 1/	11111
9	(Glucagon-Like Peptide 1 or GLP-1).mp.	16053
10	exenatide/ or liraglutide/	4971
11	(exenatide or liraglutide).mp.	4656
12	Metformin\$.mp.	36520
13	metformin/	32806
14	glibenclamide/	22919
15	(glibenclamide or glyburide).mp.	27476
16	gliclazide/	4509
17	gliclazide.mp.	4853
18	glimepiride/	3765
19	glimepiride.mp.	4678
20	glipizide/	4621
21	glipizide.mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, ui]	5067
22	sulfonylurea derivative/	7210
23	Thiazolidinediones/	17759
24	pioglitazone/	10995
25	pioglitazone.mp.	14661
26	meglitinide/	1007
27	nateglinide/	1970
28	repaglinide/	2541
29	(nateglinide or repaglinide).mp.	4368
30	exp insulin derivative/	228446
31	insulin\$.mp.	768732
32	or/1-31	817574
33	exp drug combinations/	227427

34	(drug therap\$ or drug combination\$).mp.	613891
35	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	469895
36	monotherap\$.mp.	100274
37	or/33-36	1123269
38	32 and 37	40708
39	non insulin dependent diabetes mellitus/	191497
40	(MODY or NIDDM or T2DM).mp.	28539
41	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	131718
42	((typ\$ 2 or typ\$ II) adj diabet\$).mp.	155601
43	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$).mp.	116924
44	or/39-43	338983
45	Randomized controlled trial/	669999
46	controlled clinical trial/	477961
47	randomized.ab.	580671
48	placebo.ab.	313449
49	randomly.ab.	411814
50	trial.ti.	245901
51	or/45-50	1586571
52	exp animals/ not humans.sh.	5593480
53	51 not 52	1496694
54	38 and 44 and 53	4897
55	limit 54 to humans	4574

ab = abstract; *mp* = multipurpose; *sh* = subject heading; *ti* = title

Tabelle 4-152: Ovid: Suchstrategie Cochrane 1948 bis heute: am 09. Oktober 2012

Datenbankname	EBM Reviews-Cochrane Central Register of controlled Trials
Suchoberfläche	Ovid
Datum der Suche	09. Oktober 2012
Zeitsegment	1948 bis 2012, Woche 41
Suchfilter	Suchfilter für randomisierte kontrollierte Studien (Higgins und Green, 2008)

#	Suchbegriff	Ergebnis
1	metformin/	1137
2	metformin\$.mp.	1650
3	(dipeptidyl peptidase-4 inhibitor\$ or dipeptidyl peptidase-IV inhibitor\$).mp.	152
4	(dpp-iv inhibitor\$ or dpp-4 inhibitor\$).mp.	57
5	dipeptidyl-peptidase IV inhibitors/	128
6	(vildagliptin or sitagliptin or saxagliptin or linagliptin).mp.	194
7	exp Glucagon-Like Peptide 1/	344
8	(Glucagon-Like Peptide 1 or GLP-1).mp.	478
9	(exenatide or liraglutide).mp.	141
10	dapagliflozin.mp.	11
11	glyburide/	457
12	glibenclamide.mp.	519
13	gliclazide/	139
14	gliclazide.mp.	201
15	glimepiride.mp.	206
16	glipizide/	125
17	glipizide.mp.	191
18	sulfonylurea compounds/ or sulfonylurea derivative/	465
19	exp insulin/	7497
20	insulin\$.mp.	16770
21	or/1-20	17844
22	drug therapy, combination/	21889

23	(drug therap\$ or drug combination\$).mp	64108
24	((combination\$ or oral or multiple) adj (therap\$ or agent\$ or drug\$ or treatment\$)).mp.	19590
25	monotherap\$.mp.	6821
26	22 or 23 or 24 or 25	74707
27	21 and 26	2767
28	Diabetes Mellitus, Type 2/	6524
29	(MODY or NIDDM or T2DM).mp.	1296
30	(non insulin\$ depend\$ or noninsulin\$ depend\$ or noninsulin?depend\$ of non insulin?depend).mp.	2030
31	((typ\$ 2 or typ\$ II) adj diabet\$).mp	6720
32	((late or adult\$ or matur\$ or slow or stable\$) adj diabet\$).mp.	6301
33	or/28-32	12239
34	27 and 33	1607

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.2 beschriebene Suche in Studienregistern an, und zwar einzeln für jedes Studienregister. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. *clinicaltrials.gov*), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer	23

Studienregistersuche zum zu bewertenden Arzneimittel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	29.10.2012
Suchstrategie	Dapagliflozin OR BMS-512148 [Intervention] AND „Diabetes Mellitus“ [Condition] AND (Phase II OR Phase III OR Phase IV) [Phase]
Treffer	33

Die Suche nach registrierten Studien mit Dapagliflozin zur Behandlung des Typ-2-Diabetes erfolgte anhand spezifischer Stichwörter. Die Datenbankabfrage über *ClinicalTrials.gov* ergab anhand der Suchanfrage 33 Treffer.

Da das Studienregister *clinicalstudyresults.org* nicht mehr zur Verfügung steht, wurde hier keine Suche vorgenommen.

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	29.10.2012
Suchstrategie	Dapagliflozin OR BMS-512148 [Intervention] AND „Diabetes mellitus“ [Condition] AND „All“ [Recruitment Status]
Treffer	39

Die Datenbankabfrage über das ICTRP ergab anhand der Suchanfrage 39 Treffer.

Die Suchen über die „*Advanced Search*“-Funktionen der zwei berücksichtigten Studienregister wurde für Abschnitt 4.3.1 durchgeführt.

Für die Identifikation von Studien (RCT), zum Vergleich der Sicherheit und Wirksamkeit von Dapagliflozin mit anderen Antidiabetika wurde eine gemeinsame Studienregistersuche nach Studien mit Dapagliflozin und Studien mit anderen Arzneimitteln zur Behandlung von Patienten mit Typ-2-Diabetes durchgeführt. Anschließend wurde eine MTC durchgeführt.

Die Datenbanken *clinicaltrials.gov*, *ICTRP* und *ISRCTN* wurden nach randomisierten klinischen Studien (Volltext verfügbar) untersucht. Die Suche wurde für Abschnitt 4.3.2.1 durchgeführt.

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	09.10.2012
Suchstrategie	Closed Studies Interventional Studies diabetes AND ("type 2" OR "type II") AND NOT ("type 1" OR "type I") metformin AND (dapagliflozin OR exenatide OR liraglutide OR pioglitazone OR glyburide OR glimepiride OR glipizide OR gliclazide OR nateglinide OR repaglinide OR sitagliptin OR vildagliptin OR saxagliptin OR linagliptin) Adult, Senior
Treffer	297

Die Suche nach registrierten Studien zu Dapagliflozin mit anderen Antidiabetika erfolgte anhand spezifischer Stichwörter. Die Datenbankabfrage über *clinicaltrials.gov* ergab 297 Treffer.

Da das Studienregister *clinicalstudyresults.org* nicht mehr zur Verfügung steht, wurde hier keine Suche vorgenommen.

Studienregister	<i>ICTRP</i>
Internetadresse	http://apps.who.int/trialsearch
Datum der Suche	09.10.2012
Suchstrategie	<ul style="list-style-type: none">• Title: NOT type 1 NOT gestation*• Condition: diabetes• Intervention: metformin AND dapagliflozin OR exenatide OR liraglutide OR pioglitazone OR glyburide OR glimepiride OR glipizide OR gliclazide OR nateglinide OR repaglinide OR sitagliptin OR vildagliptin OR saxagliptin OR linagliptin
Treffer	238

In der Datenbank *ICTRP* ergab die Suche über die vorgegebene Eingabemaske 238 Treffer für 186 Studien zum zu bewertenden Arzneimittel mit anderen Antidiabetika.

Studienregister	<i>ISRCTN</i>
Internetadresse	http://www.controlled-trials.com/
Datum der Suche	09.10.2012
Suchstrategie	<ul style="list-style-type: none">• Diabetes AND (metformin AND (dapagliflozin OR exenatide OR liraglutide OR pioglitazone OR glyburide OR glimepiride OR glipizide OR gliclazide OR nateglinide OR repaglinide OR sitagliptin OR vildagliptin OR saxagliptin OR linagliptin)) Treffer: 23
Treffer	23

Mittels der Datenbankabfrage bei *ISRCTN* konnten über die vorgegebene Eingabemaske 23 zutreffende Studien zum zu bewertenden Arzneimittel mit anderen Antidiabetika ermittelt werden.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund

Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrunds auf.

Tabelle 4-153 (Anhang): Im Volltext ausgeschlossene nicht relevante Publikationen (nach Literaturrecherche vom 11. Mai 2011)

	Im Volltext ausgeschlossene Publikation	Grund
1	OTTERBECK, P. E. & BANERJI, M. A. The efficacy and safety of vildagliptin in the GALIANT trial: Chronic kidney disease and other applications. Expert Review of Endocrinology and Metabolism.6 (2) (pp 143-151), 2011.Date of Publication: March 2011.	Falsche Dosierung
2	GARBER, A., HENRY, R. R., RATNER, R., HALE, P., CHANG, C. T. & BODE, B. 2011. Liraglutide, a once-daily human glucagon-like peptide 1 analogue, provides sustained improvements in glycaemic control and weight for 2 years as monotherapy compared with glimepiride in patients with type 2 diabetes. Diabetes, Obesity and Metabolism, 13, April.	Keine Vergleichstherapie von Interesse
3	GARCIA-HERNANDEZ, P., ARECHA VALETA-GRANELL, M. R., YAMAMOTO, J., FALAHATI, A., GONZALEZ-GALVEZ, G. & GRUPO DE INVESTIGADORES DE, L. [Liraglutide and glimepiride on glycaemic control in type 2 diabetes in the Mexican cohort (LEAD 3)]. [Spanish]. Revista medica del Instituto Mexicano del Seguro Social.48 (5) (pp 543-548), 2010.Date of Publication: 2010 Sep-Oct.	Keine Vergleichstherapie von Interesse
4	SCHEEN, A. J., CHARPENTIER, G., OSTGREN, C. J., HELLQVIST, A. & GAUSE-NILSSON, I. Efficacy and safety of saxagliptin in combination with metformin compared with sitagliptin in combination with metformin in adult patients with type 2 diabetes mellitus. Diabetes/Metabolism Research and Reviews.26 (7) (pp 540-549), 2010.Date of Publication: October 2010.	Vergleich nur innerhalb einer Wirkstoffgruppe
5	PFÜTZNER, A., PAZ-PACHECO, E., ALLEN, E., FREDERICH, R. & CHEN, R. 2011. Initial combination therapy with saxagliptin and metformin provides sustained glycaemic control and is well tolerated for up to 76 weeks. Diabetes, Obesity and Metabolism, 13, June.	Studie beinhaltet nur Vergleiche zwischen Monotherapien mit Metformin oder einem anderen relevanten Komparator/Placebo

	Im Volltext ausgeschlossene Publikation	Grund
6	PEREZ, A., JACKS, R., ARORA, V. & SPANHEIMER, R. Effects of pioglitazone and metformin fixed-dose combination therapy on cardiovascular risk markers of inflammation and lipid profile compared with pioglitazone and metformin monotherapy in patients with type 2 diabetes. Journal of Clinical Hypertension.12 (12) (pp 973-982), 2010.Date of Publication: December 2010.	Studie beinhaltet nur Vergleiche zwischen Monotherapien mit Metformin oder einem anderen relevanten Komparator/Placebo
7	PAREEK, A., CHANDURKAR, N., ZAWAR, S. & AGRAWAL, N. Evaluation of efficacy and tolerability of gliclazide and metformin combination: A multicentric study in patients with type 2 diabetes mellitus uncontrolled on monotherapy with sulfonylurea or metformin. American Journal of Therapeutics.17 (6) (pp 559-565), 2010.Date of Publication: November-December 2010.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
8	HADINGER, M., WERZOWA, J., VOIGT, H. C., PLEINER, J., STEMER, G., HECKING, M., DOLLER, D., HORL, W. H., WEICHHART, T. & SAEMANN, M. D. 2010. A randomized, placebo-controlled, double-blind, prospective trial to evaluate the effect of vildagliptin in new-onset diabetes mellitus after kidney transplantation. Trials [Electronic Resource], 11, 91.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
9	NAKA, K. K., PAPPAS, K., PAPATHANASSIOU, K., PAPAMICHAEL, N. D., KAZAKOS, N., KANIOGLOU, C., MAKRIYIANNIS, D., KATSOURAS, C. S., LIVERIS, K., TSATSOULIS, A. & MICHALIS, L. K. Lack of effects of pioglitazone on cardiac function in patients with type 2 diabetes and evidence of left ventricular diastolic dysfunction: A tissue doppler imaging study. Cardiovascular Diabetology.9 , 2010.Article Number: 57.Date of Publication: 23 Sep 2010.,	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
10	RASKIN, P. & MORA, P. F. Glycaemic control with liraglutide: The phase 3 trial programme. International Journal of Clinical Practice.64 (SUPPL, 167), October.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)

	Im Volltext ausgeschlossene Publikation	Grund
11	FORST, T., WEBER, M. M., LOBIG, M., LEHMANN, U., MULLER, J., HOHBERG, C., FRIEDRICH, C., FUCHS, W. & PFUTZNER, A. Pioglitazone in addition to metformin improves erythrocyte deformability in patients with Type 2 diabetes mellitus. <i>Clinical Science</i> .119 (8) (pp 345-351), 2010.Date of Publication: October 2010.	Falsche Dosierung
12	ESPOSITO, K., IDA, M. M., CIOTOLA, M., DI, P. C., SCOGNAMIGLIO, P., GICCHINO, M., PETRIZZO, M., SACCOMANNO, F., BENEDUCE, F., CERIELLO, A. & GIUGLIANO, D. Effects of a mediterranean-style diet on the need for antihyperglycemic drug therapy in patients with newly diagnosed type 2 diabetes: A randomized trial. <i>Obstetrical and Gynecological Survey</i> .65 (6) (pp 379-380), 2010.Date of Publication: June 2010.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
13	RETT, A. S. Insulin v. oral therapy for newly diagnosed type 2 diabetes. <i>Medicine Today</i> .11 (1) (pp 63+64), 2010.Date of Publication: January 2010.	Kein Volltext verfügbar
14	DIMIC, D., VELOJIC, G. M., ANTIC, S. & RADENKOVIC, S. Evaluation of the repaglinide efficiency in comparison to the glimepiride in the type 2 diabetes patients poorly regulated by the metmorfine administration. <i>Bratislava Medical Journal</i> .110 (6) (pp 335-339), 2009.Date of Publication: 2009.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
15	GOODMAN, M., THURSTON, H. & PENMAN, J. Efficacy and tolerability of vildagliptin in patients with type 2 diabetes inadequately controlled with metformin monotherapy. <i>Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme</i> .41 (5) (pp 368-373), 2009.Date of Publication: May 2009.	Falsche Dosierungsschema

	Im Volltext ausgeschlossene Publikation	Grund
16	LUND, S. S., TARNOW, L., FRANDBSEN, M., NIELSEN, B. B., HANSEN, B. V., PEDERSEN, O., PARVING, H. H. & VAAG, A. A. Combining insulin with metformin or an insulin secretagogue in non-obese patients with type 2 diabetes: 12 Month, randomised, double blind trial. <i>BMJ</i> .339 (7730) (pp 1121-1124), 2009.Date of Publication: 2009.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
17	ROSENSTOCK, J., REUSCH, J., BUSH, M., YANG, F. & STEWART, M. Potential of albiglutide, a long-acting GLP-1 receptor agonist, in type 2 diabetes: A randomized controlled trial exploring weekly, biweekly, and monthly dosing. <i>Diabetes Care</i> .32 (10) (pp 1880-1886), 2009.Date of Publication: October 2009.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält

	Im Volltext ausgeschlossene Publikation	Grund
18	<p>REINHARDT, R., LYNESS, W., ADLER, J., AHMED, A., AHMED, K., ANDERSON, D., ARGOUD, G., ARIF, A., ASHINSKY, D., AWAD, A., AZORR, M., BAILEY, D., BAKER, C., BARTILUCCI, D., BAYS, H., BEDEL, G., BELLINGER, R., BODE, B., BREakey, R., BRUSCO, O., CADE, J., CALL, R., CHANDLER, J., CHAYKIN, L., CHEUNG, D., CRABTREE, Y., CROCKETT, S., CUNNAR, J., DOW, H. J., DOWNEY, J., DREHOBL, M., DUNN, L., ERMAN, M., FITZ-PATRICK, D., FARMER, F. H., FERRERA, D. R., FRASER, N., FUSCO, F., GARBER, A., GILL, S., GILMAN, R., GILLIAM, J., GOLLAPUDI, G., HASSMAN, D., HUTCHINS, R., JAIN, R., KAIRI, R., KAYE, W., KAWLEY, A. F., KAYOTA, S., KERZNER, B., KHRONUSOVA, Y., KLONOFF, D., KUMAR, M., LALA, V., LANDGARTEN, S., LANE, W., LANG, J., LANG, R., LEICHTER, S., LEWIN, A., LIPETZ, R., LITTLEJOHN, T., LYNN, L., MADSEN, D., MAGGIACOMO, F., MALETZ, L., MALKIN, S., MARPLE, R., MCNEIL, R., MESSINA, N., MEYERS, P., MILLS, R., MORAWSKI, E., MULLEN, J., O'BARR, T., PEARLSTEIN, R., PERLMAN, N., PETERSON, G., POLLOCK, J., POPEIL, L., PRAGALOS, A., RACE, J., RADIN, D., RAVAL, P., RASKIN, P., RATNER, R., REYNOLDS, A., RIVELAND, B., COLON, L. R., SCHWARTZ, S., SHUSMAN, R., SILKINER, D., SMITH, W., SNELL, P., SOLER, N., SNYDER, B., STRINGAM, S., SUGIMOTO, D., TAMAYO, R., et al. Twice-daily dosing of a repaglinide/metformin fixed-dose combination tablet provides glycaemic control comparable to rosiglitazone/metformin tablet. <i>Diabetes, Obesity and Metabolism</i>.11 (9) (pp 865-873), 2009.Date of Publication: 2009.</p>	Keine Vergleichstherapie von Interesse

	Im Volltext ausgeschlossene Publikation	Grund
19	GAO, Y., YOON, K. H., CHUANG, L. M., MOHAN, V., NING, G., SHAH, S., JANG, H. C., WU, T. J., JOHNS, D., NORTHRUP, J. & BRODOWS, R. Efficacy and safety of exenatide in patients of Asian descent with type 2 diabetes inadequately controlled with metformin or metformin and a sulphonylurea. <i>Diabetes Research and Clinical Practice</i> .83 (1) (pp 69-76), 2009.Date of Publication: January 2009.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
20	ABE, M., OKADA, K., KIKUCHI, F. & MATSUMOTO, K. Clinical investigation of the effects of pioglitazone on the improvement of insulin resistance and blood pressure in type 2-diabetic patients undergoing hemodialysis. <i>Clinical Nephrology</i> .70 (3) (pp 220-228), 2008.Date of Publication: September 2008.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
21	BARNETT, A. H. Treatment intensification in patients with type 2 diabetes. <i>Diabetes, Obesity and Metabolism, Supplement</i> .10 (1) (pp iii), 2008.Date of Publication: June 2008.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
22	FERNANDEZ, M., TRIPLITT, C., WAJCBERG, E., SRIWIJILKAMOL, A. A., MUSI, N., CUSI, K., DEFRONZO, R. & CERSOSIMO, E. Addition of pioglitazone and ramipril to intensive insulin therapy in type 2 diabetic patients improves vascular dysfunction by different mechanisms. <i>Diabetes care</i> .31 (1) (pp 121-127), 2008.Date of Publication: Jan 2008.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)

	Im Volltext ausgeschlossene Publikation	Grund
23	MUSCH, A. Diabetes mellitus type 2: Combination with sitagliptin in moderately severe disease. [German]. Krankenhauspharmazie.29 (10) (pp 461-464), 2008.Date of Publication: October 2008.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
24	AYAD, N. M., FATTAH, W. A. & BAKRY, H. Biphasic insulin aspart 30 treatment improves glycemic control among patients with type 2 diabetes in Saudi Arabia and the Gulf region. Saudi Medical Journal.29 (6) (pp 919-920), 2008.Date of Publication: 2008.	Kein Volltext verfügbar
25	SERRA, D., HE, Y. L., BULLOCK, J., RIVIERE, G. J., BALEZ, S., SCHWARTZ, S., WANG, Y., LIGUEROS-SAYLAN, M., JARUGULA, V. & DOLE, W. P. Evaluation of pharmacokinetic and pharmacodynamic interaction between the dipeptidyl peptidase IV inhibitor vildagliptin, glyburide and pioglitazone in patients with type 2 diabetes. International Journal of Clinical Pharmacology and Therapeutics.46 (7) (pp 349-364), 2008.Date of Publication: July 2008.	Studiendauer <12 Wochen
26	LUND, S. S., TARNOW, L., FRANDBSEN, M., SMIDT, U. M., PEDERSEN, O., PARVING, H. H. & VAAG, A. A. Impact of metformin versus the prandial insulin secretagogue, repaglinide, on fasting and postprandial glucose and lipid responses in non-obese patients with type 2 diabetes. European Journal of Endocrinology.158 (1) (pp 35-46), 2008.Date of Publication: Jan 2008.	Studie beinhaltet nur Vergleiche zwischen Monotherapien mit Metformin oder einem anderen relevanten Komparator/Placebo
27	DAVIDSON, M., MEYER, P. M., HAFFNER, S., FEINSTEIN, S., D'AGOSTINO, SR., KONDOS, G. T., PEREZ, A., CHEN, Z. & MAZZONE, T. Increased high-density lipoprotein cholesterol predicts the pioglitazone-mediated reduction of carotid intima-media thickness progression in patients with type 2 diabetes mellitus. Circulation.117 (16) (pp 2123-2130), 2008.Date of Publication: Apr 2008.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält

	Im Volltext ausgeschlossene Publikation	Grund
28	LUND, S. S., TARNOW, L., FRANDBSEN, M., SMIDT, U. M., PEDERSEN, O., PARVING, H. H. & VAAG, A. A. Impact of metformin versus the prandial insulin secretagogue, repaglinide, on fasting and postprandial glucose and lipid responses in non-obese patients with type 2 diabetes. <i>European Journal of Endocrinology</i> .158 (1) (pp 35-46), 2008.Date of Publication: Jan 2008.	Doppelpublikation
29	NILSSON, P. M. 2007. Need of tight blood pressure control in type 2 diabetes. The intervention study ADVANCE gives clear information. <i>Lakartidningen</i> .104 (42) (pp 3028-3029), 2007.Date of Publication: 23 Oct 2007.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
30	No authors listed, Pioglitazone protects the type-2-diabetes patient from myocardial infarction and stroke. [German]. <i>MMW Fortschritte der Medizin</i> .149 (31-32) (pp 52-53), 2007.Date of Publication: 2 Aug 2007.	Kein Volltext verfügbar
31	No authors listed , Sitagliptin (Januvia) for type 2 diabetes. <i>The Medical letter on drugs and therapeutics</i> .49 (1251) (pp 1-3), 2007.Date of Publication: 1 Jan 2007.	Kein Volltext verfügbar
32	JACOB, A. N., SALINAS, K., ADAMS-HUET, B. & RASKIN, P. Weight gain in type 2 diabetes mellitus. <i>Diabetes, Obesity and Metabolism</i> .9 (3) (pp 386-393), 2007.Date of Publication: May 2007.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
33	No authors listed, Pioglitazone/glimepiride (Duetact) for diabetes. <i>Medical Letter on Drugs and Therapeutics</i> .49 (1253) (pp 9-11), 2007.Date of Publication: 29 Jan 2007.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
34	YAMADA, S., WATANABE, M., FUNAE, O., ATSUMI, Y., SUZUKI, R., YAJIMA, K., NAKAMURA, Y., KAWAI, T., OIKAWA, Y. & SHIMADA, A. Effect of combination therapy of a rapid-acting insulin secretagogue (glinide) with premixed insulin in type 2 diabetes mellitus. <i>Internal Medicine</i> .46 (23) (pp 1893-1897), 2007.Date of Publication: 03 Dec 2007.	Keine Vergleichstherapie von Interesse
35	DEROSA, G., D'ANGELO, A., FOGARI, E., SALVADEO, S., GRAVINA, A., FERRARI, I. & CICERO, A. F. 2007. Effects of nateglinide and glibenclamide on prothrombotic factors in naive type 2 diabetic patients treated with metformin: a 1-year, double-blind, randomized clinical trial. <i>Internal Medicine</i> , 46, 1837-1846.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
36	ERDMANN, E., DORMANDY, J., WILCOX, R., MASSI-BENEDETTI, M. & CHARBONNEL, B. PROactive 07: Pioglitazone in the treatment of type 2 diabetes: Results of the PROactive study. <i>Vascular Health and Risk Management</i> .3 (4) (pp 355-370), 2007.Date of Publication: 2007.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
37	No authors listed, Diabetes mellitus: Metformin enhances incretin effect diabetes mellitus. [German]. <i>Deutsche Apotheker Zeitung</i> .147 (41) (pp 37-38), 2007.Date of Publication: 11 Oct 2007.	Kein Volltext verfügbar
38	MUSCH, A. & HEINZL, S. Sitagliptin - Dipeptidyl peptidase-4 inhibitor for patients with diabetes mellitus type 2. [German]. <i>Medizinische Monatsschrift für Pharmazeuten</i> .30 (8) (pp 282-286), 2007.Date of Publication: Aug 2007.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
39	WHITLEY, H. P. Sitagliptin (Januvia) for the treatment of patients with type 2 diabetes. <i>American Family Physician</i> .76 (6) (pp 861), 2007.Date of Publication: 15 Sep 2007.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
40	COMASCHI, M., DEMICHELI, A., DI, P. C., BELLATRECCIA, A., MARIZ, S., ALESSI, R., BALDINI, A., BEVILACQUA, M., BORZI, V., CABASINO, F., CALCATERRA, E., CARDUCCI, A., CERNIGOI, A., CHIAVETTA, A., CHIOVATO, L., CIONI, G., CIGNARELLI, M., CORDA, A., CORDERA, R., CORSI, L., DE, R. C., DI, P. S., DONADON, W., FELLIN, R., FRANCESCONI, A., FUGAZZA, L., GHIGO, E., LEOTTA, S., MAININI, E., MAOLO, G., MARCHESI, M., MARIN, N., MARINI, F., MASOTTI, G., MUSCOGIURI, A., ORIO, F., PROVENZANO, V., RICCARDI, G., ROGNONI, C., ROSSINI, S., SERRA, R., TATTI, P., TREVISAN, R., TUCCINARDI, F. & VENEZIA, A. R. Effects of pioglitazone in combination with metformin or a sulfonylurea compared to a fixed-dose combination of metformin and glibenclamide in patients with type 2 diabetes. <i>Diabetes Technology and Therapeutics</i> .9 (4) (pp 387-398), 2007.Date of Publication: 2007.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
41	RETNAKARAN, R. Does addition of vildagliptin to metformin monotherapy improve glycemic control in patients with type 2 diabetes mellitus? <i>Nature Clinical Practice Endocrinology and Metabolism</i> .3 (8) (pp 572-573), 2007.Date of Publication: Aug 2007.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
42	BERHANU, P., PEREZ, A., YU, S. & SPANHEIMER, R. Effect of pioglitazone in combination with insulin therapy on glycaemic control, insulin dose requirement and lipid profile in patients with type 2 diabetes previously poorly controlled with combination therapy. <i>Diabetes, Obesity and Metabolism</i> .9 (4) (pp 512-520), 2007.Date of Publication: Jul 2007.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)

	Im Volltext ausgeschlossene Publikation	Grund
43	SALSALI, A. & PRATLEY, R. E. Does addition of sitagliptin to metformin monotherapy improve glyceic control in patients with type 2 diabetes mellitus? <i>Nature Clinical Practice Endocrinology and Metabolism</i> .3 (6) (pp 450-451), 2007.Date of Publication: Jun 2007.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
44	SCHWARTZ, A. V. & SELLMAYER, D. E. Thiazolidinediones: New evidence of bone loss. <i>Journal of Clinical Endocrinology and Metabolism</i> .92 (4) (pp 1232-1234), 2007.Date of Publication: Apr 2007.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
45	GLENDENNING, C. & KAUFMANN, L. Glycemic control in patients with type 2 diabetes. <i>American Family Physician</i> .75 (7) (pp 1051-1053), 2007.Date of Publication: 01 Apr 2007.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
46	GASTALDELLI, A., CASOLARO, A., PETTITI, M., NANNIPIERI, M., CIOCIARO, D., FRASCERRA, S., BUZZIGOLI, E., BALDI, S., MARI, A. & FERRANNINI, E. Effect of pioglitazone on the metabolic and hormonal response to a mixed meal in type II diabetes. <i>Clinical Pharmacology and Therapeutics</i> .81 (2) (pp 205-212), 2007.Date of Publication: Feb 2007.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
47	FONSECA, V. A., THEUMA, P., MUDALIAR, S., LEISSINGER, C. A., CLEJAN, S. & HENRY, R. R. Diabetes treatments have differential effects on nontraditional cardiovascular risk factors. <i>Journal of Diabetes and its Complications</i> .20 (1) (pp 14-20), 2006.Date of Publication: January/February 2006.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)

	Im Volltext ausgeschlossene Publikation	Grund
48	KVAPIL, M., SWATKO, A., HILBERG, C. & SHESTAKOVA, M. Biphasic insulin aspart 30 plus metformin: An effective combination in type 2 diabetes. <i>Diabetes, Obesity and Metabolism</i> .8 (1) (pp 39-48), 2006.Date of Publication: January 2006.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
49	PRATLEY, R. E., JAUFFRET-KAMEL, S., GALBREATH, E. & HOLMES, D. Twelve-week monotherapy with the DPP-4 inhibitor vildagliptin improves glycemic control in subjects with type 2 diabetes. <i>Hormone and Metabolic Research</i> .38 (6) (pp 423-428), 2006.Date of Publication: June 2006.	Keine Vergleichstherapie von Interesse
50	MAZZONE, T., MEYER, P. M., FEINSTEIN, S. B., DAVIDSON, M. H., KONDOS, G. T., D'AGOSTINO, SR., PEREZ, A., PROVOST, J. C. & HAFFNER, S. M. Effect of pioglitazone compared with glimepiride on carotid intima-media thickness in type 2 diabetes: A randomized trial. <i>Journal of the American Medical Association</i> .296 (21) (pp 2572-2581), 2006.Date of Publication: 06 Dec 2006.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält

	Im Volltext ausgeschlossene Publikation	Grund
51	KABADI, U. M. & KABADI, M. Comparative efficacy of glimepiride and/or metformin with insulin in type 2 diabetes. <i>Diabetes Research and Clinical Practice</i> .72 (3) (pp 265-270), 2006.Date of Publication: Jun 2006.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
52	BARRANCO, C. Early use of combination therapy improves outcome of patients with type 2 diabetes. <i>Nature Clinical Practice Endocrinology and Metabolism</i> .2 (4) (pp 186-187), 2006.Date of Publication: Apr 2006.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
53	MD ISA, S. H., NAJIHAH, I., NAZAIMOON, W. M. W., KAMARUDIN, N. A., UMAR, N. A., MAT, N. H. & KHALID, B. A. K. Improvement in C-reactive protein and advanced glycosylation end-products in poorly controlled diabetics is independent of glucose control. <i>Diabetes Research and Clinical Practice</i> .72 (1) (pp 48-52), 2006.Date of Publication: Apr 2006.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
54	No authors listed, Type 2 diabetic patient with reduced beta-cell function. New substance class: incretin-mimetics. [German]. <i>MMW Fortschritte der Medizin</i> .147 (31-32) (pp 55), 2005.Date of Publication: 4 Aug 2005.	Kein Volltext verfügbar
55	EBELL, M. H. Insulin monotherapy vs. combination therapy. <i>American family physician</i> .71 (5) (pp 899), 2005.Date of Publication: 1 Mar 2005.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)

	Im Volltext ausgeschlossene Publikation	Grund
56	RITZ, E. Edema and congestive heart failure from thiazolidone insulin sensitizers - Excess sodium reabsorption in the collecting duct. Journal of the American Society of Nephrology.16 (11) (pp 3139-3142), 2005.Date of Publication: 2005.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
57	FEINGLOS, M. N., SAAD, M. F., PI-SUNYER, F. X., AN, B. & SANTIAGO, O. Effects of liraglutide (NN2211), a long-acting GLP-1 analogue, on glycaemic control and bodyweight in subjects with Type 2 diabetes. Diabetic Medicine.22 (8) (pp 1016-1023), 2005.Date of Publication: Aug 2005.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
58	KRAWCZYK, A., GAJER, G., GRZESZCZAK, W. & STROJEK, K. The impact of metformin on metabolic control in obese type 2 diabetic patients treated with insulin. Preliminary observation. [Polish]. Diabetologia Doswiadczalna i Kliniczna.5 (1) (pp 47-51), 2005.Date of Publication: 2005.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
59	No authors listed, Long-term HbA1c stable. Insulin sensitizer keeps diabetes in check. [German]. MMW Fortschritte der Medizin.146 (13) (pp 60), 2004.Date of Publication: 25 Mar 2004.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
60	GENSTHALER, B. M. Better diabetes control with BOT. [German]. Pharmazeutische Zeitung.149 (38) (pp 32), 2004.Date of Publication: 16 Sep 2004.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
61	TAN, M., JOHNS, D., GALVEZ, G. G., ANTUNEZ, O., FABIAN, G., FLORES-LOZANO, F., GUAJARDO, S. Z., GARZA, E., MORALES, H., KONKOY, C. & HERZ, M. Effects of pioglitazone and glimepiride on glycemic control and insulin sensitivity in Mexican patients with type 2 diabetes mellitus: A multicenter, randomized, double-blind, parallel-group trial. Clinical Therapeutics.26 (5) (pp 680-693), 2004.Date of Publication: May 2004.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
62	TOSI, F., MUGGEO, M., BRUN, E., SPIAZZI, G., PEROBELLI, L., ZANOLIN, E., GORI, M., COPPINI, A. & MOGHETTI, P. Combination treatment with metformin and glibenclamide versus single-drug therapies in type 2 diabetes mellitus: A randomized, double-blind, comparative study. Metabolism: Clinical and Experimental.52 (7) (pp 862-867), 2003.Date of Publication: 01 Jul 2003.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)

	Im Volltext ausgeschlossene Publikation	Grund
63	WULFFELE, M. G., KOOY, A., LEHERT, P., BETS, D., OGTROP, J. C., BORGER, V. D. B., DONKER, A. J. & STEHOUSER, C. D. Combination of insulin and metformin in the treatment of type 2 diabetes. <i>Diabetes care</i> .25 (12) (pp 2133-2140), 2002.Date of Publication: Dec 2002.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
64	ROSAK, C., HAUPT, E., WALTER, T. & WERNER, J. The effect of combination treatment with acarbose and glibenclamide on postprandial glucose and insulin profiles: Additive blood glucose lowering effect and decreased hypoglycaemia. <i>Diabetes, Nutrition and Metabolism - Clinical and Experimental</i> .15 (3) (pp 143-151), 2002.Date of Publication: 2002.	Studiendauer <12 Wochen
65	TONG, P. C., CHOW, C. C., JORGENSEN, L. N. & COCKRAM, C. S. The contribution of metformin to glycaemic control in patients with Type 2 diabetes mellitus receiving combination therapy with insulin. <i>Diabetes Research and Clinical Practice</i> .57 (2) (pp 93-98), 2002.Date of Publication: 2002.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
66	LACKNAUTH, S., PATEL, R., RAPPA, L. & BRANCH, I. E. Metformin/glyburide combination therapy for 2 type diabetes. <i>P and T</i> .26 (5) (pp 268-272), 2001.Date of Publication: 2001.	Kein Volltext verfügbar
67	PONSSEN, H. H., ELTE, J. W. F., LEHERT, P., SCHOUTEN, J. P. & BETS, D. Combined metformin and insulin therapy for patients with type 2 diabetes mellitus. <i>Clinical Therapeutics</i> .22 (6) (pp 709-718), 2000.Date of Publication: 2000.	berichtet keinen Ausgangswert und keinen Folgewert für mind. einen Endpunkt

	Im Volltext ausgeschlossene Publikation	Grund
68	HERMANN, L. S., RANSTAM, J., VAALER, S. & MELANDER, A. Effects of antihyperglycaemic therapies on proinsulin and relation between proinsulin and cardiovascular risk factors in type 2 diabetes. <i>Diabetes, obesity & metabolism</i> .1 (4) (pp 227-232), 1999.Date of Publication: Jul 1999.	Studie beinhaltet nur Vergleiche zwischen Monotherapien mit Metformin oder einem anderen relevanten Komparator/Placebo
69	RITZMANN, P. Controversial aspects regarding metformin in type 2 diabetic patient: reduction of late complications in early, single drug administration; increased risk of death in later administration in combination with sulfonylurea compounds. [German]. <i>Schweizerische Rundschau für Medizin Praxis = Revue suisse de médecine Praxis</i> .88 (33) (pp 1330-1331), 1999.Date of Publication: 12 Aug 1999., 12	Kein Volltext verfügbar
70	LANDSTEDT-HALLIN, L., ARNER, P., LINS, P. E., BOLINDER, J., OLSEN, H. & GROOP, L. 1999. The role of sulphonylurea in combination therapy assessed in a trial of sulphonylurea withdrawal. Scandinavian Insulin-Sulphonylurea Study Group Research Team. <i>Diabetic Medicine</i> , 16, 827-834.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
71	LINDSTROM, T., NYSTROM, F. H., OLSSON, A. G., OTTOSSON, A. M. & ARNQVIST, H. J. The lipoprotein profile differs during insulin treatment alone and combination therapy with insulin and sulphonylureas in patients with type 2 diabetes mellitus. <i>Diabetic Medicine</i> .16 (10) (pp 820-826), 1999.Date of Publication: 1999.	Berichtet keinen Ausgangswert und keinen Folgewert für mind. einen Endpunkt
72	BANDO, Y., USHIOGI, Y., OKAFUJI, K., TOYA, D., TANAKA, N. & FUJISAWA, M. Troglitazone combination therapy in obese type 2 diabetic patients poorly controlled with alpha-glucosidase inhibitors. <i>Journal of International Medical Research</i> .27 (2) (pp 53-64), 1999.Date of Publication: 1999.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
73	MAKIMATTILA, S., NIKKILA, K. & YKI-JARVINEN, H. Causes of weight gain during insulin therapy with and without metformin in patients with Type II diabetes mellitus. <i>Diabetologia</i> .42 (4) (pp 406-412), 1999.Date of Publication: 1999.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
74	DE FINE, O. N. & ANDREASEN, A. H. The UK Prospective Diabetes Study. <i>Lancet</i> .352 (9144) (pp 1933; author reply 1934), 1998.Date of Publication: 12 Dec 1998., 12.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
75	ROBINSON, A. C., BURKE, J., ROBINSON, S., JOHNSTON, D. G. & ELKELES, R. S. The effects of metformin on glycemic control and serum lipids in insulin-treated NIDDM patients with suboptimal metabolic control. <i>Diabetes Care</i> .21 (5) (pp 701-705), 1998.Date of Publication: 1998.	Berichtet keinen Ausgangswert und keinen Folgewert für mind. einen Endpunkt
76	CEFALU, W. T., BELL-FARROW, A., WANG, Z. Q., MCBRIDE, D., DALGLEISH, D. & TERRY, J. G. Effect of glipizide GITS on insulin sensitivity, glycemic indices, and abdominal fat composition in NIDDM. <i>Drug Development Research</i> .44 (1) (pp 1-7), 1998.Date of Publication: May 1998.	Studiendauer <12 Wochen
77	BARDINI, G., MANNUCCI, E. & ROTELLA, C. M. Effect of different pharmacological formulations of gliclazide on postprandial hyperglycaemia [1]. <i>Diabetic Medicine</i> .15 (8) (pp 706), 1998.Date of Publication: 1998.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
78	FEINGLOS, M. N., THACKER, C. R., LOBAUGH, B., DEATKINE, D. D., MCNEILL, D. B., ENGLISH, J. S. & BURSEY, D. L. Combination insulin and sulfonylurea therapy in insulin-requiring type 2 diabetes mellitus. <i>Diabetes Research and Clinical Practice</i> .39 (3) (pp 193-199), 1998.Date of Publication: Mar 1998.	Berichtet keinen Ausgangswert und keinen Folgewert für mind. einen Endpunkt

	Im Volltext ausgeschlossene Publikation	Grund
79	KOVACEVIC, I., PROFOZIC, V., SKRABALO, Z., CABRIJAN, T., ZJACIC-ROTKVIC, V., GOLDONI, V., JOVIC-PASKVALIN, L., CRNCEVIC-ORLIC, Z., KOSELJ, M. & METELKO, Z. Multicentric clinical trial to assess efficacy and tolerability of acarbose (BAY G 5421) in comparison to glibenclamide and placebo. Diabetologia Croatica.26 (2) (pp 83-89), 1997.Date of Publication: 1997.	Kein Volltext verfügbar
80	NAKATOMI, I., EBERT, C. D., HEIBER, S. J., GUTNIAK, M. K., LARSSON, H., AHREN, B., JUNESKANS, O. T. & HOLST, J. J. Therapeutic levels of glucagon-like peptide-1 achieved in humans by an oral transmucosal tablet. Proceedings of the Controlled Release Society.(23) (pp 839-840), 1996.Date of Publication: 1996.	Kein Volltext verfügbar
81	MOORADIAN, A. D., ALBERT, S. G., BERNBAUM, M. & PLUMMER, S. The effect of glipizide gastrointestinal therapeutic system on islet cell hormonal responses to a test meal in NIDDM. Diabetes Care.19 (8) (pp 883-884), 1996.Date of Publication: Aug 1996.	Studiendauer <12 Wochen
82	GOLAY, A., GUILLET-DAUPHINE, N., FENDEL, A., JUGE, C. & ASSAL, J. P. The insulin-sparing effect of metformin in insulin-treated diabetic patients. Diabetes/Metabolism Reviews.11 (SUPPL, 1), 1995.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
83	ZARGAR, A. H., SHAH, N. A. & LAWAY, B. A. Effect of sulphonylurea drugs (Tolbutamine/Glibenclamide) on platelet aggregation in type-II diabetes mellitus. Journal of the Diabetic Association of India.34 (3-4) (pp 55-58), 1994.Date of Publication: 1994.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
84	CAMERINI-DAVALOS, R. A., BLOODWORTH, JR., VELASCO, C. A. & REDDI, A. S. Effect of insulin-glipizide combination on skeletal muscle capillary basement membrane width in diabetic patients. <i>Clinical Therapeutics</i> .16 (6) (pp 952-961), 1994.Date of Publication: 1994.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
85	HERMANN, L. S., SCHERSTEN, B. & MELANDER, A. Antihyperglycaemic efficacy, response prediction and Dose-Response relations of treatment with metformin and sulphonylurea, alone and in primary combination. <i>Diabetic Medicine</i> .11 (10) (pp 953-960), 1994.Date of Publication: 1994.	Studiendauer <12 Wochen
86	OFFMANN, J. & SPENGLER, M. Efficacy of 24-week monotherapy with acarbose, metformin, or placebo in dietary-treated NIDDM patients: The essen-II study. <i>American Journal of Medicine</i> .103 (6) (pp 483-490), 1997.Date of Publication: Dec 1997.	Kein Volltext verfügbar
87	MEZITIS, N. H. E., HESHKA, S., SAITAS, V., BAILEY, T. S., COSTA, R. & PI-SUNYER, F. X. Combination therapy for NIDDM with biosynthetic human insulin and glyburide. <i>Diabetes Care</i> .15 (2) (pp 265-269), 1992.Date of Publication: 1992.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
88	MOHAN, V., SNEHALATHA, C., RAMACHANDRAN, A. & VISWANATHAN, M. Combination therapy of glibenclamide and insulin in NIDDM patients with secondary failure to oral drugs. <i>The Journal of the Association of Physicians of India</i> .38 (8) (pp 537-541), 1990.Date of Publication: Aug 1990.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
89	RIDDLE, M. C., HART, J. S., BOUMA, D. J., PHILLIPSON, B. E. & YOUKER, G. Efficacy of bedtime NPH insulin with daytime sulfonylurea for subpopulation of type II diabetic subjects. <i>Diabetes Care</i> .12 (9) (pp 623-629), 1989.Date of Publication: 1989.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
90	RAMACHANDRAN, A., MOHAN, V., SNEHALATHA, M. C., CHINNIKRISHNUDU, M. & VISWANATHAN, M. Use of metformin in selected cases of NIDDM. A clinical trial. <i>Journal of the Diabetic Association of India</i> .28 (4) (pp 137-139), 1988.Date of Publication: 1988.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
91	LAWRENCE, A. M. & ABRAIRA, C. New modalities in diabetes treatment. <i>American Journal of Medicine</i> .85 (5 A) (pp 153-158), 1988.Date of Publication: 1988.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)

	Im Volltext ausgeschlossene Publikation	Grund
92	STENMAN, S., GROOP, P. H., SALORANTA, C., TOTTERMAN, K. J., FYHRQVIST, F. & GROOP, L. Effects of the combination of insulin and glibenclamide in Type 2 (non-insulin-dependent) diabetic patients with secondary failure to oral hypoglycaemic agents. <i>Diabetologia</i> .31 (4) (pp 206-213), 1988.Date of Publication: 1988.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
93	LARKINS, R. G., JERUMS, G., TAFT, J. L., GODFREY, H., SMITH, I. L. & MARTIN, T. J. Lack of the effect of gliclazide on the platelet aggregation in insulin-treated and non-insulin-treated diabetes: A two-year controlled study. <i>Diabetes Research and Clinical Practice</i> .4 (2) (pp 81-87), 1987.Date of Publication: 1987.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
94	SCHADE, D. S., MITCHELL, W. J. & GRIEGO, G. Addition of sulfonylurea to insulin treatment in poorly controlled type II diabetes. A double-blind, randomized clinical trial. <i>Journal of the American Medical Association</i> .257 (18) (pp 2441-2445), 1987.Date of Publication: 1987.	Berichtet keinen Ausgangswert und keinen Folgewert für mind. einen Endpunkt

	Im Volltext ausgeschlossene Publikation	Grund
95	JERUMS, G., MURRAY, R. M. L. & SEEMAN, E. Lack of effect of gliclazide on early diabetic nephropathy and retinopathy: A two-year controlled study. <i>Diabetes Research and Clinical Practice</i> .3 (2) (pp 71-80), 1987.Date of Publication: 1987.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
96	MAUERHOFF, T., KETELSLEGERS, J. M. & LAMBERT, A. E. Effect of glibenclamide in insulin-treated diabetic patients with a residual insulin secretion. <i>Diabete et Metabolisme</i> .12 (1) (pp 34-38), 1986.Date of Publication: 1986.	Kein Volltext verfügbar
97	HAPP, J., BEYER, J., NEST, E., FROHLICH, A., ALTHOFF, P. H. & SCHOFFLING, K. Blood sugar, serum insulin, nonesterified fatty acids and somatotropin in daily profile in adult diabetics under monotherapy with various sulfonylureas. [German]. <i>Arzneimittel-Forschung</i> .24 (8) (pp 1228-1234), 1974.Date of Publication: Aug 1974.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
98	VAILATI, G., PAGANI, G., MONTINI, M., CAPUTO, G. & SACCHETTI, G. Effect of glipizide on a double-meal test in diabetic patients. <i>Arzneimittel-Forschung</i> .22 (11) (pp 1888-1889), 1972.Date of Publication: Nov 1972.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
99	JAILLARD, J. & WAREMBOURG, H. 1972. Therapeutic trial of a new antidiabetic with vascular action: 1702 SE. [French]. <i>Lille medical : journal de la Faculte de medecine et de pharmacie de l'Universite de Lille</i> .17 (9) (pp Suppl 6:1332-1336), 1972.Date of Publication: Oct 1972., 1332-1336.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
100	HU, Y. Y., YE, S. D., ZHAO, L. L., ZHENG, M. & CHEN, Y. 2010. Hydrochloride pioglitazone decreases urinary TGF-beta1 excretion in type 2 diabetics. European Journal of Clinical Investigation, 40, 571-574.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
101	NAKA, K. K., PAPPAS, K., PAPATHANASSIOU, K., PAPAMICHAEL, N. D., KAZAKOS, N., KANIOGLOU, C., MAKRIYIANNIS, D., KATSOURAS, C. S., LIVERIS, K., TSATSOULIS, A. & MICHALIS, L. K. Lack of effects of pioglitazone on cardiac function in patients with type 2 diabetes and evidence of left ventricular diastolic dysfunction: A tissue doppler imaging study. Cardiovascular Diabetology.9 , 2010.Article Number: 57.Date of Publication: 23 Sep 2010., 57.	Doppelpublikation
102	ONUCHIN, S. G., ELSUKOVA, O. S., SOLOV'EV, O. V. & ONUCHINA, E. L. 2010. [Capabilities of hypoglycemic therapy in women with decompensated type 2 diabetes mellitus]. [Russian]. Terapevticheskii Arkhiv, 82, 34-41.	Berichtet keinen Ausgangswert und keinen Folgewert für mind. einen Endpunkt
103	BLONDE, L., DAGOGO-JACK, S., BANERJI, M. A., PRATLEY, R. E., MARCELLARI, A., BRACERAS, R., PURKAYASTHA, D. & BARON, M. 2009. Comparison of vildagliptin and thiazolidinedione as add-on therapy in patients inadequately controlled with metformin: results of the GALIANT trial--a primary care, type 2 diabetes study. Diabetes, Obesity & Metabolism, 11, 978-986.	Falsche Dosierung
104	LUND, S. S., TARNOW, L., FRANDBSEN, M., NIELSEN, B. B., HANSEN, B. V., PEDERSEN, O., PARVING, H. H. & VAAG, A. A. Combining insulin with metformin or an insulin secretagogue in non-obese patients with type 2 diabetes: 12 Month, randomised, double blind trial. BMJ.339 (7730) (pp 1121-1124), 2009.Date of Publication: 2009.	Doppelpublikation

	Im Volltext ausgeschlossene Publikation	Grund
105	PRADHAN, A. D., EVERETT, B. M., COOK, N. R., RIFAI, N. & RIDKER, P. M. 2009. Effects of initiating insulin and metformin on glycemic control and inflammatory biomarkers among patients with type 2 diabetes: the LANCET randomized trial. JAMA, 302, 1186-1194.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
106	DEROSA, G., MAFFIOLI, P., SALVADEO, S. A., FERRARI, I., GRAVINA, A., MEREU, R., PALUMBO, I., D'ANGELO, A. & CICERO, A. F. 2009. Direct comparison among oral hypoglycemic agents and their association with insulin resistance evaluated by euglycemic hyperinsulinemic clamp: the 60's study. Metabolism: Clinical & Experimental, 58, 1059-1066.	Studie beinhaltet nur Vergleiche zwischen Monotherapien mit Metformin oder einem anderen relevanten Komparator/Placebo
107	KOOY, A., DE, J. J., LEHERT, P., BETS, D., WULFFELE, M. G., DONKER, A. J. & STEHOUWER, C. D. 2009. Long-term effects of metformin on metabolism and microvascular and macrovascular disease in patients with type 2 diabetes mellitus. Archives of Internal Medicine, 169, 616-625.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
108	DEROSA, G., MEREU, R., SALVADEO, S. A., D'ANGELO, A., CICCARELLI, L., PICCINNI, M. N., FERRARI, I., GRAVINA, A., MAFFIOLI, P. & CICERO, A. F. 2009. Pioglitazone metabolic effect in metformin-intolerant obese patients treated with sibutramine. Internal Medicine, 48, 265-271.	Studie beinhaltet nur Vergleiche zwischen Monotherapien mit Metformin oder einem anderen relevanten Komparator/Placebo

	Im Volltext ausgeschlossene Publikation	Grund
109	DEROSA, G., D'ANGELO, A., FOGARI, E., SALVADEO, S., GRAVINA, A., FERRARI, I. & CICERO, A. F. 2009. Nateglinide and glibenclamide metabolic effects in naive type 2 diabetic patients treated with metformin. <i>Journal of Clinical Pharmacy & Therapeutics</i> , 34, 13-23.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
110	No authors listed, 2008. Vildagliptin and vildagliptin/metformin: new drug. Same limited efficacy as sitagliptin in type 2 diabetes. <i>Prescrire international</i> , 17, 188.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
111	FONSECA, V., BARON, M., SHAO, Q. & DEJAGER, S. 2008. Sustained efficacy and reduced hypoglycemia during one year of treatment with vildagliptin added to insulin in patients with type 2 diabetes mellitus. <i>Hormone & Metabolic Research</i> , 40, 427-430.	Keine Vergleichstherapie von Interesse
112	AHREN, B., PACINI, G., TURA, A., FOLEY, J. E. & SCHWEIZER, A. 2007. Improved meal-related insulin processing contributes to the enhancement of B-cell function by the DPP-4 inhibitor vildagliptin in patients with type 2 diabetes. <i>Hormone & Metabolic Research</i> , 39, 826-829.	Keine Vergleichstherapie von Interesse
113	No authors listed, 2007. [Combination therapy with oral antidiabetics]. [German]. <i>MMW Fortschritte der Medizin</i> , 149, 52-53.	Kein Volltext verfügbar
114	No authors listed, 2007. [New class of oral antidiabetic drugs. Effective in combination with metformin]. [German]. <i>MMW Fortschritte der Medizin</i> , 149, 17.	Kein Volltext verfügbar
115	No authors listed, 2007. Sitagliptin/metformin (Janumet) for type 2 diabetes. <i>Medical Letter on Drugs & Therapeutics</i> , 49, 45-47.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)

	Im Volltext ausgeschlossene Publikation	Grund
116	BERMUDEZ-PIRELA, V. J., CANO, C., MEDINA, M. T., SOUKI, A., LEMUS, M. A., LEAL, E. M., SEYFI, H. A., CANO, R., CISCEK, A., BERMUDEZ-ARIAS, F., CONTRERAS, F., ISRAILI, Z. H., HERNANDEZ-HERNANDEZ, R. & VALASCO, M. 2007. Metformin plus low-dose glimeperide significantly improves Homeostasis Model Assessment for insulin resistance (HOMA(IR)) and beta-cell function (HOMA(beta-cell)) without hyperinsulinemia in patients with type 2 diabetes mellitus. <i>American Journal of Therapeutics</i> , 14, 194-202.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
117	WILCOX, R., BOUSSER, M. G., BETTERIDGE, D. J., SCHERNTHANER, G., PIRAGS, V., KUPFER, S., DORMANDY, J. & PROACTIVE, I. 2007. Effects of pioglitazone in patients with type 2 diabetes with or without previous stroke: results from PROactive (PROspective pioglitAzone Clinical Trial In macroVascular Events 04). <i>Stroke</i> , 38, 865-873.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
118	No authors listed, 2007. Pioglitazone/glimeperide (Duetact) for diabetes. <i>Medical Letter on Drugs & Therapeutics</i> , 49, 9-11.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
119	CHALMERS, J., PERKOVIC, V., JOSHI, R. & PATEL, A. 2006. ADVANCE: breaking new ground in type 2 diabetes. <i>Journal of Hypertension - Supplement</i> , 24, S22-S28.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)

	Im Volltext ausgeschlossene Publikation	Grund
120	YU, D., MURDOCH, S. J., PARIKH, S. J., MARCOVINA, S. M., COBITZ, A., CHEN, H. & BRUNZELL, J. D. 2006. Rosiglitazone increases LDL particle size and buoyancy and decreases C-reactive protein in patients with type 2 diabetes on statin therapy. <i>Diabetes & Vascular Disease Research</i> , 3, 189-196.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
121	No authors listed, 2006. [Therapy with incretin mimetic exenatide. Soon HbA1c value and weight decrease permanently]. [German]. <i>MMW Fortschritte der Medizin</i> , 148, 55.	Kein Volltext verfügbar
122	GÖKE, B. 2006. [Improved blood sugar control plus weight loss]. [German]. <i>MMW Fortschritte der Medizin</i> , 148, 51.	Kein Volltext verfügbar
123	No authors listed, 2006. Pioglitazone/metformin (Actoplus met). <i>Medical Letter on Drugs & Therapeutics</i> , 48, 9-11	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
124	TSENG, C. H. & HUANG, T. S. 2005. Pioglitazone with sulfonylurea: glycemic and lipid effects in Taiwanese type 2 diabetic patients. <i>Diabetes Research & Clinical Practice</i> , 70, 193-194.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
125	No authors listed, 2004. [2-year data of large clinical comparative studies. Type 2 diabetes: lasting metabolic control with pioglitazone]. [German]. <i>MMW Fortschritte der Medizin</i> , 146, 48-49.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
126	No authors listed, 2003. [Blood sugar inspite of maximal metformin dosage. New fixed combination forces HbA1c down]. [German]. MMW Fortschritte der Medizin, 145, 54.	Kein Volltext verfügbar
127	No authors listed, 2002. [Effective reduction of the HbA1c level, fewer side effects. First fixed combination for type 2 diabetic patients]. [German]. MMW Fortschritte der Medizin, 144, 49.	Kein Volltext verfügbar
128	ARONOFF, S., ROSENBLATT, S., BRAITHWAITE, S., EGAN, J. W., MATHISEN, A. L. & SCHNEIDER, R. L. 2000. Pioglitazone hydrochloride monotherapy improves glycemic control in the treatment of patients with type 2 diabetes: a 6-month randomized placebo-controlled dose-response study. The Pioglitazone 001 Study Group. Diabetes Care, 23, 1605-1611.	Doppelpublikation
129	NOZUE, T., MICHISHITA, I., MINAGAWA, F. & GENDA, A. 1999. Troglitazone directly increases HDL cholesterol levels. Diabetes Care, 22, 355-356.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
130	LANDSTEDT-HALLIN, L., ARNER, P., LINS, P. E., BOLINDER, J., OLSEN, H. & GROOP, L. 1999. The role of sulphonylurea in combination therapy assessed in a trial of sulphonylurea withdrawal. Scandinavian Insulin-Sulphonylurea Study Group Research Team. Diabetic Medicine, 16, 827-834.	Doppelpublikation
131	No authors listed, 1995. Metformin for non-insulin-dependent diabetes mellitus. Medical Letter on Drugs & Therapeutics, 37, 41-42.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
132	LINS, P. E., LUNDBLAD, S., PERSSON-TROTZIG, E. & ADAMSON, U. 1988. Glibenclamide improves the response to insulin treatment in non-insulin-dependent diabetics with second failure to sulfonylurea therapy. Acta Medica Scandinavica, 223, 171-179.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)

	Im Volltext ausgeschlossene Publikation	Grund
133	ZHENG, G. F., WANG, J. P., ZHANG, H., HU, Z. X., LIU, J., XIAO, J. Z., CHEN, S. M., CAO, H. B., LI, G. W., HU, Y. H. & PAN, X. R. 1995. Clinical observation on glucobay treatment for NIDDM. Chinese Journal of Endocrinology, 11, 163-164.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
134	WANG, W. L., ZHANG, B. Z., ZHANG, S. L., XU, D. F. & WEI, X. H. 1994. The clinical evaluation of domestic gliclazide tablets in patients with NIDDM. Chinese Journal of Pharmacoepidemiology, 3, 119-121.	Kein Volltext verfügbar
135	MCALPINE, L. G., MCALPINE, C. H., WACLAWSKI, E. R., STORER, A. M., KAY, J. W. & FRIER, B. M. 1988. A comparison of treatment with metformin and gliclazide in patients with non-insulin-dependent diabetes. European Journal of Clinical Pharmacology, 34, 129-132.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
136	LU, S. C., ZHU, Y. D., SHEN, J. J. & DU, X. D. 2000. Efficacy observations of repaglinide in the treatment of type 2 diabetes among 68 cases. Modern Applied Pharmacy, 17, 236-238.	Kein Volltext verfügbar
137	XU, W. C., CHEN, C. R. & CHEN, Y. S. 2001. Combination therapy with bedtime insulin and daytime oral hypoglycaemic agents in type 2 diabetic patients. Hebei Medicine, 23, 23-24.	Kein Volltext verfügbar
138	KAWAMORI, R., KINOSHITA, J., IKEDA, M., KUBOTA, M., WADA, M., KANDA, T., IKEBUCHI, M., TODO, R. & YAMASAKIY PS 28 Oral Therapy in NIDDM Pioglitazone ameliorates insulin resistance in NIDDM [abstract]. Diabetologia, Vol.40, pp.A 306, 1997.	Berichtet keinen Ausgangswert und keinen Folgewert für mind. einen Endpunkt

	Im Volltext ausgeschlossene Publikation	Grund
139	GREGORIO, F., AMBROSI, F., CAPODICASA, E., MANFRINI, S., TESTA, I. & FILIPPONI, P. 1994. Metformin-sulphonylurea combination therapy: the influence of regional adiposity on biguanide-induced haemostatic improvement. Balance in type II diabetic patients. Annals of Experimental and Clinical Medicine, Vol.2, pp.136-45, 1994., 45.	Kein Volltext verfügbar
140	ANON Implantiruemyi dozator insulina povyshaet kachestvo zhizni pri insulinnezavisimom sakharnom diabete. Mezhdunarodnyi Zhurnal Meditsinskoi Praktiki, 27p. 1997.	Keine RCT (gepoolte Analysen wurden ebenfalls ausgeschlossen)
141	YANG, J. K., DI, F. S., HE, R. H., ZHU, X. S., WANG, D. Q., YANG, M. G., WANG, Y. G., YANG, T., LIUXH & CHEN, J. W. 2002. Clinical study on rosiglitazone monotherapy of early type 2 diabetes. China Pharmacy, 13, 608-610.	Kein Volltext verfügbar
142	FU, L. P. & GU, W. 2003. The profile of serum free fatty acids in patients with type 2 diabetes and the intervention effect of pioglitazone. Oa1/2-O1/2N[S], 25, 259-262.	Kein Volltext verfügbar
143	WANG, Y. P., WANG, Y., LI, Z. L., WANG, S. L. & ZHANG, S. M. 2003. Clinical observation of effect of repaglinide (NovoNorn) combined with NPH (Novolin N) on type 2 diabetes mellitus patients. Tianjin Medical Journal, 31, 294-296.	Studie beinhaltet keinen Vergleich zwischen Kombinationsarmen aus 2 oder weniger Komparatoren von Interesse. Die Studie wurde ausgeschlossen, falls kein nutzbarer Vergleich zwischen einer Mono- oder Kombinationsbehandlung vorgenommen wurde, oder falls die Studie eine Tripletherapie ohne Insulin enthält
144	CHEN, X. L., KE, Q. H. & LIU, C. Y. 2003. Efficacy of combination of multiple drugs in the management of hypertension in elderly patients with type 2 diabetes mellitus. Chinese Journal of Geriatrics, 22, 520-522.	Kein Volltext verfügbar
145	LIU, H. X., QIN, A. P. & HE, B. X. 2003. Clinical observation of efficacy of rosiglitazone combined with insulin in treatment of elderly type 2 diabetes mellitus. Journal of Chinese Physician, 5, 1502-1504.	Kein Volltext verfügbar

	Im Volltext ausgeschlossene Publikation	Grund
146	DEJAGER, S., RAZAC, S., FOLEY, J. E. & SCHWEIZER, A. Vildagliptin in drug-naive patients with type 2 diabetes: A 24-week, double-blind, randomized, placebo-controlled, multiple-dose study. <i>Hormone and Metabolic Research</i> .39 (3) (pp 218-223), 2007.Date of Publication: Mar 2007.	Keine Vergleichstherapie von Interesse
147	DEROSA, G., D'ANGELO, A., FOGARI, E., SALVADEO, S., GRAVINA, A., FERRARI, I. & CICERO, A. F. 2007. Effects of nateglinide and glibenclamide on prothrombotic factors in naive type 2 diabetic patients treated with metformin: a 1-year, double-blind, randomized clinical trial. <i>Internal Medicine</i> , 46, 1837-1846.	Doppelpublikation
148	PI-SUNYER, F. X., SCHWEIZER, A., MILLS, D. & DEJAGER, S. Efficacy and tolerability of vildagliptin monotherapy in drug-naive patients with type 2 diabetes. <i>Diabetes Research and Clinical Practice</i> .76 (1) (pp 132-138), 2007.Date of Publication: Apr 2007.	keine Vergleichstherapie von Interesse
149	ROSENSTOCK, J., AGUILAR-SALINAS, C., KLEIN, E., NEPAL, S., LIST, J. & CHEN, R. Effect of saxagliptin monotherapy in treatment-naive patients with type 2 diabetes. <i>Current Medical Research and Opinion</i> .25 (10) (pp 2401-2411), 2009.Date of Publication: October 2009.	Doppelpublikation
150	MORETTO, T. J., MILTON, D. R., RIDGE, T. D., MACCONELL, L. A., OKERSON, T., WOLKA, A. M. & BRODOWS, R. G. Efficacy and tolerability of exenatide monotherapy over 24 weeks in antidiabetic drug-naive patients with type 2 diabetes: A randomized, double-blind, placebo-controlled, parallel-group study. <i>Clinical Therapeutics</i> .30 (8) (pp 1448-1460), 2008.Date of Publication: August 2008.	Verweis auf Monotherapie
151	FERRANNINI, E., RAMOS, S. J., SALSALI, A., TANG, W. & LIST, J. F. Dapagliflozin monotherapy in type 2 diabetic patients with inadequate glycemic control by diet and exercise: A randomized, double-blind, placebo-controlled, phase 3 trial. <i>Diabetes Care</i> .33 (10) (pp 2217-2224), 2010.Date of Publication: October 2010.	Verweis auf Monotherapie

	Im Volltext ausgeschlossene Publikation	Grund
152	DEL PRATO, S., BARNETT, A. H., HUISMAN, H., NEUBACHER, D., WOERLE, H. J. & DUGI, K. A. Effect of linagliptin monotherapy on glycaemic control and markers of beta-cell function in patients with inadequately controlled type 2 diabetes: A randomized controlled trial. <i>Diabetes, Obesity and Metabolism</i> .13 (3) (pp 258-267), 2011.Date of Publication: March 2011.	Verweis auf Monotherapie
153	MUDALIAR, S., CHANG, A. R., ARODA, V. R., CHAO, E., BURKE, P., BAXI, S., GRIVER, K. A., O'CONNOR, D. T. & HENRY, R. R. Effects of intensive insulin therapy alone and with added pioglitazone on renal salt/water balance and fluid compartment shifts in type 2 diabetes. <i>Diabetes, Obesity and Metabolism</i> .12 (2) (pp 133-138), 2010.Date of Publication: February 2010.	Verweis auf Insulin-Add on
154	ROSENSTOCK, J., AGUILAR-SALINAS, C., KLEIN, E., NEPAL, S., LIST, J. & CHEN, R. Effect of saxagliptin monotherapy in treatment-naive patients with type 2 diabetes. <i>Current Medical Research and Opinion</i> .25 (10) (pp 2401-2411), 2009.Date of Publication: October 2009.	Verweis auf Monotherapie
155	LIST, J. F., WOO, V., MORALES, E., TANG, W. & FIEDOREK, F. T. Sodium-glucose cotransport inhibition with dapagliflozin in type 2 diabetes. <i>Diabetes Care</i> .32 (4) (pp 650-657), 2009.Date of Publication: April 2009.	Verweis auf Monotherapie
156	MOHAN, V., YANG, W., SON, H. Y., XU, L., NOBLE, L., LANGDON, R. B., AMATRUDA, J. M., STEIN, P. P. & KAUFMAN, K. D. Efficacy and safety of sitagliptin in the treatment of patients with type 2 diabetes in China, India, and Korea. <i>Diabetes Research and Clinical Practice</i> .83 (1) (pp 106-116), 2009.Date of Publication: January 2009.	Verweis auf Monotherapie
157	NONAKA, K., KAKIKAWA, T., SATO, A., OKUYAMA, K., FUJIMOTO, G., KATO, N., SUZUKI, H., HIRAYAMA, Y., AHMED, T., DAVIES, M. J. & STEIN, P. P. Efficacy and safety of sitagliptin monotherapy in Japanese patients with type 2 diabetes. <i>Diabetes Research and Clinical Practice</i> .79 (2) (pp 291-298), 2008.Date of Publication: Feb 2008.	Verweis auf Monotherapie

	Im Volltext ausgeschlossene Publikation	Grund
158	HANEFELD, M., HERMAN, G. A., WU, M., MICKEL, C., SANCHEZ, M. & STEIN, P. P. Once-daily sitagliptin, a dipeptidyl peptidase-4 inhibitor, for the treatment of patients with type 2 diabetes. <i>Current Medical Research and Opinion</i> .23 (6) (pp 1329-1339), 2007.Date of Publication: June 2007.	Verweis auf Monotherapie
159	YILMAZ, H., GURSOY, A., SAHIN, M. & GUVENER, D. N. Comparison of insulin monotherapy and combination therapy with insulin and metformin or insulin and rosiglitazone or insulin and acarbose in type 2 diabetes. <i>Acta Diabetologica</i> .44 (4) (pp 187-192), 2007.Date of Publication: December 2007.	Verweis auf Insulin-Add on
160	TERAMOTO, T., YAMADA, N., SHIRAI, K. & SAITO, Y. Effects of pioglitazone hydrochloride on Japanese patients with type 2 diabetes mellitus. <i>Journal of atherosclerosis and thrombosis</i> .14 (2) (pp 86-93), 2007.Date of Publication: Apr 2007.	Verweis auf Monotherapie
161	ZIB, I., JACOB, A. N., LINGVAY, I., SALINAS, K., MCGAVOCK, J. M., RASKIN, P. & SZCZEPANIAK, L. S. Effect of pioglitazone therapy on myocardial and hepatic steatosis in insulin-treated patients with type 2 diabetes. <i>Journal of Investigative Medicine</i> .55 (5) (pp 230-236), 2007.Date of Publication: Jul 2007.	Verweis auf Insulin-Add on
162	GOLDSTEIN, B. J., FEINGLOS, M. N., LUNCEFORD, J. K., JOHNSON, J. & WILLIAMS-HERMAN, D. E. Effect of initial combination therapy with sitagliptin, a dipeptidyl peptidase-4 inhibitor, and metformin on glycemic control in patients with type 2 diabetes. <i>Diabetes Care</i> .30 (8) (pp 1979-1987), 2007.Date of Publication: Aug 2007.	Verweis auf Monotherapie
163	SCOTT, R., WU, M., SANCHEZ, M. & STEIN, P. Efficacy and tolerability of the dipeptidyl peptidase-4 inhibitor sitagliptin as monotherapy over 12 weeks in patients with type 2 diabetes. <i>International Journal of Clinical Practice</i> .61 (1) (pp 171-180), 2007.Date of Publication: Jan 2007.	Verweis auf Monotherapie
164	ASCHNER, P., KIPNES, M. S., LUNCEFORD, J. K., SANCHEZ, M., MICKEL, C. & WILLIAMS-HERMAN, D. E. Effect of the dipeptidyl peptidase-4 inhibitor sitagliptin as monotherapy on glycemic control in patients with type 2 diabetes. <i>Diabetes Care</i> .29 (12) (pp 2632-2637), 2006.Date of Publication: December 2006.	Verweis auf Monotherapie

	Im Volltext ausgeschlossene Publikation	Grund
165	RAZ, I., CHEN, Y., WU, M., HUSSAIN, S., KAUFMAN, K. D., AMATRUDA, J. M., LANGDON, R. B., STEIN, P. P. & ALBA, M. Efficacy and safety of sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes. <i>Current Medical Research and Opinion</i> .24 (2) (pp 537-550), 2008.Date of Publication: Feb 2008.	Verweis auf Monotherapie
166	KHAN, M., MURRAY, F. T., KARUNARATNE, M. & PEREZ, A. Pioglitazone and reductions in post-challenge glucose levels in patients with type 2 diabetes. <i>Diabetes, Obesity and Metabolism</i> .8 (1) (pp 31-38), 2006.Date of Publication: Jan 2006.	Verweis auf Monotherapie
167	HUNG, Y. J., PEI, D., KUO, S. W., HSIEH, C. H., HE, C. T., SHEN, D. C. & SHEU, W. H. H. Pioglitazone monotherapy improves daily plasma glucose and insulin concentrations in type 2 diabetic subjects: A double-blind, placebo-controlled trial. <i>Journal of Medical Sciences</i> .26 (1) (pp 19-24), 2006.Date of Publication: Feb 2006.	Verweis auf Monotherapie
168	MATTOO, V., ECKLAND, D., WIDEL, M., DURAN, S., FAJARDO, C., STRAND, J., KNIGHT, D., GROSSMAN, L., OAKLEY, D. & TAN, M. Metabolic effects of pioglitazone in combination with insulin in patients with type 2 diabetes mellitus whose disease is not adequately controlled with insulin therapy: Results of a six-month, randomized, double-blind, prospective, multicenter, parallel-group study. <i>Clinical Therapeutics</i> .27 (5) (pp 554-567), 2005.Date of Publication: May 2005.	Verweis auf Insulin-Add on
169	CHARBONNEL, B. H., MATTHEWS, D. R., SCHERNTHANER, G., HANEFELD, M. & BRUNETTI, P. A long-term comparison of pioglitazone and gliclazide in patients with Type 2 diabetes mellitus: A randomized, double-blind, parallel-group comparison trial. <i>Diabetic Medicine</i> .22 (4) (pp 399-405), 2005.Date of Publication: Apr 2005.	Verweis auf Monotherapie
170	TAN, M. H., JOHNS, D., STRAND, J., HALSE, J., MADSBAD, S., ERIKSSON, J. W., CLAUSEN, J., KONKOY, C. S. & HERZ, M. Sustained effects of pioglitazone vs. glibenclamide on insulin sensitivity, glycaemic control, and lipid profiles in patients with Type 2 diabetes. <i>Diabetic Medicine</i> .21 (8) (pp 859-866), 2004.Date of Publication: Aug 2004.	Verweis auf Monotherapie

	Im Volltext ausgeschlossene Publikation	Grund
171	HORTON, E. S., FOLEY, J. E., SHEN, S. G. & BARON, M. A. Efficacy and tolerability of initial combination therapy with nateglinide and metformin in treatment-naive patients with type 2 diabetes. <i>Current Medical Research and Opinion</i> .20 (6) (pp 883-889), 2004.Date of Publication: Jun 2004.	Verweis auf Monotherapie
172	STROWIG, S. M., AVILES-SANTA, M. L. & RASKIN, P. Comparison of insulin monotherapy and combination therapy with insulin and metformin or insulin and troglitazone in type 2 diabetes. <i>Diabetes care</i> .25 (10) (pp 1691-1698), 2002.Date of Publication: Oct 2002.	Verweis auf Insulin-Add on
173	SCHERBAUM, W. A. & GOKE, B. Metabolic efficacy and safety of once-daily pioglitazone monotherapy in patients with type 2 diabetes: A double-blind, placebo-controlled study. <i>Hormone and Metabolic Research</i> .34 (10) (pp 589-595), 2002.Date of Publication: 01 Oct 2002.	Verweis auf Monotherapie
174	GARBER, A. J., LARSEN, J., SCHNEIDER, S. H., PIPER, B. A. & HENRY, D. Simultaneous glyburide/metformin therapy is superior to component monotherapy as an initial pharmacological treatment for type 2 diabetes. <i>Diabetes, Obesity and Metabolism</i> .4 (3) (pp 201-208), 2002.Date of Publication: 2002.	Verweis auf Monotherapie
175	HERMANN, L. S., KALEN, J., KATZMAN, P., LAGER, I., NILSSON, A., NORRHAMN, O., SARTOR, G. & UGANDER, L. Long-term glycaemic improvement after addition of metformin to insulin in insulin-treated obese type 2 diabetes patients. <i>Diabetes, Obesity and Metabolism</i> .3 (6) (pp 428-434), 2001.Date of Publication: 2001.	Verweis auf Insulin-Add on

	Im Volltext ausgeschlossene Publikation	Grund
176	ROSENBLATT, S., MISKIN, B., GLAZER, N. B., PRINCE, M. J., ROBERTSON, K. E., CORBIN, R. P., GILDERMAN, L., GILLIE, E., GOLLAPUDI, M., GOVE, R. C., JONES, S., LEWIN, A. J., LODEWICK, P. A., MEMBRENO, L., MENEGHINI, L., MILLER, D., MIDDLETON, D., MOHAN, M. K., MOORE, C., NUNEZ, M., FARMER, M., O'BARR, T., JR., PAPPAS, J., PHILLIPSON, B., RIPLEY, P., ROSS, D. C., ROSS, N., SOLOMON, E., STUCCIO-WHITE, N., SMITH, T., TOH, P., WHELE, S., EGAN, J. W., MATHISEN, A. L. & SCHNEIDER, R. L. The impact of pioglitazone on glycemic control and atherogenic dyslipidemia in patients with type 2 diabetes mellitus. <i>Coronary Artery Disease</i> .12 (5) (pp 413-423), 2001.Date of Publication: 2001.	Verweis auf Monotherapie
177	HORTON, E. S., CLINKINGBEARD, C., GATLIN, M., FOLEY, J., MALLOWS, S. & SHEN, S. Nateglinide alone and in combination with metformin improves glycemic control by reducing mealtime glucose levels in type 2 diabetes. <i>Diabetes Care</i> .23 (11) (pp 1660-1665), 2000.Date of Publication: 2000.	Verweis auf Monotherapie
178	ARONOFF, S., ROSENBLATT, S., BRAITHWAITE, S., EGAN, J. W., MATHISEN, A. L., SCHNEIDER, R. L., GABLES, C., RUOFF, G., WEISS, S., ZAVORAL, J. H., FRANCIS, B., DEABATE, A., GILDERMAN, L., MANGIONE, A., MURRAY, J., GREMILLION, D., WOLFLEY, G., KHAIRI, R., MAGGIACOMO, F. P., WILLIAMS, D. L., BUSICK, E., GREEN, A., LEVINE, J. H., SKOBELOFF, E. & WILLIAMS, T. Pioglitazone hydrochloride monotherapy improves glycemic control in the treatment of patients with type 2 diabetes: A 6-month randomized placebo-controlled dose-response study. <i>Diabetes Care</i> .23 (11) (pp 1605-1611), 2000.Date of Publication: 2000.	Verweis auf Monotherapie
179	AVILES-SANTA, L., SINDING, J. & RASKIN, P. Effects of metformin in patients with poorly controlled, insulin-treated type 2 diabetes mellitus: A randomized, double-blind, placebo-controlled trial. <i>Annals of Internal Medicine</i> .131 (3) (pp 182-188), 1999.Date of Publication: 03 Aug 1999.	Verweis auf Insulin-Add on

	Im Volltext ausgeschlossene Publikation	Grund
180	TURNER, R. C., CULL, C. A., FRIGHI, V. & HOLMAN, R. R. Glycemic control with diet, sulfonylurea, metformin, or insulin in patients with type 2 diabetes mellitus. Progressive requirement for multiple therapies (UKPDS 49). Journal of the American Medical Association.281 (21) (pp 2005-2012), 1999.Date of Publication: 02 Jun 1999.	Verweis auf Monotherapie
181	GOLDBERG, R. B., EINHORN, D., LUCAS, C. P., RENDELL, M. S., DAMSBO, P., HUANG, W. C., STRANGE, P. & BRODOWS, R. G. A randomized placebo-controlled trial of repaglinide in the treatment of type 2 diabetes. Diabetes Care.21 (11) (pp 1897-1903), 1998.Date of Publication: 1998.	Verweis auf Monotherapie
182	ROSENSTOCK, J., SAMOLS, E., MUCHMORE, D. B. & SCHNEIDER, J. Glimepiride, a new once-daily sulfonylurea: A double-blind placebo- controlled study of NIDDM patients. Diabetes Care.19 (11) (pp 1194-1199), 1996.Date of Publication: Nov 1996.	Verweis auf Monotherapie
183	GOLDBERG, R. B., HOLVEY, S. M. & SCHNEIDER, J. A dose-response study of glimepiride in patients with NIDDM who have previously received sulfonylurea agents. Diabetes Care.19 (8) (pp 849-856), 1996.Date of Publication: Aug 1996., Aug.	Verweis auf Monotherapie
184	LEWITT, M. S., YU, V. K. F., RENNIE, G. C., CARTER, J. N., MAREL, G. M., YUE, D. K. S. & HOOPER, M. J. Effects of combined insulin-sulfonylurea therapy in type II patients. Diabetes Care.12 (6) (pp 379-383), 1989.Date of Publication: 1989.	Verweis auf Insulin-Add on
185	CASNER, P. R. Insulin-glyburide combination therapy for non-insulin-dependent diabetes mellitus: A long-term double-blind, placebo-controlled trial. Clinical Pharmacology and Therapeutics.44 (5) (pp 594-603), 1988.Date of Publication: 1988.	Verweis auf Insulin-Add on
186	CAMERINI-DAVALOS, R. A., VELASCO, C. A., GLASSER, M. & BLOODWORTH, JR. Sulfonylurea-induced decrease of muscle capillary basement membrane thickness in diabetes. Diabetes Research and Clinical Practice.5 (2) (pp 113-123), 1988.Date of Publication: 1988.	Verweis auf Monotherapie

	Im Volltext ausgeschlossene Publikation	Grund
187	CAMERINI-DAVALOS, R. A., VELASCO, C. A., REDDI, A. S., GLASSER, M. & BLOODWORTH, JR. Delay of progression of diabetic microangiopathy. <i>Metabolism: Clinical and Experimental</i> .37 (2 SUPPL, 1), 1988.	Verweis auf Monotherapie
188	FALKO, J. M. & OSEI, K. Combination insulin/glyburide therapy in type II diabetes mellitus. Effects on lipoprotein metabolism and glucoregulation. <i>American Journal of Medicine</i> .79 (3 B) (pp 92-101), 1985.Date of Publication: 1985.	Verweis auf Insulin-Add on
189	No authors listed, U.K. prospective diabetes study. II. Reduction in HbA(1c) with basal insulin supplement, sulfonylurea, or biguanide therapy in maturity-onset diabetes. A multicenter study. <i>Diabetes</i> .34 (8) (pp 793-798), 1985.Date of Publication: 1985., 1985.	Verweis auf Monotherapie
190	OSEI, K., O'DORISIO, T. M. & FALKO, J. M. Concomitant insulin and sulfonylurea therapy in patients with type II diabetes. Effects on glucoregulation and lipid metabolism. <i>American Journal of Medicine</i> .77 (6) (pp 1002-1009), 1984.Date of Publication: 1984.	Verweis auf Insulin-Add on
191	TURNER, R. C., MANN, J. I. & ICETON, G. UK prospective study of therapies of maturity-onset diabetes. I. Effect of diet, sulphonylurea, insulin or biguanide therapy on fasting plasma glucose and body weight over one year. <i>Diabetologia</i> .24 (6) (pp 404-411), 1983.Date of Publication: 1983., 1983.	Verweis auf Monotherapie
192	VILSBOLL, T., ROSENSTOCK, J., YKI-JARVINEN, H., CEFALU, W. T., CHEN, Y., LUO, E., MUSSER, B., ANDRYUK, P. J., LING, Y., KAUFMAN, K. D., AMATRUDA, J. M., ENGEL, S. S. & KATZ, L. 2010. Efficacy and safety of sitagliptin when added to insulin therapy in patients with type 2 diabetes. <i>Diabetes, Obesity & Metabolism</i> , 12, 167-177.	Verweis auf Insulin-Add on
193	SCHIEL, R. & MULLER, U. A. 2008. Efficacy and treatment satisfaction of once-daily insulin glargine plus one or two oral antidiabetic agents versus continuing premixed human insulin in patients with Type 2 diabetes previously on long-term conventional insulin therapy: the SWITCH Pilot Study. <i>Experimental & Clinical Endocrinology & Diabetes</i> , 116, 58-64.	Verweis auf Insulin-Add on

	Im Volltext ausgeschlossene Publikation	Grund
194	YAMANOUCHI, T., SAKAI, T., IGARASHI, K., ICHIYANAGI, K., WATANABE, H. & KAWASAKI, T. 2005. Comparison of metabolic effects of pioglitazone, metformin, and glimepiride over 1 year in Japanese patients with newly diagnosed Type 2 diabetes. <i>Diabetic Medicine</i> , 22, 980-985.	Verweis auf Monotherapie
195	ROSENSTOCK, J., EINHORN, D., HERSHON, K., GLAZER, N. B., YU, S. & STUDY, G. 2002. Efficacy and safety of pioglitazone in type 2 diabetes: a randomised, placebo-controlled study in patients receiving stable insulin therapy. <i>International Journal of Clinical Practice</i> , 56, 251-257.	Verweis auf Insulin-Add on
196	HIRSCH, I. B. 1999. Metformin added to insulin therapy in poorly controlled type 2 diabetes. <i>Diabetes Care</i> , 22, 854.	Verweis auf Insulin-Add on
197	RELIMPIO, F., PUMAR, A., LOSADA, F., MANGAS, M. A., ACOSTA, D. & ASTORGA, R. 1998. Adding metformin versus insulin dose increase in insulin-treated but poorly controlled Type 2 diabetes mellitus: an open-label randomized trial. <i>Diabetic Medicine</i> , 15, 997-1002.	Verweis auf Insulin-Add on
198	MATTHEWS, D., JENDLE, J., NAUCK, M., FRID, A., HERMANSEN, K., DURING, M., ZDRAVKOVIC, M. & STRAUSS, B. J. Liraglutide, a once-daily human GLP-1 analogue, reduces fat percentage and visceral and subcutaneous adipose tissue compared with glimepiride when added to metformin in subjects with Type 2 diabetes. <i>Diabetic Medicine</i> . Conference: Diabetes UK's Annual Professional Conference 2009 Glasgow United Kingdom. Conference Start: 20090311 Conference End: 20090313. Conference Publication: (var.pagings).26 (pp 21-22), 2009. Date of Publication: March 2009.	Berichtet keinen Ausgangswert und keinen Folgewert für mind. einen Endpunkt
199	ARECHA VALETA, R., SECK, T., CHEN, Y., KROBOT, K. J., O'NEILL, E. A., DURAN, L., KAUFMAN, K. D., WILLIAMS-HERMAN, D. & GOLDSTEIN, B. J. Efficacy and safety of treatment with sitagliptin or glimepiride in patients with type 2 diabetes inadequately controlled on metformin monotherapy: A randomized, double-blind, non-inferiority trial. <i>Diabetes, Obesity and Metabolism</i> . 13 (2) (pp 160-168), 2010. Date of Publication: February 2011.	Bezieht Sulfonylharnstoff mit ein (24-wöchiges Netzwerk), aber keine Angaben zum systolischen Blutdruck

	Im Volltext ausgeschlossene Publikation	Grund
200	COMASCHI, M., CORSI, A., DI, P. C., BELLATRECCIA, A. & MARIZ, S. The effect of pioglitazone as add-on therapy to metformin or sulphonylurea compared to a fixed-dose combination of metformin and glibenclamide on diabetic dyslipidaemia. <i>Nutrition, Metabolism and Cardiovascular Diseases</i> .18 (5) (pp 373-379), 2008.Date of Publication: June 2008.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
201	UMPIERREZ, G., ISSA, M. & VLAJNIC, A. Glimepiride versus pioglitazone combination therapy in subjects with type 2 diabetes inadequately controlled on metformin monotherapy: Results of a randomized clinical trial. <i>Current Medical Research and Opinion</i> .22 (4) (pp 751-759), 2006.Date of Publication: Apr 2006.	Bezieht Sulfonylharnstoff mit ein (24-wöchiges Netzwerk), aber keine Angaben zum systolischen Blutdruck
202	MARRE, M., HOWLETT, H., LEHERT, P. & ALLAVOINE, T. Improved glycaemic control with metformin-glibenclamide combined tablet therapy (Glucovance) in Type 2 diabetic patients inadequately controlled on metformin. <i>Diabetic Medicine</i> .19 (8) (pp 673-680), 2002.Date of Publication: 2002	Studiendauer <12 Wochen
203	MARRE, M., VAN, G. L., USADEL, K. H., BALL, M., WHATMOUGH, I. & GUITARD, C. Nateglinide improves glycaemic control when added to metformin monotherapy: Results of a randomized trial with type 2 diabetes patients. <i>Diabetes, Obesity and Metabolism</i> .4 (3) (pp 177-186), 2002.Date of Publication: 2002.	Bezieht Sulfonylharnstoff mit ein (24-wöchiges Netzwerk), aber keine Angaben zum systolischen Blutdruck
204	MOSES, R. Repaglinide in combination therapy with metformin in Type 2 diabetes. <i>Experimental and Clinical Endocrinology and Diabetes</i> .107 (SUPPL, 4), 1999.	Bezieht Sulfonylharnstoff mit ein (24-wöchiges Netzwerk), aber keine Angaben zum systolischen Blutdruck
205	MOSES, R., CARTER, J., SLOBODNIUK, R., DONNELLY, T., BOYAGES, S., MOFFITT, P., COLAGIURI, S., HOPKINS, H. & KIDSON, W. Effect of repaglinide addition to metformin monotherapy on glycemic control in patients with type 2 diabetes. <i>Diabetes Care</i> .22 (1) (pp 119-124), 1999.Date of Publication: Jan 1999.	Bezieht Sulfonylharnstoff mit ein (24-wöchiges Netzwerk), aber keine Angaben zum systolischen Blutdruck

	Im Volltext ausgeschlossene Publikation	Grund
206	FORST, T., UHLIG-LASKE, B., RING, A., GRAEFEMODY, U., FRIEDRICH, C., HERBACH, K., WOERLE, H. J. & DUGI, K. A. 2010. Linagliptin (BI 1356), a potent and selective DPP-4 inhibitor, is safe and efficacious in combination with metformin in patients with inadequately controlled Type 2 diabetes. <i>Diabetic Medicine</i> , 27, 1409-1419.	Studiendauer <12 Wochen
207	RISTIC, S., COLLOBER-MAUGEAIS, C., PECHER, E. & CRESSIER, F. 2006. Comparison of nateglinide and gliclazide in combination with metformin, for treatment of patients with Type 2 diabetes mellitus inadequately controlled on maximum doses of metformin alone. <i>Diabetic Medicine</i> , 23, 757-762.	Bezieht Sulfonylharnstoff mit ein (24-wöchiges Netzwerk), aber keine Angaben zum systolischen Blutdruck
208	NAUCK, M., HERMANSEN, K., FRID, A., SHAH, N., TANKOVA, T., MITHA, I., DURING, M., ZDRAVKOVIC, M. & MATTHEWS, D. Liraglutide, a once-daily human GLP-1 analogue, in type 2 diabetes provides similar glycaemic control with reduced body weight compared with glimepiride when added to metformin. <i>Journal of Diabetes.Conference: 3rd International Congress on Prediabetes and the Metabolic Syndrome Nice France.Conference Start: 20090401 Conference End: 20090404.Conference Publication: (var.pagings).1 (pp A21), 2009.Date of Publication: April 2009</i>	Doppelpublikation zu Nauck 2009
209	JENDLE, J., NAUCK, M. A., MATTHEWS, D. R., FRID, A., HERMANSEN, K., DURING, M., ZDRAVKOVIC, M., STRAUSS, B. J. & GARBER, A. J. Weight loss with liraglutide, a once-daily human glucagon-like peptide-1 analogue for type 2 diabetes treatment as monotherapy or added to metformin, is primarily as a result of a reduction in fat tissue. <i>Diabetes, Obesity and Metabolism</i> .11 (12) (pp 1163-1172), 2009.Date of Publication: December 2009.	Doppelpublikation zu Nauck 2009
210	BOLLI, G., DOTTA, F., COLIN, L., MINIC, B. & GOODMAN, M. Comparison of vildagliptin and pioglitazone in patients with type 2 diabetes inadequately controlled with metformin. <i>Diabetes, Obesity and Metabolism</i> .11 (6) (pp 589-595), 2009.Date of Publication: 2009.	Doppelpublikation zu Bolli 2008

	Im Volltext ausgeschlossene Publikation	Grund
211	HERMANSEN, K., KOLOTKIN, R. L., HAMMER, M., ZDRAVKOVIC, M. & MATTHEWS, D. 2010. Patient-reported outcomes in patients with type 2 diabetes treated with liraglutide or glimepiride, both as add-on to metformin. Primary care diabetes, 4, 113-117.	Doppelpublikation zu Nauck 2009
212	SECK, T., NAUCK, M., SHENG, D., SUNGA, S., DAVIES, M. J., STEIN, P. P., KAUFMAN, K. D. & AMATRUDA, J. M. Safety and efficacy of treatment with sitagliptin or glipizide in patients with type 2 diabetes inadequately controlled on metformin: A 2-year study. International Journal of Clinical Practice.64 (5) (pp 562-576), 2010.Date of Publication: April 2010.	Doppelpublikation zu Nauck 2007
213	AHREN, B., FOLEY, J. E., FERRANNINI, E., MATTHEWS, D. R., ZINMAN, B., DEJAGER, S. & FONSECA, V. A. 2010. Changes in prandial glucagon levels after a 2-year treatment with vildagliptin or glimepiride in patients with type 2 diabetes inadequately controlled with metformin monotherapy. Diabetes Care, 33, 730-732.	Doppelpublikation zu Matthews 2010

In der aktualisierten Literaturrecherche vom 04. Juni 2012 wurden zusätzlich 8 Vollpublikationen gesichtet und begründet ausgeschlossen.

Tabelle 4-154 (Anhang): Im Volltext ausgeschlossene nicht relevante Publikationen (aktualisierte Literaturrecherche vom 04. Juni 2012)

	Im Volltext ausgeschlossene Publikation	Grund
1	HAAK, T., MEINICKE, T., JONES, R., VON, E. M. & WOERLE, H. J. 2011. Combination of linagliptin and metformin improves glycemic control in type 2 diabetes: A randomized trial with an open-label arm in patients with poor glycemic control. Diabetes.Conference: 71st Scientific Sessions of the American Diabetes Association San Diego, CA United States.Conference Start: 20110624 Conference End: 20110628.Conference Publication: (var.pagings), 60, July.	Verweis auf Monotherapie
2	WILLIAMS-HERMAN, D., XU, L., TENG, R., GOLM, G. T., JOHNSON, J., DAVIES, M. J., KAUFMAN, K. D. & GOLDSTEIN, B. J. 2012. Effect of initial combination therapy with sitagliptin and metformin on beta-cell function in patients with type 2 diabetes. Diabetes, Obesity and Metabolism, 14, January.	Verweis auf Monotherapie

	Im Volltext ausgeschlossene Publikation	Grund
3	CHIEN, M. N., LEE, C. C., CHEN, W. C., LIU, S. C., LEUNG, C. H. & WANG, C. H. 2011. Effect of sitagliptin as add-on therapy in elderly type 2 diabetes patients with inadequate glycemic control in Taiwan. <i>International Journal of Gerontology</i> , 5, June.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
4	FINEMAN, M., FLANAGAN, S., TAYLOR, K., AISPORNA, M., SHEN, L. Z., MACE, K. F., WALSH, B., DIAMANT, M., CIRINCIONE, B., KOTHARE, P., LI, W. I. & MACCONELL, L. Pharmacokinetics and pharmacodynamics of exenatide extended-release after single and multiple dosing. <i>Clinical Pharmacokinetics</i> .50 (1) (pp 65-74), 2011.Date of Publication: 2011.	Die Studie bezieht sich nicht auf die Zielpopulation (inadequater glykämischer Kontrolle zum Zeitpunkt des Studieneinschlusses mit Diät und Bewegung allein (Monotherapie) Metformin-Monotherapie (Metformin <i>Add-on</i>) Insulin mit bis zu zwei weiteren oralen Antidiabetika (Insulin <i>Add-on</i>)
5	SECK, T. L., ENGEL, S. S., WILLIAMS-HERMAN, D. E., SISK, C. M., GOLM, G. T., WANG, H., KAUFMAN, K. D. & GOLDSTEIN, B. J. 2011. Sitagliptin more effectively achieves a composite endpoint for A1C reduction, lack of hypoglycemia and no body weight gain compared with glipizide. <i>Diabetes Research & Clinical Practice</i> , 93, e15-e17.f	Berichtet keinen Ausgangswert und keine Folgewerte für mind. einen Endpunkt
6	SCHONDORF, T., MUSHOLT, P. B., HOHBERG, C., FORST, T., LEHMANN, U., FUCHS, W., LOBIG, M., MULLER, J. & PFUTZNER, A. 2011. The fixed combination of pioglitazone and metformin improves biomarkers of platelet function and chronic inflammation in type 2 diabetes patients: results from the PIOfix study. <i>Journal of Diabetes Science & Technology</i> , 5, 426-432.	Berichtet keinen Ausgangswert und keine Folgewerte für mind. einen Endpunkt
7	HAAK, T., MEINICKE, T., JONES, R., WEBER, S., EYNATTEN, M. V. & WOERLE, H. J. 2012. Initial	Verweis auf Monotherapie

	Im Volltext ausgeschlossene Publikation	Grund
	combination of linagliptin and metformin improves glycaemic control in type 2 diabetes: A randomized, double-blind, placebo-controlled study. <i>Diabetes, Obesity and Metabolism</i> , 14, June	
8	PFÜTZNER, A., SCHONDORF, T., TSCHOPE, D., LOBMANN, R., MERKE, J., MULLER, J., LEHMANN, U., FUCHS, W. & FORST, T. 2011. PIOfix-study: effects of pioglitazone/metformin fixed combination in comparison with a combination of metformin with glimepiride on diabetic dyslipidemia. <i>Diabetes Technology & Therapeutics</i> , 13, 637-643.	Bezieht Sulfonylharnstoff mit ein (24-wöchiges Netzwerk), aber keine Angaben zum systolischen Blutdruck

Tabelle 4-155 (Anhang): Im Volltext ausgeschlossene nicht relevante Publikationen (aktualisierte Literaturrecherche vom 09. Oktober 2012)

	Im Volltext ausgeschlossene Publikation	Grund
1	SRIVASTAVA S., SAXENA GN., KESHWANI P., GUPTA R. 2012. Comparing the efficacy and safety profile of sitagliptin versus glimepiride in patients of type 2 diabetes mellitus inadequately controlled with metformin alone. <i>J Assoc Physicians India</i> , 60, 27-30.	Dosierung und/oder Dosierungsschema nicht zugelassen
2	DEROSA G., RAGONESI PD., CARBONE A., FOGARI E., BIANCHI L., BONAVENTURA A., ROMANO D., CICERO AF., MAFFIOLI P. 2012 Vildagliptin added to metformin on β -cell function after a euglycemic hyperinsulinemic and hyperglycemic clamp in type 2 diabetes patients. <i>Diabetes Technol Ther.</i> , 14(6):475-84	Die Studie bezieht sich nicht auf die Zielpopulation (inadequate glykämische Kontrolle zum Zeitpunkt des Studieneinschlusses mit Metformin-Monotherapie (Metformin <i>Add-on</i>))
3	KROBOT KJ., FERRANTE SA., DAVIES MJ., SECK T., MEININGER GE., WILLIAMS-HERMAN D., KAUFMAN KD., GOLDSTEIN BJ. 2012 Lower risk of hypoglycemia with sitagliptin compared to glipizide when either is added to metformin therapy: a pre-specified analysis adjusting for the most recently measured HbA(1c) value. <i>Curr Med Res Opin.</i> 28(8):1281-7.	Berichtet keinen Ausgangswert und keine Folgewerte für mind. einen Endpunkt
4	ROSENSTOCK J., AGGARWAL N., POLIDORI D., ZHAO Y., ARBIT D., USISKIN K., CAPUANO G., CANOVATCHEL W. 2012 Dose-ranging effects of canagliflozin, a sodium-glucose cotransporter 2 inhibitor, as add-on to metformin in subjects with type 2 diabetes. <i>Diabetes care.</i> 35(6):1232-8.	Studiendauer $\leq 17,9$ Wochen

	Im Volltext ausgeschlossene Publikation	Grund
5	ROSS SA., RAFEIRO E., MEINICKE T., TOORAWA R., WEBER-BORN S., WOERLE HJ. 2012 Efficacy and safety of linagliptin 2.5 mg twice daily versus 5 mg once daily in patients with type 2 diabetes inadequately controlled on metformin: a randomised, double-blind, placebo-controlled trial. <i>Curr Med Res Opin.</i> 28(9):1465-74.	Studiendauer $\leq 17,9$ Wochen

Anhang 4-D: Liste der abgebrochenen Studien

Benennen Sie in der nachfolgenden Tabelle die abgebrochenen Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-156 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
--						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
--						
aktivkontrolliert, andere Vergleichstherapie(n)						
D1690C 00007	nein	ja	nein	nein	Ja [EUCTR2008-004916-12-DE] (ICTRP Search Portal, 2012g)	nein
MB102 020	nein	ja	nein	(Bristol-Myers Squibb Company, 2010b)	Ja [NCT00726505] (ICTRP Search Portal, 2012j) (Clinicaltrials.gov, 2012c)	nein
a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.						
B: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.						
C: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder –ergebnisse.						

Anhang 4-E: Liste der laufenden Studien

Benennen Sie in der nachfolgenden Tabelle die laufenden Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-157 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
MB102-054*	nein	ja	nein	nein	ja [NCT01095653] (Clinicaltrials.gov, 2012f)	nein
MB102-055	nein	ja	nein	nein	ja [NCT01095666] (Clinicaltrials.gov, 2012)}	nein
MB102-073	nein	ja	nein	nein	ja [NCT01137474] (Clinicaltrials.gov, 2012~)	nein
MB102-077	nein	ja	nein	nein	ja [NCT01195662] (Clinicaltrials.gov, 2012b)	nein
CV1811-69	nein	ja	nein	nein	ja [NCT01606007] (Clinicaltrials.gov, 2012d)	nein
CV1811-68	nein	ja	nein	nein	Ja [NCT01619059] (clinicaltrials.gov, 2012j)	nein
D1690C-00018 (LT II)	nein	ja	nein	nein	ja [NCT01031680] (Clinicaltrials.gov, 2012r) (ICTRP Search Portal, 2012i)	nein
D1690C	nein	ja	nein	nein	ja	nein

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
00019 (LT II)					[NCT01042977] (Clinicaltrials.gov, 2012h) (ICTRP Search Portal, 2012h)	
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
D1690C 00004 (LT II)	nein	ja	nein	nein	ja [NCT00660907] (Clinicaltrials.gov, 2012a) [EUCTR2007-005220-33-GB] (ICTRP Search Portal, 2012a)	nein
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.</p> <p>b: Studie, für die der Unternehmer Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

* Die Studie ist abgeschlossen. Es liegen jedoch keine Ergebnisse vor.

Anhang 4-F: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-158 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-158 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-158 (Anhang): Studiendesign und -methodik für Studie D1690C00004

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Das primäre Ziel (52 Wochen): Beurteilung, ob nach einer doppelblinden 52 Wochen langen Behandlung, die Veränderung des HbA1c (vom Ausgangswert) durch eine Behandlung mit Metformin+Dapagliflozin einer Behandlung mit Metformin+Glipizid (Sulfonylharnstoff) nicht unterlegen ist bei Patienten mit Typ-2-Diabetes, die eine unzureichende Blutzuckerkontrolle unter einer Metformin-Monotherapie (1500 mg oder höher) haben</p> <p>Sekundäre Hauptziele:</p> <ul style="list-style-type: none"> - Gewichtsreduktion unter Metformin+Dapagliflozin im Vergleich zu Metformin+Glipizid nach 52 Wochen - Anzahl von Hypoglykämien unter Metformin+Dapagliflozin im Vergleich zu Metformin+Glipizid nach 52 Wochen <p>Andere sekundäre Ziele: Wirksamkeit: Vergleich des Effekts von Metformin+Dapagliflozin mit Metformin+Glipizid auf zusätzliche Gewichts- und glykämische Endpunkte Sicherheit: Beurteilung der Sicherheit und Verträglichkeit (mittels Erhebung von unerwünschten Ereignissen, Laborwerten, EKG, Pulsmessung, Blutdruckmessung, Hypoglykämien, Kreatininclearance und Werten der körperlichen Untersuchung)</p> <p>Ziele der Verlängerungsphase:</p> <ul style="list-style-type: none"> - Beurteilung der Sicherheit und Verträglichkeit (wie nach den ersten 52 Wochen) - Beurteilung der Nachhaltigkeit von Metformin+Dapagliflozin <p>Tertiäres Ziel: Vergleich von Metformin+Dapagliflozin mit Metformin+Glipizid an zusätzlichen Wirksamkeitspunkten</p> <p>Ziele (nach 104 Wochen): Beurteilung der gleichen Sicherheits- und Verträglichkeitsendpunkte nach 104 Wochen wie nach 52 Wochen Beurteilung der Erhaltung der Wirksamkeit von Metformin+Dapagliflozin nach 104-wöchiger Behandlung</p>
-	Methoden	
3	Studiendesign	

Item ^a	Charakteristikum	Studieninformation
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Internationale, multizentrische, randomisierte (1:1), parallele, aktivkontrollierte und doppelblinde 52-wöchige Phase-III-Studie mit einer 52-wöchigen Verlängerungsphase I und einer 104-wöchigen Verlängerungsphase II
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht zutreffend
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien bei der 1. Visite (<i>Enrollment</i>):</p> <ul style="list-style-type: none"> - Einverständniserklärung - Diagnose: Typ-2-Diabetes - Männer oder Frauen, Alter ≥ 18 Jahre - Frauen im gebärfähigem Alter müssen entsprechende Methoden zur Verhütung benutzen - Behandlung mit Metformin (stabile Dosis von 1500 mg oder höher pro Tag) für mindestens 8 Wochen vor der 1. Visite (Erlaubt war zusätzlich ein weiteres orales Antidiabetikum mit halber Maximaldosierung) <p>Einschlusskriterien bei der 2. Visite (Laborwerte von der 1. Visite):</p> <ul style="list-style-type: none"> - HbA1c $>6,5\%$ und $\leq 10,0\%$ (Patienten mit HbA1c $>6,5\%$ und $<7,0\%$ wurden nicht weiter eingeschlossen, falls schon über 25% der eingeschlossenen Patienten einen HbA1c von unter 7% hatten) - FPG ≤ 270 mg/dL - C-Peptid Level $\geq 1,0$ ng/mL <p>Einschlusskriterium bei der 4. Visite (Metformin <i>dose-stabilization</i> Phase):</p> <ul style="list-style-type: none"> - Gleichbleibende Metformin-Behandlung mit 1500 mg/Tag oder höher für mindestens 8 Wochen <p>Falls nicht notwendig, da bereits eine Monotherapie mit Metformin gegeben wurde, konnte die Metformin <i>dose-stabilization</i>-Phase übersprungen werden</p> <p>Ausschlusskriterien bei der 1. Visite:</p> <ul style="list-style-type: none"> - Type-1-Diabetes, vorbelastet mit diabetischer Ketoazidose oder hyperosmolarem, nicht ketotischem Koma - Vorbelastet mit Diabetes insipidus - Symptome eines schlecht

Item ^a	Charakteristikum	Studieninformation
		<p>eingestellten Diabetes</p> <ul style="list-style-type: none"> - Vorgeschichte einer labilen oder einer schnell fortschreitenden Nierenkrankheit - Vorliegen einer erblich bedingten renalen Glukosurie - Schwanger oder stillend - BMI >45,0 kg/m² - Insulintherapie innerhalb des Jahres des Einschluss - Teilnahme an einer klinischen Studie mit Dapagliflozin und/oder mit einem anderen SGLT2 Inhibitor - Behandlung mit Glucocorticoiden innerhalb 30 Tagen vor Einschluss - Adipositaschirurgie - Gewichtsreduzierende Medikation - Behandlung mit antiviralen Medikamenten / HIV-Behandlung - Unverträglichkeit / Kontraindikation oder eine potentielle Allergie gegen Dapagliflozin, Metformin, Glipizid oder den Placeboinhaltsstoffen - kongestive Herzinsuffizienz - Bedeutende kardiovaskuläre Vorgeschichte innerhalb der letzten 6 Monate vor der 1. Visite - Schwere Atemwegserkrankungen - Schwerer unkontrollierter Bluthochdruck (systolischer Blutdruck \geq180 mm Hg und/oder diastolischer Blutdruck \geq110 mm Hg) - Patienten mit Risiko einer Dehydrierung (nach Einschätzung des Prüfarztes) - Vorgeschichte von Hämoglobinopathie - Drogen oder Alkoholmissbrauch innerhalb der letzten 12 Monate - Bösartiger Tumor innerhalb der letzten 5 Jahre - Beteiligung an der Planung und Durchführung der Studie - Teilnahme an einer klinischen Studie während der letzten 90 Tage vor der 1. Visite - Blut-, Plasma- oder Blutplättchenspende innerhalb der

Item ^a	Charakteristikum	Studieninformation
		<p>letzten 3 Monate vor der 1. Visite</p> <ul style="list-style-type: none"> - Schlechte Protokoll- oder Medikations-<i>Compliance</i> <p>Ausschlusskriterien bei der 2. Visite (Laborwerte von der 1. Visite):</p> <ul style="list-style-type: none"> - Serum Kreatinin <60 mL/min - UACR (Urin Albumin: Kreatin Verhältnis) >1800 mg/g - Lebererkrankung und/oder eine abnormale Leberfunktion - Gesamteiweiß im Urin >2 mg/dL - Kreatinin Kinase ≥ 3 * ULN (<i>upper limit of normal</i>) - Hämoglobin $\leq 11,0$ g/dL für Männer und $\leq 10,0$ g/dL für Frauen - Abnormale TSH- Werte (Thyroid stimulierende Hormonwerte) - Positives serologisches Ergebnis einer infektiösen Lebererkrankungen - Klinisch bedeutende Abnormitäten, die aufgrund der Labortests oder der körperlichen Untersuchung entdeckt wurden <p>Ausschlusskriterien bei der 5. Visite (Randomisierung):</p> <ul style="list-style-type: none"> - Klinisch bedeutende Abnormitäten, die aufgrund des EKG oder der körperlichen Untersuchung entdeckt wurden
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische, internationale Studie durchgeführt an 95 Zentren in 10 Ländern (Frankreich, Deutschland, Italien, Niederlande, Spanien, Schweden, UK, Argentinien, Brasilien, Südafrika)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Woche 0-18 Titrations-Phase: <i>Open-label</i> Metformin (1500, 2000 oder 2500 mg/Tag+Dapagliflozin (Dosisanpassung in 3 Schritten 2,5, 5 oder 10 mg; entweder bis zum optimalen Effekt oder zur höchsten verträglichen Dosis)</p> <p>versus</p> <p><i>Open-label</i> Metformin (1500, 2000 oder 2500 mg/Tag)+Glipizid (Dosisanpassung in 3 Schritten 5, 10 oder 20 mg; entweder bis zum optimalen Effekt oder zur höchsten verträglichen Dosis)</p> <p>Woche 18-52 Maintenance-Phase: <i>Open-label</i> Metformin (1500, 2000 oder 2500 mg/Tag)+stabile Dosis Dapagliflozin</p>

Item ^a	Charakteristikum	Studieninformation
		versus <i>Open-label</i> Metformin (1500, 2000 oder 2500 mg/Tag)+stabile Dosis Glipizid
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primär: Veränderung vom Ausgangswert des HbA1c zu Woche 52</p> <p>Sekundäre Hauptendpunkte: Veränderung des Körpergewichts vom Ausgangswert zu Woche 52</p> <p>Anzahl der Patienten mit mindestens einer Hypoglykämie</p> <p>Anzahl der Patienten mit einer Gewichtsreduktion von mindestens 5% in Woche 52</p> <p>Sekundäre Endpunkte: Veränderung des Körpergewichts bei Patienten mit BMI ≥ 30 kg/m²</p> <p>Veränderung des Körpergewichts bei Patienten mit BMI ≥ 27 kg/m</p> <p>Veränderung des Bauchumfangs</p> <p>Veränderung des HbA1c bei Patienten mit einem HbA1c-Ausgangswert $\geq 7\%$</p> <p>Anteil der Patienten mit einem HbA1c $< 7\%$ bei Patienten mit einem HbA1c-Ausgangswert $\geq 7\%$</p> <p>Anteil der Patienten mit einem HbA1c $< 6,5\%$</p> <p>Veränderung des FPG vom Ausgangswert bis Woche 52</p> <p>Veränderung der AUC (0-180 Minuten) für PPG in der OGTT-Subgruppe</p> <p>Veränderung des systolischen Blutdrucks vom Ausgangswert bis Woche 52</p> <p>Veränderung des systolischen und diastolischen Blutdrucks bei Patienten mit einem systolischen Blutdruck > 140 mmHg zu Beginn der Studie</p> <p>Sicherheit: Unerwünschte Ereignisse, Laborwerte, EKG, Puls, Blutdruck, Hypoglykämien, Kreatininclearance und Ergebnisse der körperlichen Untersuchung</p> <p>Anzahl der Patienten mit einem Ödem innerhalb der 52 Wochen</p> <p>Tertiäre Endpunkte:</p> <p>Behandlungszufriedenheit anhand des DTSQc in Woche 52</p> <p>Behandlungszufriedenheit anhand des DTSQc in Woche 26</p> <p>Veränderung des HbA1c bei Patienten mit einem HbA1c-Ausgangswert $> 6,5\%$ und $< 8\%$</p> <p>Veränderung des HbA1c bei Patienten mit einem HbA1c-Ausgangswert $\geq 8\%$ und $< 9\%$</p> <p>Veränderung des HbA1c bei Patienten mit</p>

Item ^a	Charakteristikum	Studieninformation
		<p>einem HbA1c-Ausgangswert $\geq 9\%$</p> <p>Veränderung von <i>Fasting</i> Insulin, C-Peptid und Proinsulin</p> <p>Veränderung von HOMA-2</p> <p>Veränderung der AUC (0-180 Minuten) für Insulin und C-Peptid</p> <p>Veränderung von TC, LDL-C, HDL-C, TG und FFS (freie Fettsäuren)</p> <p>Ziele nach 104 Wochen:</p> <p>Sicherheit: wie nach 52 Wochen</p> <p>Wirksamkeit:</p> <p>Veränderung des HbA1c von der 52. bis zur 104. Woche</p> <p>Veränderung vom Ausgangswert des HbA1c zu Woche 104</p> <p>Veränderung vom Ausgangswert des Körpergewichts zu Woche 104</p> <p>Anteil der Patienten mit einer Gewichtsreduzierung von mindestens 5% nach 104 Wochen</p> <p>Hypoglykämien innerhalb 104 Wochen</p> <p>Veränderung des FPG nach 104 Wochen</p> <p>Studienabbrüche aufgrund unzureichender Blutzuckerkontrolle</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Endpunkte zum Blutdruck wurden von den tertiären Endpunkten zu den sekundären Endpunkten verschoben
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Annahmen, um die Nichtunterlegenheit (anhand der Veränderung des HbA1c) zu beweisen:</p> <p>Standardabweichung 1,25%, einseitiges Signifikanzniveau von 0,025 bei einer Nichtunterlegenheitsschranke von 0,35%; <i>Power</i>: 90%, 5%ige Ausschlussrate</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Patienten erhalten einen E-code (Randomisierungsnummer) nach dem unterschriebenen <i>Informed Consent</i>
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Balancierte Blöcke 1:1
9	<p>Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment)</p> <p>Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war</p>	Zuteilung zur Studienmedikation wird durch das <i>interactive web response system</i> (IWRS) durchgeführt (unabhängig und zentral)

Item ^a	Charakteristikum	Studieninformation
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	AstraZeneca stellt Randomisierungsnummern sowie deren <i>Kit</i> -Nummern zur Verfügung. IWRS führt Zuteilung durch.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) ja b) ja c) ja Randomisierung und identische Studienmedikation (<i>double-dummy</i> Technik)
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Verwendung von gleich vielen und identisch aussehenden Studienmedikamenten für Patienten und das medizinischen Personal
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Eine hierarchische Testprozedur wurde zwischen den primären und sekundären Hauptendpunkten verwendet, um den Fehler 1. Art zu kontrollieren. Die Nichtunterlegenheit wurde erst getestet (einseitig, alpha: 0,025), dann die Überlegenheit von Metformin+Dapagliflozin in den sekundären Hauptendpunkten (zweiseitig, alpha: 0,05). Der primäre Endpunkt wurde analysiert mit einem ANCOVA-Modell, andere stetige sekundäre Endpunkte wurden ebenfalls mit einem ANCOVA-Modell analysiert. Diskrete Variablen wurden anhand der Anzahl, Anteil und dem dazugehörigen Konfidenzintervall zusammengefasst. Wirksamkeitsendpunkte wurden anhand des <i>Full Analysis Set</i> , der primäre Endpunkt anhand des <i>Full Analysis Set</i> und des <i>Per-Protocol Analysis Set</i> untersucht. Sicherheitsendpunkte wurden anhand des <i>Safety Analysis Set</i> untersucht. Nach 104 Wochen wurden Wirksamkeitsendpunkte mit Modellen mit wiederholten Messungen analysiert.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Subgruppen: Veränderung des HbA1c nach Geschlecht, Alter, Rasse, Region, HbA1c-Ausgangswert und eGFR-Ausgangswert
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums	a) 406 vs 408 b) 400 vs 401 c) 400 vs 401

Item ^a	Charakteristikum	Studieninformation
	berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 17)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studienbeginn: 31.03.2008 Studienende (regulär): 15.12.2009 Studienende Verlängerungsphase I: 15.12.2010 Studienende Verlängerungsphase II: läuft noch
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde auf 104 Wochen (Verlängerungsphase I) und anschließend auf 208 Wochen verlängert (Verlängerungsphase II). Diese Verlängerungsphase dauert noch an.
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

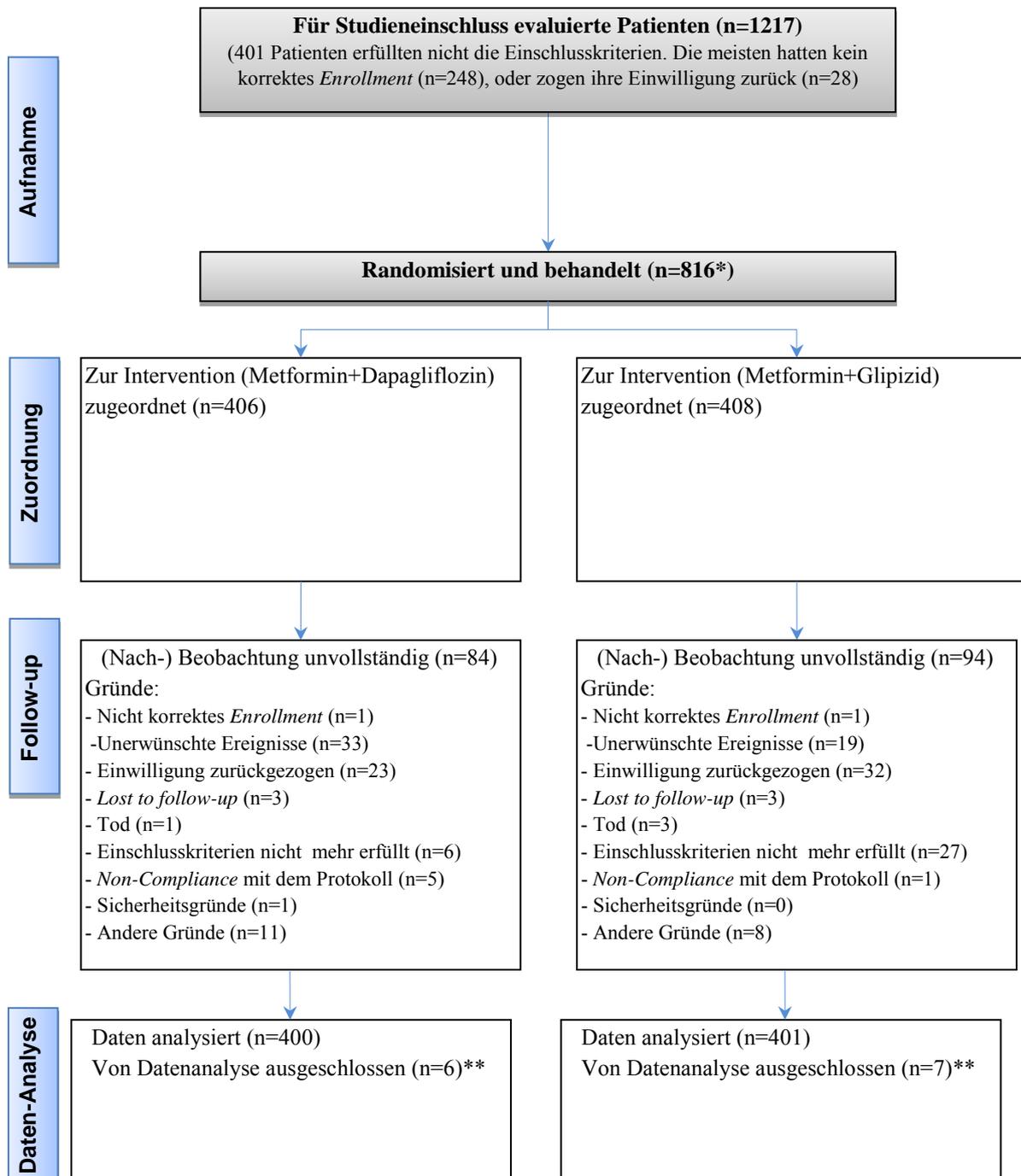


Abbildung 17: Flussdiagramm Studiendesign und -methodik für Studie D1690C00004

*2 Patienten erhielten keine Studienmedikation, **13 Patienten wurden ausgeschlossen, weil sie keinen Ausgangswert bzw. keinen Wert in einem Wirksamkeitsendpunkt hatten

Studien für die MTC

Tabelle 4-159 (Anhang): Studiendesign und -methodik für Studie MB102014 (Bailey, 2010)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Die Studie untersuchte die Wirksamkeit und Sicherheit von Dapagliflozin, verabreicht zusätzlich zu Metformin, in erwachsenen Patienten mit Typ-2-Diabetes mellitus (T2DM), die durch Metformin allein nicht ausreichend kontrolliert waren.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Bei der Studie handelt es sich um eine Phase-III-, randomisierte, doppelblinde, parallelgruppen-geführte (im Verhältnis 1:1) ^b , placebokontrollierte Sicherheits- und Wirksamkeitsstudie.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderungen der Methodik nach dem Studienbeginn
4	Probanden / Patienten	Mit Metformin allein nicht ausreichend kontrollierte Typ-2-Diabetiker
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Erwachsene Patienten im Alter von 18-77 Jahren mit Typ-2-Diabetes, mit HbA1c-Wert von 7-10%, C-Peptid-Konzentration $\geq 0,34$ nmol/L, BMI ≤ 45 kg/m², und die mit einer stabilen Dosis Metformin (≥ 1500 mg/Tag) mindestens für 8 Wochen vor der Eingangsuntersuchung behandelt wurden.</p> <p>Ausschlusskriterien: Serumkreatinin ≥ 133 μmol für Männer oder ≥ 124 μmol für Frauen (konsistent mit Metformin-Anwendung), Urinalbumin/Kreatininverhältnis $> 203,4$ mg/mmol, Aspartat-Aminotransferase oder Alanin-Transferase größer als das Dreifache der Obergrenze des Normalwertes, Kreatinkinase größer als das Dreifache der Obergrenze des Normalwertes, Symptome vom schlecht kontrollierten Diabetes (einschließlich ausgeprägter Polyurie und Polydypsie mit $> 10\%$ Gewichtsverlust während der letzten 3 Monate vor dem Studienbeginn), klinisch relevante Nieren-, Leber-, hämatologische, onkologische, endokrine, psychiatrische oder rheumatische Erkrankung, jüngstes kardiovaskuläres Ereignis (in den letzten 6 Monaten) oder kongestives Herzinsuffizienz (gemäß <i>New York Heart Association class III/IV</i>), und</p>

Item ^a	Charakteristikum	Studieninformation
		systolischer Blutdruck ≥ 180 mmHg oder diastolischer Blutdruck ≥ 110 mmHg.
4b	Studienorganisation und Ort der Studiendurchführung	Internationale Studie mit 80 Zentren: 30 in USA, 21 in Kanada, 11 in Argentinien, 10 in Mexico, 8 in Brasilien
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Dapagliflozin 10 mg/Tag (n=135) verglichen mit Placebo einmal täglich (n=137), beide zusätzlich zu einer bestehenden nicht-verblindeten Metformin-Therapie.
6	Zielkriterien	Veränderung des HbA1c-Wertes (%) in Woche 24 im Vergleich zu Studienbeginn
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Das primäre Zielkriterium war: Veränderung des HbA1c-Wertes (%) in Woche 24 im Vergleich zu Studienbeginn. Die sekundären Zielkriterien waren: Änderung der Nüchternplasmaglukose (FPG)-Konzentration und des Körpergewichts nach 24 Wochen, Änderung von FPG nach 1 Woche, Anteil der Patienten mit glykämischer Response, und Änderung des HbA1c-Wertes (%) nach 24 Wochen bei Patienten mit HbA1c-Ausgangswert $\geq 9\%$.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderung der Zielkriterien nach Studienbeginn
7	Fallzahl	Randomisierungszahl: Dapagliflozin 10 mg/Tag: n=135, Placebo: n=137
7a	Wie wurden die Fallzahlen bestimmt?	Berechnung der Stichprobengröße und der statistischen <i>Power</i> wurde wie folgt durchgeführt: 129 Patienten mit Untersuchungen nach dem Studieneintritt pro Behandlungsarm bedeuten 90% statistische <i>Power</i> , um Unterschiede in HbA1c-Mittelwerten zwischen jeder Dapagliflozin-Behandlungsgruppe und der Placebo-Gruppe von 0,5% zu erkennen, unter der Annahme einer Standardabweichung (SD) von 1,1% (Buse, 2005). Haben 5% der Patienten keine Messungen nach dem Studieneintritt, sind 544 Patienten (136 in jeder Gruppe) notwendig für die Randomisierung.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Computerbasierte Generierung
8a	Methode zur Generierung der zufälligen Zuteilung	Computerbasierte Generierung mit implementierter zentraler, telefonbasierter interaktiver Anlage (<i>interactive voice response system</i> , IVRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung war computergeneriert und wurde nach Prüfzentrum stratifiziert. Ein

Item ^a	Charakteristikum	Studieninformation
		Prüfzentrum wurde als Block zugeordnet, als der erste Patient des Zentrums randomisiert wurde.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Probanden bekamen eine eindeutige, fortlaufende Nummer von dem zentralen IVRS zugewiesen. Der Prüfarzt ersuchte IVRS telefonisch um die Zuordnung.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Probanden wurden von den Prüfarzten der verschiedenen Zentren in die Studie aufgenommen. Den qualifizierten Probanden wurde eine eindeutige fortlaufende Nummer von dem zentralen IVRS zugewiesen. Während Medikamentenverteilung wies IVRS jedem Patienten eine Set-Nummer, entsprechend den auf den Studienmedikations-Packungen und -Flaschen aufgedruckten Nummern.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Die Probanden, Prüfarzte und Sponsorpersonal waren verblindet bezüglich Behandlung, HbA1c und Uringlukose-Konzentrationen. Randomisierungsschema wurde in einem sicheren Ort aufbewahrt. Im Falle eines medizinischen Notfalls, für den das Wissen über die Behandlung des Patienten entscheidend war, war dem Prüfarzt gestattet, die Verblindung des Patienten zu brechen.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Placebo- und Dapagliflozin-Filmtabletten waren ähnlich in Farbe, Form, Größe, Textur und Geschmack.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Analysen der kontinuierlichen Ergebnisse basierten auf separaten Modellen der Kovarianzanalyse (ANCOVA), die Behandlungsgruppe als Effekt und der Ausgangswert als Kovariate, für fehlende Daten wurde die letzte Beobachtung übertragen (<i>last observation carried forward</i>). Als Teil der sekundären Analysen, der Vergleich der Anteile der Patienten, die einen therapeutischen glykämischen Response erhielten, wurde mit einer logistischen Regression, basierend auf einer etablierten Methodik mit Ausgangswertanpassung, durchgeführt (Zhang, 2008; Tsiatis, 2008). Nach Korrektur der Ausgangswerte und Differenzen der durchschnittlichen Veränderungen vom Ausgangswert

Item ^a	Charakteristikum	Studieninformation
		<p>gegenüber der Woche 24, wurden p-Werte für primäre und sekundäre Endpunkte berechnet, um den Behandlungseffekt von Dapagliflozin gegenüber Placebo zu vergleichen.</p> <p>Für die primäre Analyse (Änderung von HbA1c vom Ausgangswert zu Woche 24) wurden Dapagliflozin-Gruppen und die Placebo-Gruppe bei $\alpha=0,019$ Dunnett-Prozedur anwendend verglichen. War der primäre Endpunkt signifikant, wurden die Änderungen vom Ausgangswert in sekundären Endpunkten sequenziell bei $\alpha=0,05$ in folgender Reihenfolge getestet: FPG-Konzentration in Woche 24, Körpergewicht in Woche 24, Anteil der Patienten mit HbA1c <7,0% in Woche 24, HbA1c-% in Woche 24 bei Patienten mit Ausgangs-HbA1c $\geq 9\%$, Körpergewicht in Woche 24 bei Patienten mit Ausgangs-BMI $\geq 27 \text{ kg/m}^2$ (Supplementary webappendix, Bailey, 2010), FPG-Konzentration in Woche 1 und der Anteil der Patienten, die HbA1c $\leq 6,5\%$ in Woche 24 erreichten.</p> <p>In Übereinstimmung mit dem Studiendesign wurden keine p-Werte für exploratorische Endpunkte generiert. Für Sicherheit wurden nur zusammenfassende Statistiken gemeldet. Die statistischen Analysen wurden mit SAS/STAT Version 8.2 ausgeführt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Subgruppenanalysen vom primären Endpunkt und Nebenwirkungen (vorgeplante exploratorische Ziele):</p> <ul style="list-style-type: none"> • Rasse (asiatisch; schwarz / Afroamerikaner; weiß; andere) • Geschlecht (weiblich; männlich) • Ethnische Zugehörigkeit (hispanisch / Latino; nicht-hispanisch / Latino) (nur Wirksamkeit) • HbA1c-Ausgangswert (<8%; >8% and <9%; >9%) (nur Wirksamkeit) • Alter (≤ 65 Jahre; >65 Jahre) • Geografische Region (Nord-Amerika; Süd-Amerika; Europa; Asien / Pazifik) (nur Wirksamkeit) • Geschlecht / Alter (weiblich ≤ 50 Jahre; weiblich >50 Jahre)
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 18)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben,	Siehe <i>Flow-Chart</i> (Abbildung 18)

Item ^a	Charakteristikum	Studieninformation
	c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 18)
14	Aufnahme / Rekrutierung	Die Probanden wurden vom 18. September, 2007 bis zum 10. April, 2008 rekrutiert. Der 24-Wochen-Stichtag war der 4. November, 2008.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Probanden wurden über 24 Wochen beobachtet, mit Nachbeobachtungen nach 1, 2, 4, und alle 4 Wochen danach.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend – der 24-Wochen-Periode folgte eine 78-Wochen-Verlängerungsperiode
a: nach CONSORT 2010.		

Anmerkungen: ^aNach CONSORT 2010 checklist (Moher, 2010); ^b Randomisierung erfolgte zu drei Dosen von Dapagliflozin (2,5 mg, 5 mg, 10 mg) und Placebo im Verhältnis 1:1:1:1, aber nur Dapagliflozin 10 mg-Arm und Placebo-Arm wurden für die Nutzenbewertung einbezogen

Abkürzungen: HbA1c, Glykiertes Hämoglobin; IVRS, Interactive voice response system; SD, Standardabweichung; T2DM, Type-2- Diabetes mellitus; FPG, Nüchternplasmaglukose; USA, United States of America

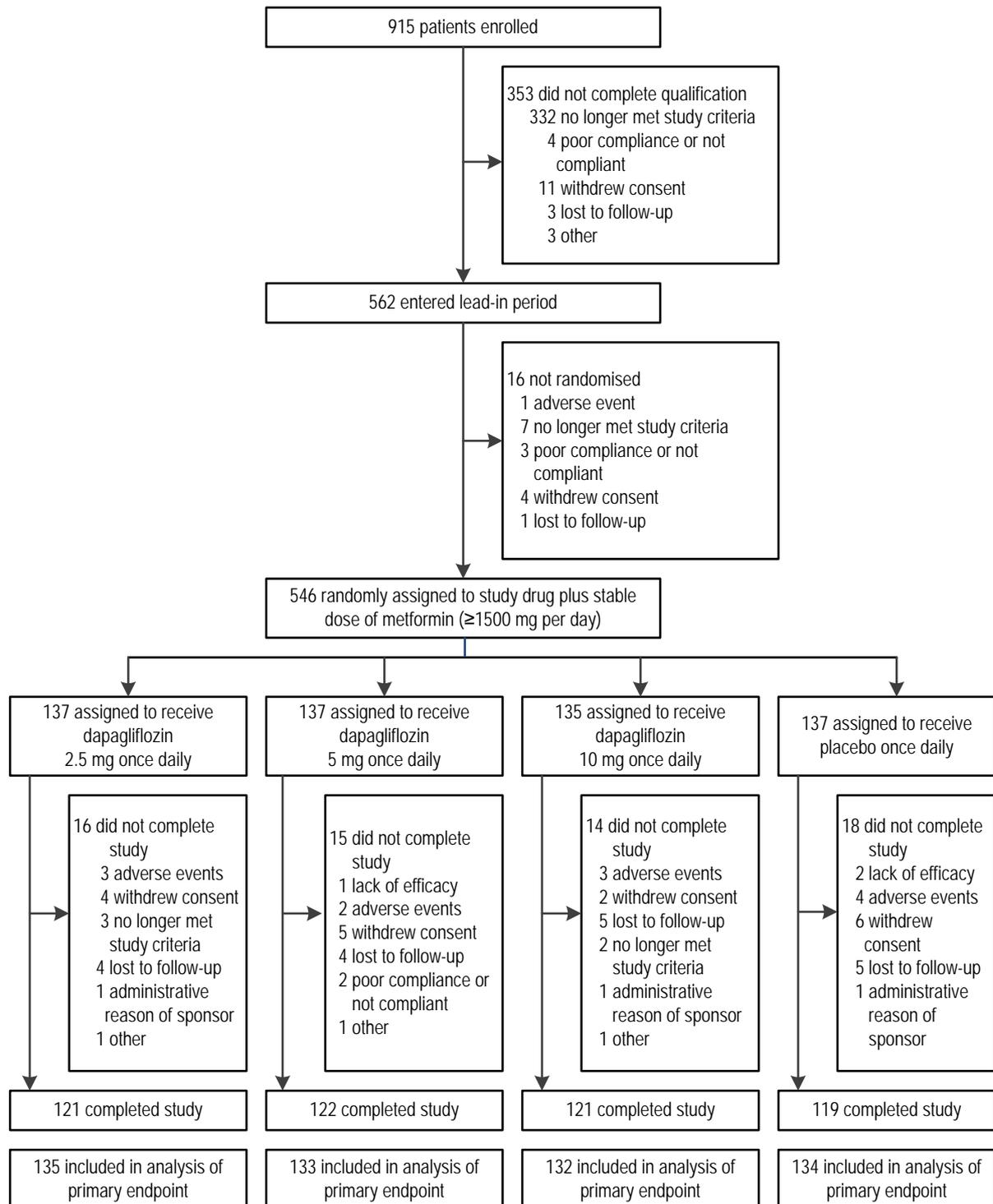


Abbildung 18: Flussdiagramm Studiendesign und -methodik für Studie MB102014 (Bailey, 2010)

Anmerkung: Randomisierung erfolgte zu drei Dosen von Dapagliflozin (2,5 mg, 5 mg, 10 mg) und Placebo im Verhältnis 1:1:1:1, aber nur Dapagliflozin 10 mg-Arm und Placebo-Arm wurden für die Nutzenbewertung einbezogen

Tabelle 4-160 (Anhang): Studiendesign und -methodik für die Studie D1690C00012 (Bolinder, 2012)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Diese Studie untersuchte die Gewichtsabnahme mit Dapagliflozin als Zusatztherapie zu Metformin bei erwachsenen Patienten mit Typ-2-Diabetes mellitus (T2DM), die mit Metformin allein nicht ausreichend kontrolliert waren.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Bei der Studie handelt es sich um eine 24-wöchige, Phase III-, randomisierte, doppelblinde, parallelgruppen-geführte (im Verhältnis 1:1), placebokontrollierte Studie mit einer noch laufenden zentrum- und patientenverblindete 78-wöchigen Verlängerungsperiode.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<ol style="list-style-type: none"> 1. Kriterien und Vorgehen für Studienabbrüche wegen erhöhter Leberenzymwerte wurden zu einer detailliertere Handhabung geändert, einschließlich zusätzlicher Blutentnahmen im Falle anhaltend erhöhter Leberwertabweichungen. Diese Änderung erfolgte wegen Ausweitung der bestehenden Verfahren nach Einführung von <i>US Food and Drug Administration (FDA) Guidance for Industry, Drug-Induced Liver Injury: Pre-marketing Clinical Evaluation</i>. 2. In der Sicherheitsanalyse wurde ‚randomisierte‘ zu ‚alle‘ Patienten, die mindestens eine Dosis von der doppelblinden Studienmedikation erhielten, geändert. Diese Änderung wurde vor der Entblindung der Studiendaten vorgenommen, und wurde vorgenommen, um Patienten in die Sicherheitsanalyse einzuschliessen, die aus Versehen doppelblinde Studienmedikation erhielten, aber nicht randomisiert waren.
4	Probanden / Patienten	Mit Metformin allein nicht ausreichend kontrollierte Typ-2-Diabetiker
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Die Probanden hatten Typ-2-Diabetes, das Alter für Männer war ≥ 30 Jahre und für Frauen ≥ 55 Jahre, sie wurden mit Metformin mit einer stabilen Dosis von ≥ 1500 mg/Tag für mindestens 12 Wochen vor dem Einschluß in die Studie behandelt; und sie</p>

Item ^a	Charakteristikum	Studieninformation
		<p>hatten eine nicht ausreichende glykämische Kontrolle, definiert als HbA1c $\geq 6,5\%$ und $\leq 8,5\%$.</p> <p>Ausschlusskriterien:</p> <p>Probanden mit Typ-1-Diabetes; Diabetes insipidus; kortikosteroid-induzierte Typ-2-Diabetes; Vorgeschichte von diabetischer Ketoazidose oder nichtketotischem hyperosmolarem Koma; schlecht kontrolliertes Diabetes charakterisiert mit Polyurie/Polydipsie mit $>5\%$ Gewichtsverlust; Insulinbehandlung in den letzten 6 Monaten vor der ersten Studienvisite mit Ausnahme von vorübergehender Insulineinnahme für <7 Tage; Körpergewichtsänderungen $>5\%$ in den letzten drei Monaten vor dem Eintritt in die Studie; systolischer Blutdruck ≥ 180 mmHg und/oder diastolischer Blutdruck ≥ 110 mmHg; kardiovaskuläres Ereignis in den letzten 6 Monaten vor dem Eintritt in die Studie; kongestive Herzinsuffizienz; angeborene Nierenglykosurie, signifikante Nieren-, Leber-, Atem-, hämatologische, onkologische, endokrine, immunologische (einschließlich Überempfindlichkeit gegen Studienmedikation), und Alkohol- und/oder Drogenmissbrauch-Erkrankungen; Schwangerschaft und/oder Stillen; Adipositaschirurgie; Gebrauch von Abführmitteln in den letzten 30 Tagen vor dem Eintritt in die Studie.</p>
4b	Studienorganisation und Ort der Studiendurchführung	Europäische Studie mit 40 Zentren in 5 Ländern (Bulgarien, Tschechische Republik, Ungarn, Polen und Schweden)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Dapagliflozin 10 mg/Tag (n=91) verglichen mit Placebo einmal täglich (n=91), beide zusammen zu der üblichen Dosis des nicht-verblindeten Metformins.
6	Zielkriterien	Änderung des Körpergewichts (kg) in Woche 24 im Vergleich zum Ausgangswert
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Das primäre Zielkriterium war:</p> <p>Änderung des Körpergewichts (kg) in Woche 24 im Vergleich zum Ausgangswert.</p> <p>Die sekundären Zielkriterien waren:</p> <p>Änderung im Taillenumfang (cm) in Woche 24 im Vergleich zum Ausgangswert; Änderung von Körperfettanteil in Woche 24 im Vergleich zum Ausgangswert, gemessen mit <i>Dual-X-Ray</i>-Absorptiometrie; und der Anteil von Probanden mit Gewichtsabnahme $\geq 5\%$ nach 24 Wochen.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderung der Zielkriterien nach Studienbeginn

Item ^a	Charakteristikum	Studieninformation
7	Fallzahl	Randomisierungszahl: Dapagliflozin 10 mg/Tag: n=91, Placebo: n=91
7a	Wie wurden die Fallzahlen bestimmt?	Die Berechnungen der Stichprobengröße basierten auf früheren 12-Wochen Daten der Körpergewichtsänderungen mit Dapagliflozin (Zhang, 2010), in denen die durchschnittliche, placebokorrigierte Änderung im Körpergewicht in der 10 mg-Dapagliflozingruppe 1,3 kg in Woche 12 betrug, und die Standardabweichung aller Dapagliflozin-Gruppen 2,6 kg betrug. Es wurde erwartet, daß die 24-Wochen Daten der aktuellen Studie größere Gewichtsabnahme (2 kg) wie auch größere Variabilität aufzeigen würden. Bei einer angenommenen ca. 50%igen Zunahme in der Variabilität wurde eine Standardabweichung von 4,0 kg ausgewählt. Um eine Differenz von 2 kg zwischen den Behandlungsgruppen nachzuweisen, waren 86 auswertbare Probanden je Behandlungsgruppe für 90% statistische Power mit einem zweiseitigen Signifikanzniveau von 0,05 erforderlich. Angenommen, dass 5% von den randomisierten Probanden von der primären Analyse wegen fehlender Daten ausgeschlossen würden, mussten mindestens 182 Patienten randomisiert werden.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Computerbasierte Generierung
8a	Methode zur Generierung der zufälligen Zuteilung	Randomisierung erfolgte nach einem prädefinierten computergenerierten Randomisierungsschema
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung erfolgte in zwei Strata, Männer und Frauen, und wurde in balancierten Blockgrößen von vier durchgeführt, um ungefähr gleiche Anzahl der Patienten in den Behandlungsgruppen und innerhalb der Strata zu gewährleisten.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Zuordnung der Patienten in die Behandlungsgruppen erfolgte nach einem prädefinierten computergenerierten Randomisierungsschema.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Probanden wurden von den Prüfarzten in die Studie aufgenommen und zu der Behandlung nach einem prädefinierten computergenerierten Randomisierungsschema zugewiesen. Das Randomisierungsschema wurde von dem Studiensponsor zur Verfügung gestellt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b)	Die Probanden, Prüfarzte und

Item ^a	Charakteristikum	Studieninformation
	diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Sponsorpersonal waren verblindet bezüglich der Behandlungszuteilungscodes, die in einem sicheren Ort aufbewahrt wurden. Im Falle eines medizinischen Notfalls, für den das Wissen über die Behandlung des Patienten entscheidend war, war dem Prüfarzt gestattet, die Verblindung des Patienten zu brechen.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Placebo- und Dapagliflozin (10 mg)-Filmtabletten waren identisch in Aussehen, Geruch und Geschmack, und waren in identischen Flaschen abgepackt.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die primären, sekundären Haupt- und exploratorischen Endpunkte wurden mit der <i>Full analysis set</i>-Population durchgeführt.</p> <p>Für glykämische Variablen wurden die Beobachtungen nach Beginn einer Notfalltherapie von der Analyse ausgeschlossen. Diese wie andere fehlende glykämische Daten und die nicht-glykämischen Variablen in Woche 24 wurden mit den Daten der letzten Beobachtung ersetzt (LOCF-Methode).</p> <p>Zwei Analysensets wurden definiert: Sicherheitsanalysenset umfasste alle Patienten, die mindestens eine Dosis des Studienmedikaments erhielten. <i>Full analysis set</i> umfasste alle randomisierten Patienten, die mindestens eine Dosis des Studienmedikaments erhielten und die neben dem Ausgangswert mindestens einen weiteren Wert zumindest für eine Wirksamkeitsvariable hatten.</p> <p><i>Closed testing procedure</i> wurde verwendet, um die Typ-1-Fehlerrate bei $\leq 0,05$ (zweiseitig) für den primären und drei sekundäre Hauptendpunkte zu prüfen. Falls der primäre Endpunkt signifikant mit einem Niveau von $< 0,05$ war, wurden die Ergebnisse der sekundären Hauptendpunkte dann mittels Hochbergs Methode interpretiert (Hochberg, 1988).</p> <p>Kontinuierliche Endpunkte wurden mittels Kovarianzanalyse bewertet, mit Behandlungsarm und Geschlecht als feste Effekte (<i>fixed effects</i>) und Ausgangswert als Kovariate. Die Quoten wurden mittels logistischer Regressionsanalyse mit Anpassung für die Ausgangswerte und Geschlecht analysiert wie von Zhang et al. (Zhang, 2008) beschrieben. P-Werte wurden für primäre Endpunkte und sekundäre Hauptendpunkte, und Konfidenzintervalle mit nominalen p-Werten für exploratorische Endpunkte berichtet.</p>

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Subgruppenanalysen vom primären Endpunkt und Nebenwirkungen (vorgeplante exploratorische Ziele): <ul style="list-style-type: none"> • Geschlecht (weiblich; männlich) • Alter (<65 Jahre; >65 Jahre) • Ethnische Zugehörigkeit (alle Teilnehmer waren weiß; deshalb dieselbe Ergebnisse wie in der Hauptanalyse) • Ausgangswert HbA1c (<8%; >8% und <9%; >9%) (nur Wirksamkeit) • Ausgangswert BMI <27; >27 und <30; >30 kg/m² (nur Wirksamkeit) • Eingangswert der geschätzten glomerulären Filtrationsrate (>30 und <60; >60 und <90; und >90 mL/min/1.73 m²) (nur Wirksamkeit)
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 19)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die <ul style="list-style-type: none"> a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden 	Siehe <i>Flow-Chart</i> (Abbildung 19)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 19)
14	Aufnahme / Rekrutierung	Die 24-wöchige Studie wurde vom 13. Februar 2009 an durchgeführt, und die 78-wöchige Verlängerungsstudie ist noch nicht abgeschlossen. Der Stichtag für die 24-wöchige Studie war der 3. Juni 2010.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Probanden wurden über 24 Wochen beobachtet, mit Untersuchungen nach Randomisierung in Wochen 0, 4, 8, 16 und 24.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

Abkürzungen: BMI, *Body-Mass-Index*; FDA, *Food and Drug Administration*; HbA1c, Glykiertes Hämoglobin; IVRS, *Interactive voice response system*; LOCF, *Last observation carried forward*; RCT, Randomisierte kontrollierte Studie; SD, Standardabweichung; T2DM, Type-2-Diabetes mellitus; USA, United States of America

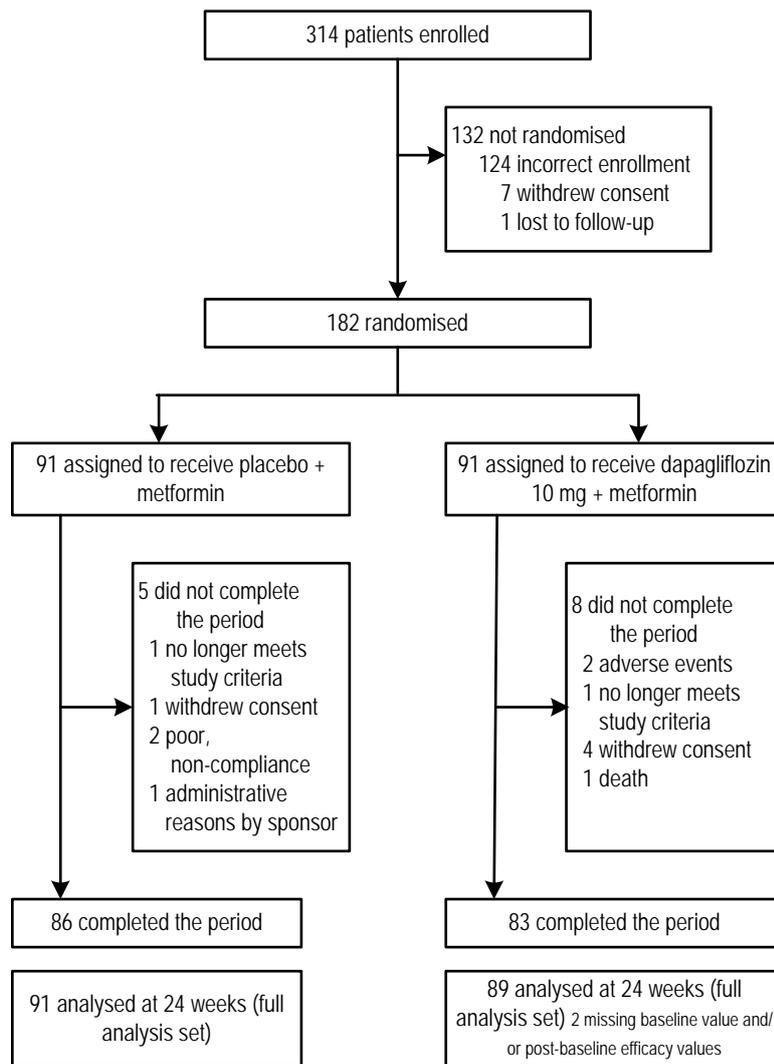


Abbildung 19: Flussdiagramm Studiendesign und -methodik für die Studie D1690C00012 (Bolinder, 2012)

Tabelle 4-161 (Anhang): Studiendesign und -methodik für die Studie von Scott 2008

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Studienziel war die Überprüfung der Wirkung von Sitagliptin bei Patienten mit Typ-2-Diabetes unter Metformin-Behandlung aber unzureichend durch Metformin allein kontrolliert
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Es handelte sich um eine multinationale, randomisierte, doppelblinde, parallele (im Verhältnis 1:1), placebokontrollierte Wirksamkeits/Sicherheitsstudie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nach Beginn der Studie wurden keine Änderungen vorgenommen
4	Probanden / Patienten	Patienten mit Typ-2-Diabetes, unzureichend durch Metformin kontrolliert
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: Männer und Frauen mit Typ-2-Diabetes (Alter 18-75 Jahre), die Metformin als Monotherapie in einer stabilen Dosis von ≥ 1500 mg/Tag für mindestens 10 Wochen vor der Eingangsuntersuchung erhielten, aber unzureichende glykämische Kontrolle zeigten (definiert durch $HbA1c \geq 7$ und $\leq 11\%$) Ausschlusskriterien: Typ1-Diabetes, Insulingebrauch innerhalb 8 Wochen vor Eingangsuntersuchung, Kontraindikation für TZDs oder Metformin, sich verschlechternde Nierenfunktion (Kreatininclearance < 60 ml/min), Alanin-Aminotransferase oder Aspartat-Aminotransferase mehr als zweifach über dem Normalwert, Nüchternblutglukosewert > 270 mg/dL vor der Randomisierung
4b	Studienorganisation und Ort der Studiendurchführung	nicht berichtet
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	nicht zutreffend
6	Zielkriterien	Veränderungen der HbA1c-Ausgangswerte im Vergleich zu den Werten in Woche 18 in Prozenten
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Veränderungen der HbA1c-Ausgangswerte im Vergleich zu den Werten in Woche 18 in Prozenten Sekundäre Zielkriterien: Nüchternblutglukosewert (FPG), Nüchternseruminsulin,

Item ^a	Charakteristikum	Studieninformation
		Nüchternserumproinsulin, Nüchternplasmalipide
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nach Studienbeginn wurden keine Änderungen vorgenommen
7	Fallzahl	Anzahl der randomisierten Patienten; Sitagliptin 100 mg/Tag n=94 Placebo n=92
7a	Wie wurden die Fallzahlen bestimmt?	nicht berichtet
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	nicht berichtet
8a	Methode zur Generierung der zufälligen Zuteilung	nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar, die Autoren beschreiben die Studie als doppelblind, aber beschreiben nicht explizit wer verblindet war und wie die Verblindung durchgeführt wurde
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht zutreffend
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Wirksamkeitsanalysen basierten auf der Patientenpopulation, die nach Randomisierung mindestens eine Dosis des Studienmedikaments erhalten hatte und nach Eingangsuntersuchung an mindestens einer weiteren Untersuchung teilgenommen hat. Durch Analyse der Kovarianz mittels ANCOVA Modell wurden die Behandlungsgruppen bezüglich kontinuierlichen Wirksamkeitsparametern verglichen. Schwerpunkt lag auf Veränderungen zwischen Ausgangswerten und Werten in Woche 18, mit den Ausgangswerten als Kovariate. Die Betrachtung fehlender Daten erfolgte mit der

Item ^a	Charakteristikum	Studieninformation
		LOCF-Methode. Mittels ANCOVA wurden <i>least squares</i> (LS) Mittelwertsänderungen (oder prozentuale Änderungen) zwischen Ausgangswerten und Werten in Woche 18 (95% Konfidenzintervall) für jeden Wirksamkeitsendpunkt über alle Behandlungsgruppen berechnet. Die Studie war angelegt um zu zeigen, dass Sitagliptin-der Placebo-Behandlung überlegen ist. Um Zwischen-Gruppenunterschiede aufzuzeigen wurden Interferenzttests mit p-Werten spezifiziert, dann Veränderungen in den Sitagliptin- vs. Placebo-Gruppen verglichen. Der Rosiglitazon-Behandlungsarm wurde nur zu Abschätzungszwecken eingeschlossen. Differenzen in LS Mittelwertsänderungen (prozentualen Änderungen) und 95% Konfidenzintervall wurden berechnet und damit Rosiglitazon- und Placebo- bzw. Rosiglitazon- und Sitagliptin-Gruppen aufgrund der Zwischen-Gruppenunterschiede miteinander zu vergleichen. Ein p-Wert $\leq 0,05$ wurde als statistisch signifikant angesehen.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Subgruppen nach Geschlecht, Alter (< oder ≥ 65 Jahre), Rasse, Ausgangs-BMI, HbA1c-Ausgangswert, Dauer der Diabeteserkrankung
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 20)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 20)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 20)
14	Aufnahme / Rekrutierung	nicht berichtet
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Patienten wurden über 18 Wochen beobachtet, mit Untersuchungen nach Randomisierung in Woche 0, 3, 6, 12 und 18
14b	Informationen, warum die Studie endete oder beendet wurde	nicht zutreffend
a: nach CONSORT 2010.		

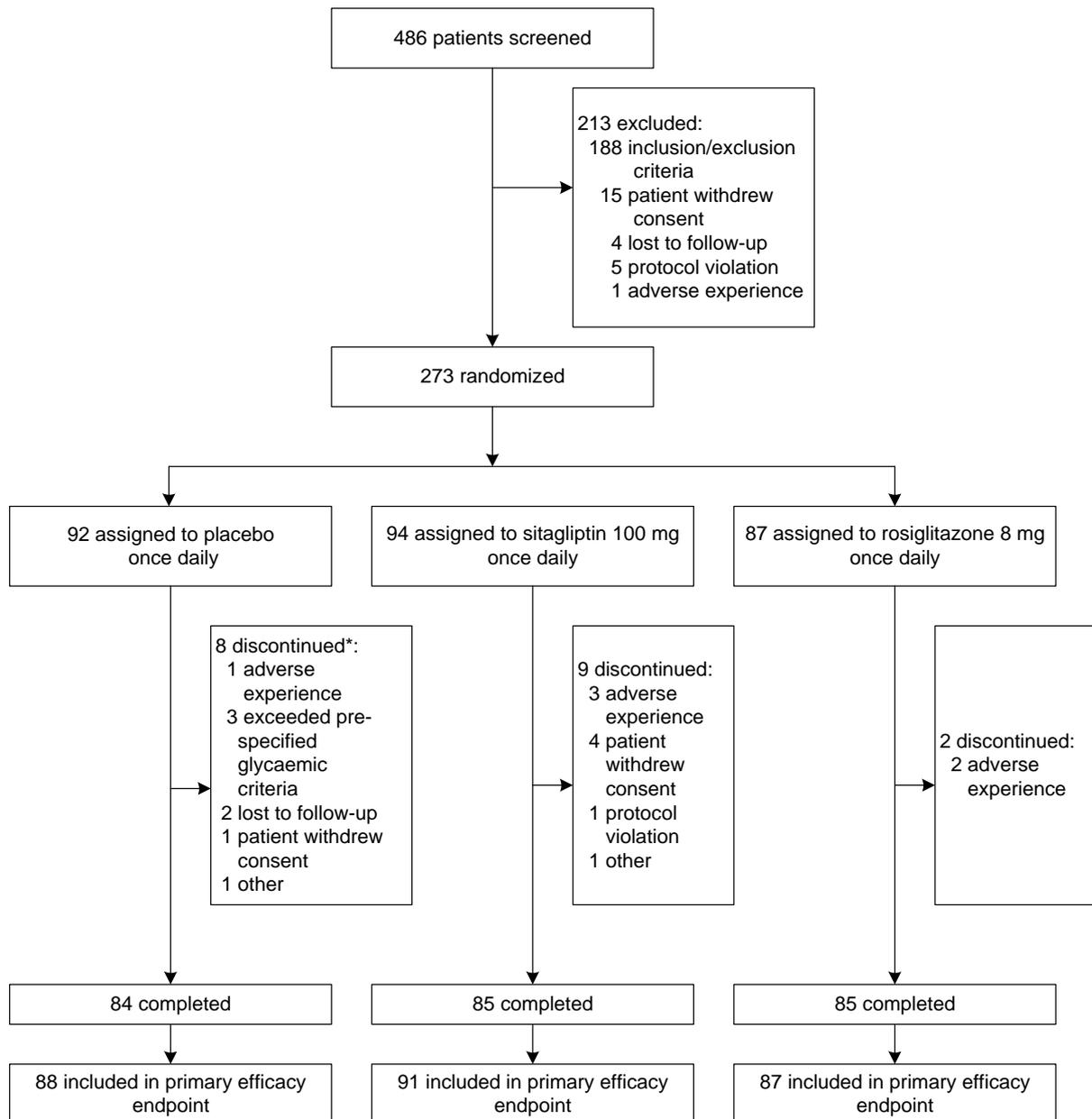


Abbildung 20: Flussdiagramm Studiendesign und -methodik für die Studie von Scott 2008.

Anmerkung: Randomisierung erfolgte für Placebo, Sitagliptin und Rosiglitazon im Verhältnis 1:1:1, aber nur der Placebo-Arm und der Sitagliptin-Arm wurden für die Nutzenbewertung einbezogen

Tabelle 4-162 (Anhang): Studiendesign und -methodik für die Studie von Charbonnel 2006

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Diese Studie wurde angelegt um die Wirksamkeit und Sicherheit des Dipeptidyl-Peptidase-4-Inhibitors, Sitagliptin zu überprüfen. Sitagliptin wurde erwachsenen Typ-2-Diabetes Patienten zusätzlich zu Metformin verabreicht sofern diese durch Metformin allein nicht ausreichend therapiert waren.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Die Studie war eine multinationale, randomisierte, doppelblinde, placebokontrollierte, parallel angelegte (Verhältnis 1:2) Wirksamkeits- und Sicherheitsstudie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Es wurden keine Änderungen vorgenommen
4	Probanden / Patienten	Typ-2-Diabetiker mit unzureichender glykämischer Kontrolle unter Metformin-Behandlung (≥ 1500 mg/Tag)
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>In die Studie aufgenommen und randomisiert wurden Männer und Frauen, im Alter von 19-78 Jahren mit Typ-2-Diabetes-Erkrankung und unzureichender glykämischer Kontrolle (definiert durch $HbA1c \geq 7$ und $< 10\%$), die entweder bei Eintritt in die Studie oder nach einer Dosisfindungsphase mit einer stabilen Dosis Metformin als Monotherapie von mindestens 1500 mg/Tag behandelt wurden. Patienten, die kein orales antihyperglykämisch Medikament (OHA), eine andere OHA-Monotherapie oder Metformin mit einem anderen OHA verabreicht bekamen, konnten zur Teilnahme an der Studie ausgewählt werden, wenn ihre HbA1c-Werte den Auswahlkriterien entsprachen.</p> <p>Ausschlusskriterien:</p> <p>Patienten mit Typ-1-Diabetes, Insulintherapie innerhalb 8 Wochen vor Auswahl für die Studie, Verschlechterung der Nierenfunktion, Nüchternblutzuckerwert (FPG) $> 14,4$ mmol/L zum Zeitpunkt der Randomisierung</p> <p>Während der Studie war die Einnahme anderer OHA verboten, erlaubt war die Behandlung mit fettspiegel- und blutdrucksenkenden Medikamenten, Schilddrüsenhormonersatztherapie und empfängnisverhütenden Medikamenten, ohne</p>

Item ^a	Charakteristikum	Studieninformation
		dabei ihre Dosierung zu verändern.
4b	Studienorganisation und Ort der Studiendurchführung	nicht berichtet
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Sitagliptin 100 mg täglich bzw. Placebo einmal täglich zusätzlich zur gewohnten Dosis Metformin
6	Zielkriterien	Veränderungen in den HbA1c-Werten zwischen Ausgangsuntersuchung und Untersuchung nach 24 Wochen in Prozenten
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Veränderungen in den HbA1c-Werten zwischen Ausgangsuntersuchung und Untersuchung nach 24 Wochen in Prozenten Sekundäre Zielkriterien: Veränderungen zwischen Ausgangsuntersuchung und Untersuchung nach 24 Wochen bei : Nüchternplasmaglukosewerten (FPG), Glukose-, Insulin-, C-Peptidkonzentrationen, gemessen unmittelbar vor, und 60 bzw. 120 min nach einer Standardmahlzeit. Total Cholesterin, Triglyzeriden, LDL-, HDL-, Nicht-HDL-Cholesterin und das Verhältnis von Triglyzeriden zu HDL-Cholesterin. Bestimmt wurden sowohl Mittelwerte für Glukose, Insulin, C-Peptide als auch <i>area under curve</i> (AUC) Werte bzw. das Verhältnis Insulin AUC zu Glukose AUC nach einer standardisierten Morgenmahlzeit.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurden keine Veränderungen vorgenommen
7	Fallzahl	Anzahl der randomisierten Teilnehmer: Saxagliptin n=464, für Placebo n= 237
7a	Wie wurden die Fallzahlen bestimmt?	nicht berichtet
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	nicht berichtet
8a	Methode zur Generierung der zufälligen Zuteilung	nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	nicht berichtet

Item ^a	Charakteristikum	Studieninformation
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar: die Autoren beschreiben die Studie als verblindet, beschreiben aber nicht wer verblindet wurde und wie die Verblindung durchgeführt wurde,
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht zutreffend
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die Wirksamkeitsanalyse basierte auf der Patientenpopulation, die nach Randomisierung mindestens eine Dosis des Studienmedikaments erhalten und sowohl an der Eingangsuntersuchung wie an mindestens einer weiteren Untersuchung teilgenommen haben.</p> <p>Kontinuierliche Wirksamkeitsparameter wurden mittels ANCOVA-Modell analysiert, wobei in den Behandlungsgruppen Veränderungen zwischen Eingangswerten (Tag1/ Randomisierung) und Werten in Woche 24 verglichen wurden. Die Analysen wurden bezüglich Ausgangswerten und vorausgegangener oder nicht durchgeführten hyperglykämischen Therapien angepasst. Fehlende Daten wurden gemäß LOCF behandelt. Um die Beeinflussung von Wirksamkeitsvergleichen durch die Behandlung nach Notfallindikation auszuschließen, wurden Daten nach Notfallindikation wie fehlende Daten behandelt.</p> <p>Die primäre Hypothese für diese Wirksamkeitsstudie war: die Gabe von 100 mg Sitagliptin führt, verglichen mit Placebo-Einnahme zu einer größeren Reduktion der HbA1c-Werte in Woche 24. Bestimmt durch die statistische Signifikanz der Differenz bei den kleinsten Quadraten (<i>least-squares mean change</i>) zwischen HbA1c-Ausgangswerten und Werten in Woche 24 Saxagliptin versus Placebo.</p> <p>Die Tests für die sekundären Endpunkte wurden in der Reihenfolge FPG, Glukose 2h nach der Mahlzeit durchgeführt, vorausgesetzt die primäre Wirksamkeitshypothese war erfüllt. Ferner wurde das Verhältnis der Patienten in jeder Behandlungsgruppe, die das Ziel HbA1c-</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Wert <7% in Woche 24 erreichten, bestimmt. Eine Zeit-bis-glykämischer Notfall-Analyse wurde mittels Kaplan-Meier-Schätzung und <i>log-rank</i> Test durchgeführt. Außerdem wurde das Verhältnis der Patienten in jeder Behandlungsgruppe bestimmt, die während der Studie eine glykämische Notfalltherapie erhielten.</p> <p>Sicherheitsanalysen wurden bei allen randomisierten Patienten durchgeführt, die mindestens eine Dosis des Studienmedikaments/Placebo erhalten hatten.</p> <p>Interferenzttests wurden durchgeführt um statistische Signifikanzlevel zu Gruppenunterschieden zu bestimmen und zwar für Hypoglykämie, vorher festgelegte spezielle gastrointestinale unerwünschte Ereignisse (abdominale Schmerzen, Diarrhö, Übelkeit und Erbrechen) und Veränderungen in Körpergewicht</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	nicht zutreffend
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 21)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 21)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 21)
14	Aufnahme / Rekrutierung	nicht berichtet.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Nach Studienaufnahme fanden Untersuchungen in Woche 0, 6, 12, 18 und 24 statt
14b	Informationen, warum die Studie endete oder beendet wurde	nicht berichtet
a: nach CONSORT 2010.		

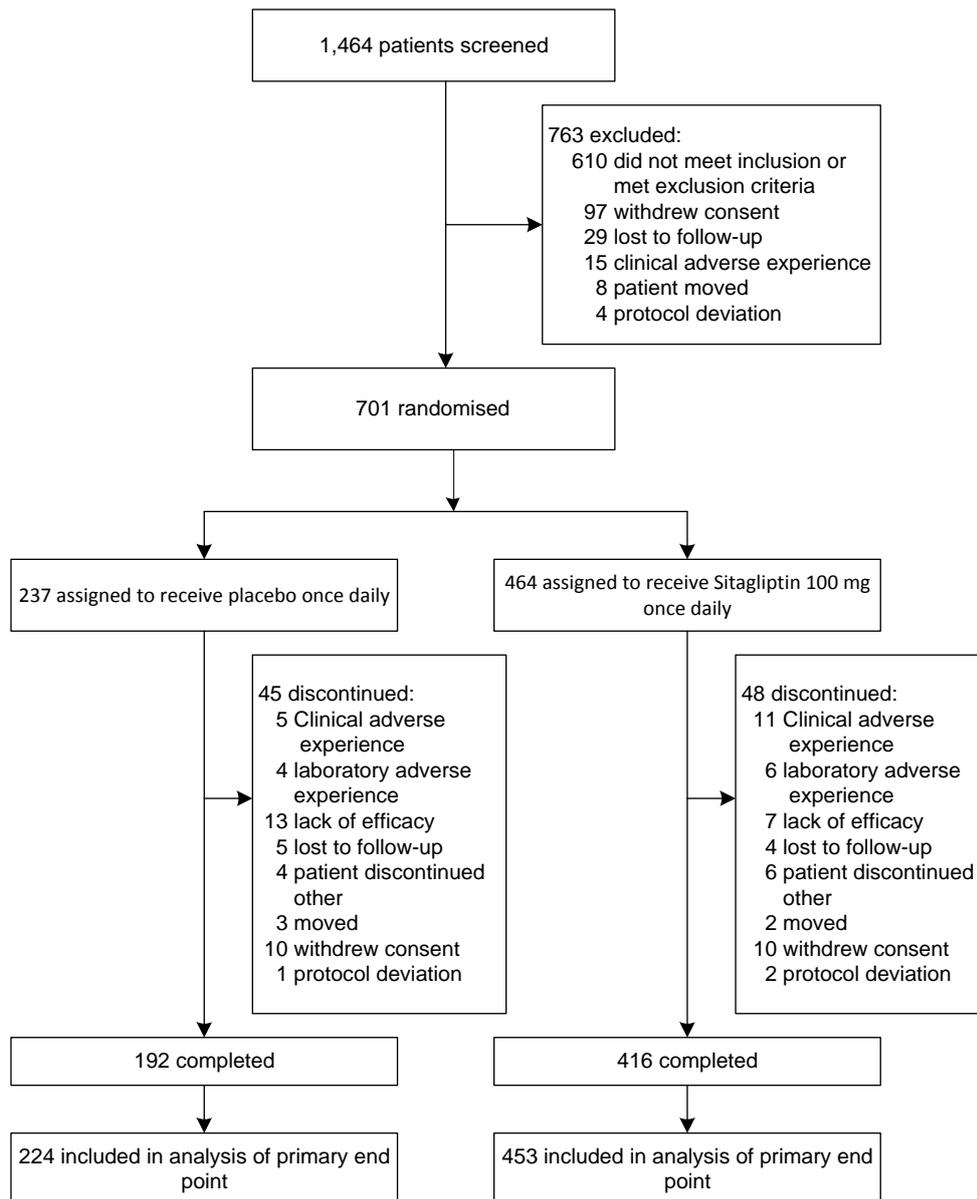


Abbildung 21: Flussdiagramm Studiendesign und -methodik für die Studie von Charbonnel 2006

Tabelle 4-163 (Anhang): Studiendesign und -methodik für die Studie von DeFronzo, 2009

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Diese 24-wöchige Studie untersuchte die Wirksamkeit und Sicherheit von Saxagliptin als <i>add on</i> Therapie bei Patienten mit Typ-2-Diabetes, die durch Metformin allein nicht ausreichend kontrolliert waren.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, placebokontrollierte Parallel-(1:1:1:1) Wirksamkeits/Sicherheitsstudie ^b
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Die Methodik wurde im Verlauf der Studie nicht verändert.
4	Probanden / Patienten	Typ-2-Diabetiker, die durch Metformin allein unzureichend therapiert sind
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Männer und Frauen im Alter von 18 bis 77 Jahren mit unzureichender Blutzuckerkontrolle (HbA1c $\geq 7,0$ und $\leq 10,0\%$), vorbehandelt mit einer stabilen Dosis Metformin (≥ 1500 g aber nicht >2550 mg/Tag) für mindestens 8 Wochen vor der Eingangsuntersuchung, Nüchtern-C-Peptid-Konzentration $\geq 1,0$ ng/mL, BMI ≤ 40 kg/m²</p> <p>Ausschlusskriterien.</p> <p>Patienten wurden von der Studie ausgeschlossen, wenn sich ein oder mehrere der folgenden Symptome einer schlecht eingestellten Diabetes zeigten:</p> <p>Diabetische Ketoazidose, hyperosmolares nichtketonisches Koma, Gebrauch eines anderen antihyperglykämischen Medikaments (bis 8 Wochen vor Studienbeginn) oder Insulin (1 Jahr vor Studienbeginn), kardiovaskuläre Erkrankung (6 Monate vor Studienbeginn), Herzinsuffizienz (gemäß <i>New York Heart Association stage III/IV</i>), in EKG linke ventrikuläre Auswurfraction $\leq 40\%$, chronischer oder wiederholter Gebrauch von Kortikosteroiden, Alkohol oder Drogenmissbrauch innerhalb des vergangenen Jahres, Behandlung mit systemischen Cytochrom P 450 3A4-Inhibitoren oder Induktoren, aktuelle Lebererkrankung, klinisch relevante Anomalitäten in der Eingangsuntersuchung bei Leber- und Nierenwerten, endokrinen, metabolischen, haematologischen, immunologischen Parametern. Ebenso wurden Schwangere und stillende Frauen von</p>

Item ^a	Charakteristikum	Studieninformation
		Studie ausgeschlossen.
4b	Studienorganisation und Ort der Studiendurchführung	152 Zentren (92 in USA, 23 in Kanada, 9 in Mexiko, 8 in Argentinien, 7 in Brasilien, 5 in Puerto Rico, 4 in Australien, 2 in Taiwan, 2 in Chile)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Saxagliptin 2,5 mg, Saxagliptin 5 mg, Saxagliptin 10 mg, Placebo alle zusammen mit Metformin (≥ 1500 mg aber nicht >2550 mg) einmal am Tag
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Veränderung des HbA1c-Wertes zwischen der Eingangsuntersuchung und nach 24 Wochen Sekundäre Zielkriterien: Veränderungen zwischen Eingangsuntersuchung und 24 Wochen-Werten bei: Nüchternblutglukose (FPG), Prozentsatz der Patienten mit HbA1c $< 7,0\%$, postprandiale Glukose (PPG) 3-h AUC während eines 75-g oralen Glukosetoleranztest (OGTT), der orale Glukosetoleranztest wurde 30 min nach Verabreichung der Studienmedikamente durchgeführt. Andere Wirksamkeitskriterien beinhalteten 2-h postprandiale Plasmaglukose (OGTT), nüchtern und postprandiale Glukagon, Insulin C-Peptiden HOMA-2 abgeleiteten Indizes für Insulinresistenz und β -Zellfunktion (Matthews 1985) Indizes für Insulin und β -Zellsensitivität abgeleitet vom OGTT (Mari 2001, Phillips 1994)
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurden keine Änderungen vorgenommen.
7	Fallzahl	2,5 mg Saxagliptin einmal täglich n=192 5 mg Saxagliptin einmal täglich n=191 10 mg Saxagliptin einmal täglich n=181 Placebo einmal täglich n=179
7a	Wie wurden die Fallzahlen bestimmt?	153 Patienten pro Behandlungsarm bedeuten mindestens 90% statische Power um Unterschiede in HbA1c-Mittelwerten zwischen den verschiedenen Behandlungsgruppen von 0,5% zu detektieren (SD 1,2%). Bei einer angenommenen Studienabbruchrate von 15% müssen somit 720 Patienten randomisiert werden (180 pro Behandlungsarm)
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	nicht berichtet
8	Randomisierung, Erzeugung der Behandlungsfolge	

Item ^a	Charakteristikum	Studieninformation
8a	Methode zur Generierung der zufälligen Zuteilung	<i>Interactive voice response</i> System
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Blockmutationen
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Das Vorgehen zur Randomisierung sowohl für die Patienten wie auch für die Behältnisse wurde im Randomisierungszentrum innerhalb des <i>Drug Supply Management Department</i> von BMS durchgeführt. Die Kit-Nummerierung war nicht fortlaufend und entsprach den aufgedruckten Nummern der Verpackungen und Flaschen der Studienmedikamente.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Prüfarzt, BMS-Personal und Studienteilnehmer waren während der Studiendauer verblindet. HbA1c-, Fibrinogen-, hsCRP- und IL6-Werte in den Laborberichten waren während der Studiendauer für die Studienbeteiligten verdeckt.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Saxagliptin-Tabletten und Placebo-Verabreichung waren äusserlich identisch
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Wirksamkeitsanalysen wurden mit der randomisierten Patientenpopulation durchgeführt, die mindestens eine Dosis des Studienmedikamentes erhalten hatte und bei der mindestens eine Untersuchung nach der Eingangsuntersuchung stattgefunden hat. Jede Saxagliptin-Gruppe wurde mit der Placebo-Gruppe bezüglich Veränderungen der Studienmessgrößen mittels ANCOVA verglichen, dabei dienten der Behandlungsarm als Effektor und Ausgangswerte als Kovariate. Punktschätzer und 95% Konfidenzintervall innerhalb eines Behandlungsarmes, sowie zwischen den Behandlungsgruppen in Woche 24 wurden mit der Methode der kleinsten Quadrate berechnet. Sequentielle Tests wurden zur Analyse sekundären Endpunkte angewendet. Andere kontinuierliche Wirksamkeitsvariablen wurden durch deskriptive statistische Methoden beschrieben. Der Prozentsatz der Patienten in den verschiedenen Behandlungsarmen, die nach 24 Wochen einen therapeutischen <i>Response</i> zeigten, wurden mit Hilfe des Fisher-Testes verglichen. Um fehlende Daten zu berücksichtigen, wurde die LOCF-

Item ^a	Charakteristikum	Studieninformation
		Methode verwendet. Wirksamkeits- und Sicherheitsmessungen nach einer Notfallintervention wurden nicht berücksichtigt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 22)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 22)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 22)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Studie begann im Juli 2005 und endete im August 2006. Messungen wurden alle 2 Wochen in der ersten 8 Wochen und danach alle 4 Wochen vorgenommen
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde protokollgemäß nach 24 Wochen beendet.
a: nach CONSORT 2010.		

^aRandomisierungsverhältnis 1:1:1:1 Saxagliptin (2,5 mg, 5 mg, 10 mg Placebo) zusätzlich zu 1500–2550 mg Metformin einmal täglich

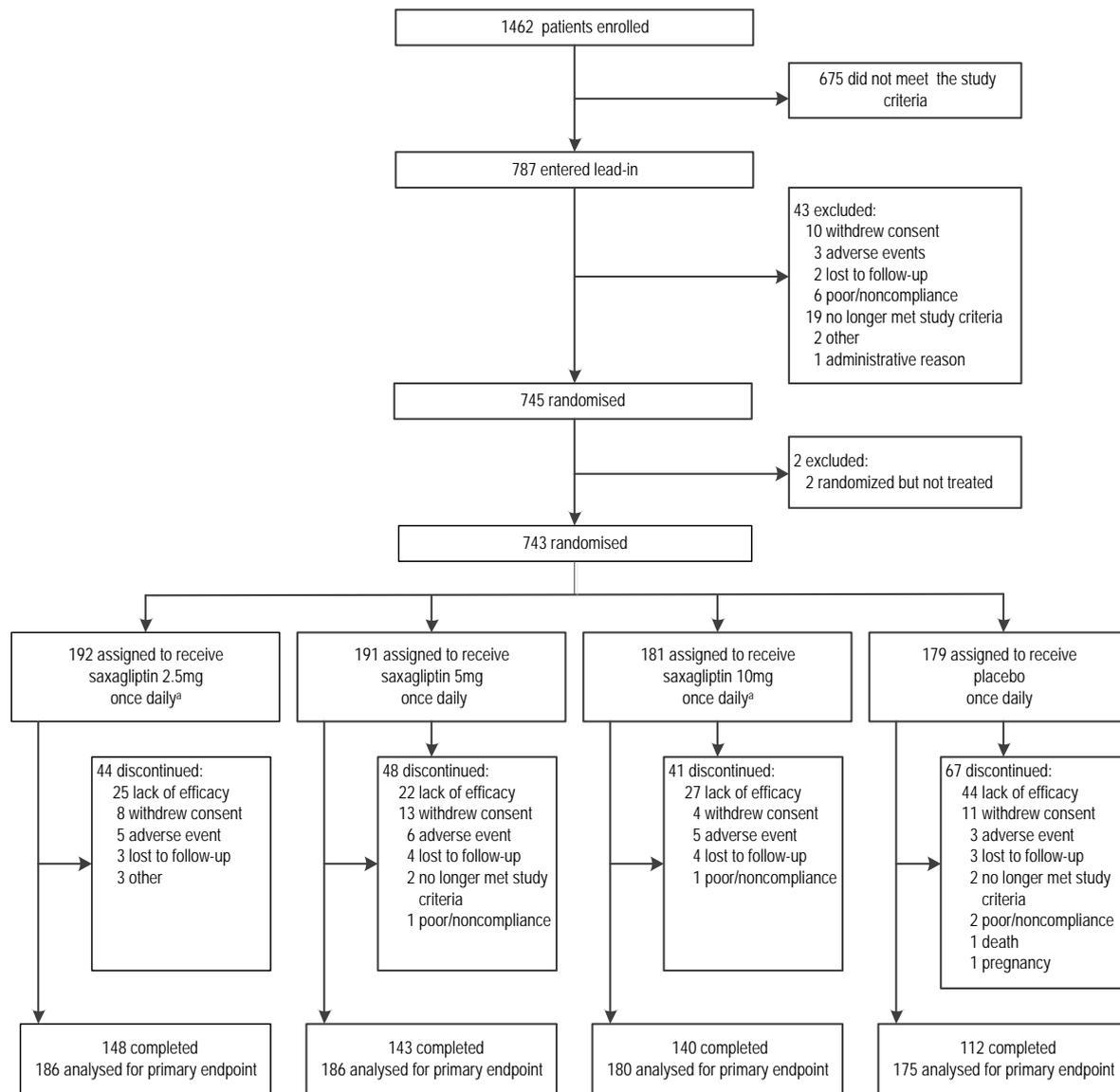


Abbildung 22: Flussdiagramm Studiendesign und -methodik für die Studie von DeFronzo, 2009

Anmerkung: ^a Dieser Arm war nicht in die Analyse eingeschlossen

Tabelle 4-164 (Anhang): Studiendesign und -methodik für die Studie von Bosi, 2007

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Studienziel war die Wirksamkeit und Sicherheit von Vildagliptin, einem neuen Dipeptidyl-Peptidase-4-Inhibitor zu bewerten. Vildagliptin wurde als Zusatztherapie zu Metformin bei Patienten mit Typ-2-Diabetes für 24 Wochen verabreicht.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Die Studie war eine multizentrische, randomisierte, doppelblinde, parallel angelegte (im Verhältnis 1:1) ^b , placebokontrollierte Wirksamkeits- und Sicherheitsstudie.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderungen wurden nach dem Studienbeginn vorgenommen
4	Probanden / Patienten	Typ-2-Diabetiker mit unzureichender glykämischer Kontrolle unter Metformin-Behandlung
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Patienten mit Typ-2-Diabetes, die Metformin als Monotherapie für mindestens 3 Monate erhalten hatten und die mit einer stabilen Dosis von ≥ 1500 mg täglich für mindestens 4 Wochen vor der Eingangsvisite behandelt wurden. Die Patienten sollten einen HbA1c-Wert im Bereich 7,5-11,0% bei der Eingangsuntersuchung haben, und wenn sie nicht zu dieser Zeit die maximal verträgliche Dosis erhielten, willigten sie ein, ihre Metformin-Dosis auf 2000 mg bei der ersten Visite zu erhöhen. Männer und Frauen (falls im gebärfähigen Alter, verwendeten sie eine medizinisch anerkannte Verhütungsmethode), im Alter von 18-78 Jahren, mit einem BMI im Bereich von 22–45 kg/m², und mit FPG <15 mmol/L waren teilnahmeberechtigt.</p> <p>Ausschlusskriterien:</p> <p>Patienten waren ausgeschlossen, wenn sie eins oder mehrere von den folgenden Symptomen hatten; Typ-1- oder eine andere Form von Diabetes, akute metabolische diabetische Komplikationen innerhalb der vergangenen 6 Monate, kongestive Herzinsuffizienz, die pharmakologische Behandlung erforderte, Herzinfarkt, instabile Angina oder koronare Bypassoperation innerhalb der vergangenen 6 Monate. Lebererkrankung wie Cirrhosis oder chronische aktive Hepatitis ebenfalls schlossen die Teilnahme aus, wie auch</p>

Item ^a	Charakteristikum	Studieninformation
		Nierenerkrankung oder –funktionsstörung mit Serumkreatininspiegel ≥ 132 $\mu\text{mol/L}$ (Männer) oder ≥ 123 $\mu\text{mol/L}$ (Frauen).
4b	Studienorganisation und Ort der Studiendurchführung	International in 109 Zentren, 79 in USA, 8 in Frankreich, 6 in Italien und 16 in Schweden.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Vildagliptin 100 mg täglich (in gleichmässig verteilten Dosen), oder Placebo, zusätzlich zu der gewohnten Dosis des nicht-verblindeten Metformins.
6	Zielkriterien	Änderung des HbA1c-Wertes zwischen Screening und Untersuchung nach 24 Wochen in Prozenten.
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Die Änderung des HbA1c-Wertes zwischen Ausgangswert und Studienendpunkt. Für Patienten, die vorzeitig abbrachen, wurden die letzten Beobachtungen übertragen. Alle Wirksamkeitsresultate beziehen sich auf die primäre ITT-Population. Sekundäre Wirksamkeitsparameter: FPG, Nüchternplasmalipide und Körpergewicht.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurden keine Änderungen der Zielkriterien nach dem Studienbeginn vorgenommen.
7	Fallzahl	Anzahl der randomisierten Teilnehmer: Vildagliptin 100 mg/Tag: n=185; Placebo: n=182.
7a	Wie wurden die Fallzahlen bestimmt?	Nicht berichtet
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Nicht berichtet
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die	Unklar: die Autoren beschreiben die Studie als verblindet, beschreiben aber nicht ausdrücklich wer verblindet wurde und wie

Item ^a	Charakteristikum	Studieninformation
	Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	die Verblindung durchgeführt wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Veränderungen der Basiswerte in primären und sekundären Endpunkten wurden mittels eines Modells der Kovarianzanalyse (ANCOVA) verglichen, mit Behandlung und gepoolte Studienzentren (<i>pooled centre</i>) als Klassifikationsvariablen und Basiswert als Kovariate. Die Analysen wurden mit zweiseitigen Tests und einer statistischen Signifikanz von 0,05 durchgeführt. Multiples Testen wurde mittels <i>Hochberg's multiple testing step-up</i> -Methode korrigiert, um eine zweiseitige Signifikanzniveau von 0,05 zu erhalten. Die berichteten Sicherheits- und Verträglichkeitsdaten schliessen alle Patienten ein, die mindestens eine Dosis der Studienmedikation erhalten hatten und bei denen mindestens eine Untersuchung nach der Eingangsuntersuchung stattgefunden hatte.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht zutreffend
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 23)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 23)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 23)
14	Aufnahme / Rekrutierung	Nicht berichtet
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Patienten wurden über 24 Wochen beobachtet, mit Untersuchungen nach Randomisierung in Wochen 0, 4, 12, 16 und 24.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

Anmerkungen: ^b, Randomisiert wurde zu zwei Dosen von Vildagliptin (50 mg, 100 mg) und Placebo, d. h. im Verhältnis 1:1:1, aber nur der Vildagliptin 100 mg-Arm und der Placebo-Arm wurden für die Nutzenbewertung einbezogen.

Abkürzungen: RCT, Randomisierte kontrollierte Studie; T2DM, Typ-2-Diabetes mellitus; BMI, *Body-Mass-Index*; HbA1c, Glykiertes Hämoglobin; FPG, Nüchternplasmaglukose; ITT, *Intention to treat*, US, United States.

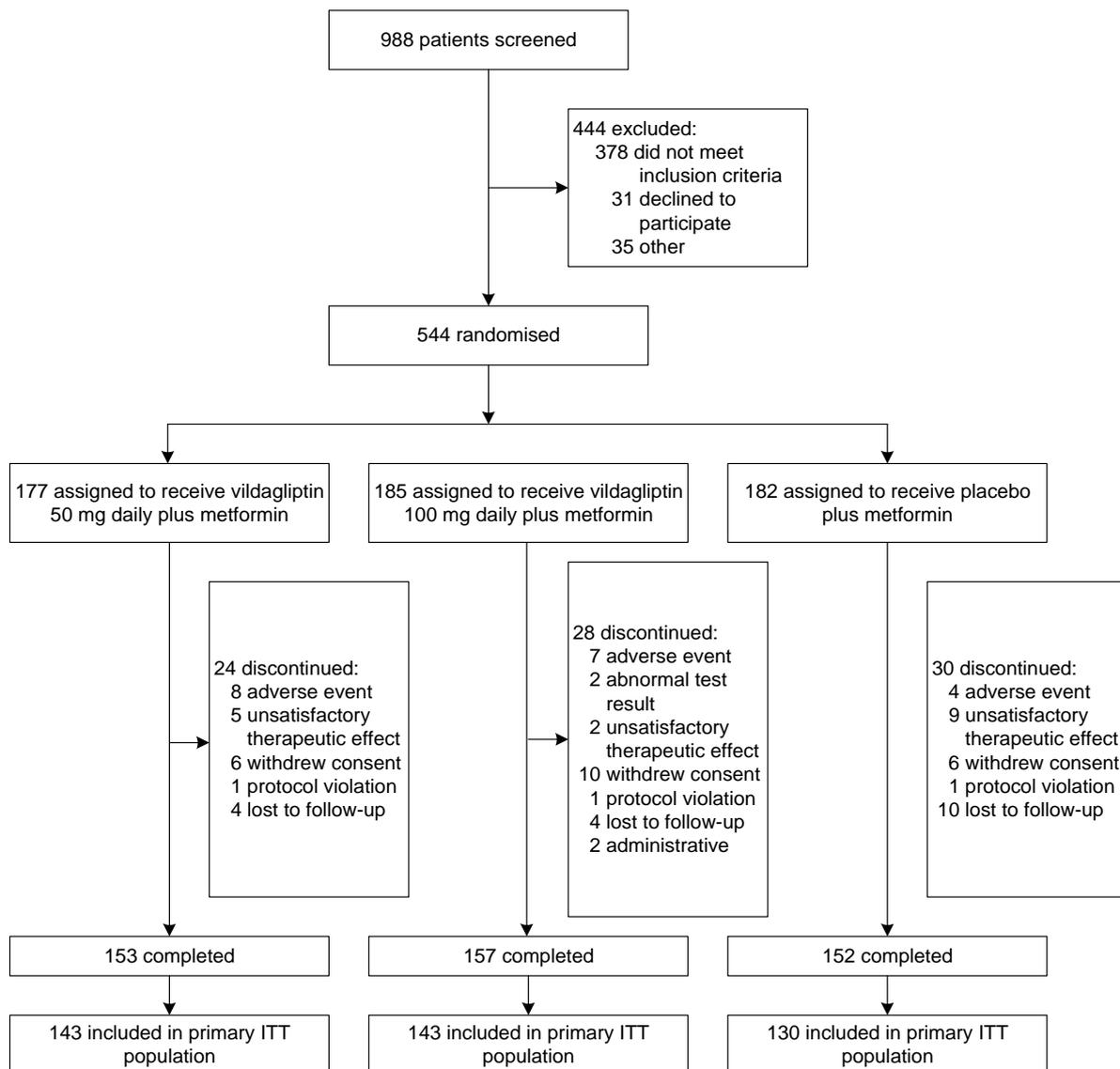


Abbildung 23: Flussdiagramm Studiendesign und -methodik für die Studie von Bosi 2007.

Anmerkungen: Randomisiert wurde zu zwei Dosen von Vildagliptin (50 mg, 100 mg) und Placebo, d. h. im Verhältnis 1:1:1, aber nur der Vildagliptin 100 mg-Arm und der Placebo-Arm wurden für die Nutzenbewertung einbezogen.

Abkürzung: ITT, Intention to treat.

Tabelle 4-165 (Anhang): Studiendesign und -methodik für die Studie von Taskinen 2011

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Ziel war es, die Wirksamkeit und Sicherheit des potenten und selektiven Dipeptidyl-Peptidase-4 (DPP-4) Inhibitors Linagliptin zu bewerten. Linagliptin wurde als Zusatztherapie zu Metformin bei Patienten mit Typ-2-Diabetes und unzureichender glykämischer Kontrolle verabreicht.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, placebokontrollierte. Parallelstudie (Verhältnis 3:1) zu Wirksamkeit und Sicherheit
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nach Studienbeginn erfolgten keine Änderungen
4	Probanden / Patienten	Typ-2-Diabetiker durch Metformin allein unzureichend kontrolliert
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Männer und Frauen im Alter von 18-80 Jahren mit Typ-2-Diabetes und einem BMI $\leq 40 \text{ kg/m}^2$, therapiert mit Metformin ($\geq 1500 \text{ mg/Tag}$, oder maximal tolerierte Dosis) und nicht mehr als einem anderen oralen Antidiabetikum. Das Antidiabetikum musste in unveränderter Dosis 10 Wochen vor Studienzustimmung verabreicht worden sein, die Dosis an Metformin sollte ≥ 12 Wochen vor Randomisierung stabil geblieben sein. Bei der Eingangsuntersuchung lag der Schwellenwert für die Aufnahme in die Studie für HbA1c bei 7,0-10,0% für die Patienten, die Metformin als Monotherapie erhalten hatten (ungefähr Zweidrittel der Patienten) oder bei 6,5-9,0% für die Patienten, die ein zusätzliches Medikament erhalten hatten. Nach Beginn der Placebo-Gabe lag der HbA1c-Schwellenwert für alle Patienten bei 7,0-10,0%.</p> <p>Ausschlusskriterien:</p> <p>Behandlung mit Rosiglitazon, Pioglitazon, glukose-ähnliche Peptid-1-Analogen, Insulin, oder Medikamenten gegen Fettleibigkeit innerhalb der letzten 3 Monaten, Änderungen in der Dosierung einer Schilddrüsenhormonbehandlung innerhalb der letzten 6 Wochen oder Behandlung mit systemischen Steroiden nach Studienzustimmung. Patienten wurden ausgeschlossen bei Leberfunktionsstörungen (Serumwerte von Alanin-Aminotransferase, Aspartat-Aminotransferase oder alkalischer</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Phosphatase mehr als dreifach über Normalwert, Nierenversagen oder Nierenerkrankung (Serumkreatinin $\geq 135 \mu\text{mol/L}$, Herzinfarkt, Schlaganfall oder transitorische ischämische Attacke innerhalb von 6 Monaten nach Studienzustimmung. Akute oder chronische metabolische Azidose, instabile oder akute Herzinsuffizienz, vererbte Galaktoseintoleranz, Dehydrierung. Patienten durften an keiner anderen Medikamentenstudie in den vergangenen zwei Monaten teilgenommen haben.</p>
4b	Studienorganisation und Ort der Studiendurchführung	82 Zentren in 10 Ländern (Tschechische Republik, Finnland, Griechenland, Indien, Israel, Mexiko, Neuseeland, Russland, Schweden und USA
5	<p>Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p>Linagliptin 5 mg einmal täglich oder Placebo zusammen mit Metformin ($\geq 1500 \text{ mg}$) einmal täglich</p> <p>Während der ersten 12 Wochen der Studie konnte eine Notfallbehandlung (Sulfonylharnstoff) durchgeführt werden, wenn der Patient einen bestätigten Glukosewert $>13.3 \text{ mmol/L}$ nach Fasten über Nacht hatte. Während den letzten 12 Wochen der Studie konnte die Notfallbehandlung ab einem Glukosewert von 11.1 mmol/L nach Fasten über Nacht oder einem zufällig festgestellten Wert von $>22,2 \text{ mmol/L}$ eingeleitet werden. Die Patienten die eine Notfallbehandlung erhielten, blieben in der Studie und wurden nicht entblindet.</p>
6	Zielkriterien	
6a	<p>Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten</p>	<p>Primäres Zielkriterium: Änderung des HbA1c-Wertes zwischen Ausgangsmessung und Messung in Woche 24 in Prozent</p> <p>Sekundäre Zielkriterien: Änderungen zwischen Ausgangswerten und Werten in Woche 24 für Nüchternserumglukose (FPG), 2h FPG, Veränderung der HbA1c- und FPG-Werte über die Zeit, Prozentsatz der Patienten, die während der Behandlung einen HbA1c-Wert $<7,0\%$ und $<6,5\%$ erreichten und der Prozentsatz der Patienten, die eine Verringerung des HbA1c-Wertes um $\geq 0,5\%$ erreichten.</p> <p>Andere Endpunkte waren das Verhältnis der Patienten, die eine Notfallbehandlung benötigten, die Veränderung des BMI im Vergleich zum Ausgangswert, Veränderungen bei Biomarkern für Insulin, β-Zellfunktion und postprandialer Glukosekontrolle.</p>

Item ^a	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nach Beginn der Studie wurden keine Änderungen der Zielkriterien vorgenommen
7	Fallzahl	Randomisiert wurden : Linagliptin 5 mg einmal täglich n=523 Placebo einmal täglich n=177
7a	Wie wurden die Fallzahlen bestimmt?	Eine Populationsgröße von 600 Patienten (450 Linagliptin und 150 Placebo) war erforderlich um eine statistische Power von 95% sicherzustellen, um damit einen Behandlungsunterschied von 0.6% bei einer Standardabweichung von 1,6% für die Veränderungen der HbA1c-Werte zu detektieren.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgt im Verhältnis 3:1 für Linagliptin oder Placebo, diese Allokation war stratifiziert durch das Ausmaß der glykämischen Kontrolle bei Start der Placebo-Gabe (HbA1c <8,5% oder ≥8,5%) entsprechend einer Monotherapie oder Kombinationstherapie bei Studienanmeldung
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar. Die Autoren beschreiben die Studie als doppelblind, berichten aber nicht explizit wer verblindet war oder wie die Verblindung durchgeführt wurde
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Der primäre Endpunkt: Veränderungen zwischen HbA1c-Ausgangswerten und Werten nach 24-wöchiger Behandlung wurden mittels ANCOVA-Modell mit den fixen Faktoren Behandlung, <i>wash-out</i> und HbA1c-Ausgangswert als lineare Kovariate,

Item ^a	Charakteristikum	Studieninformation
		<p>zwischen den Linagliptin- und Placebo-Gruppen verglichen. Die Analyse wurde mit der <i>full analysis set</i> Population, durchgeführt (fehlende Daten wurden gemäß der LOCF-Methode berücksichtigt). Diese Gruppe umfasst alle randomisierten Patienten, die mindestens eine Dosis des Studienmedikaments erhalten haben und neben der HbA1c-Ausgangswert-Messung an mindestens einer weiteren HbA1c-Bestimmung teilgenommen haben. Auswirkungen aller Methoden beim Umgang mit fehlenden Daten wurden mittels Sensitivitätsanalysen untersucht. ANCOVA und logistische Regressionsverfahren wurden verwendet, um alle kontinuierlichen und kategorischen sekundären und Sicherheitsendpunkte zu analysieren. Zusätzlich wurden Kaplan-Meier-Kurven erstellt um den Einfluss einer Notfallmedikamentierung zu beschreiben. Hauptsächlich wurde mit deskriptiven statistischen Methoden gearbeitet. Für die Bewertung der Sicherheit und von Biomarkern wurden keine fehlenden Daten eingebracht.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 24)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 24)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 24)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Daten zur Patientenrekrutierung wurden nicht berichtet. Die Studie dauerte 24 Wochen, Messungen wurden alle 6 Wochen durchgeführt.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete protokollgemäß nach 24 Wochen
a: nach CONSORT 2010.		

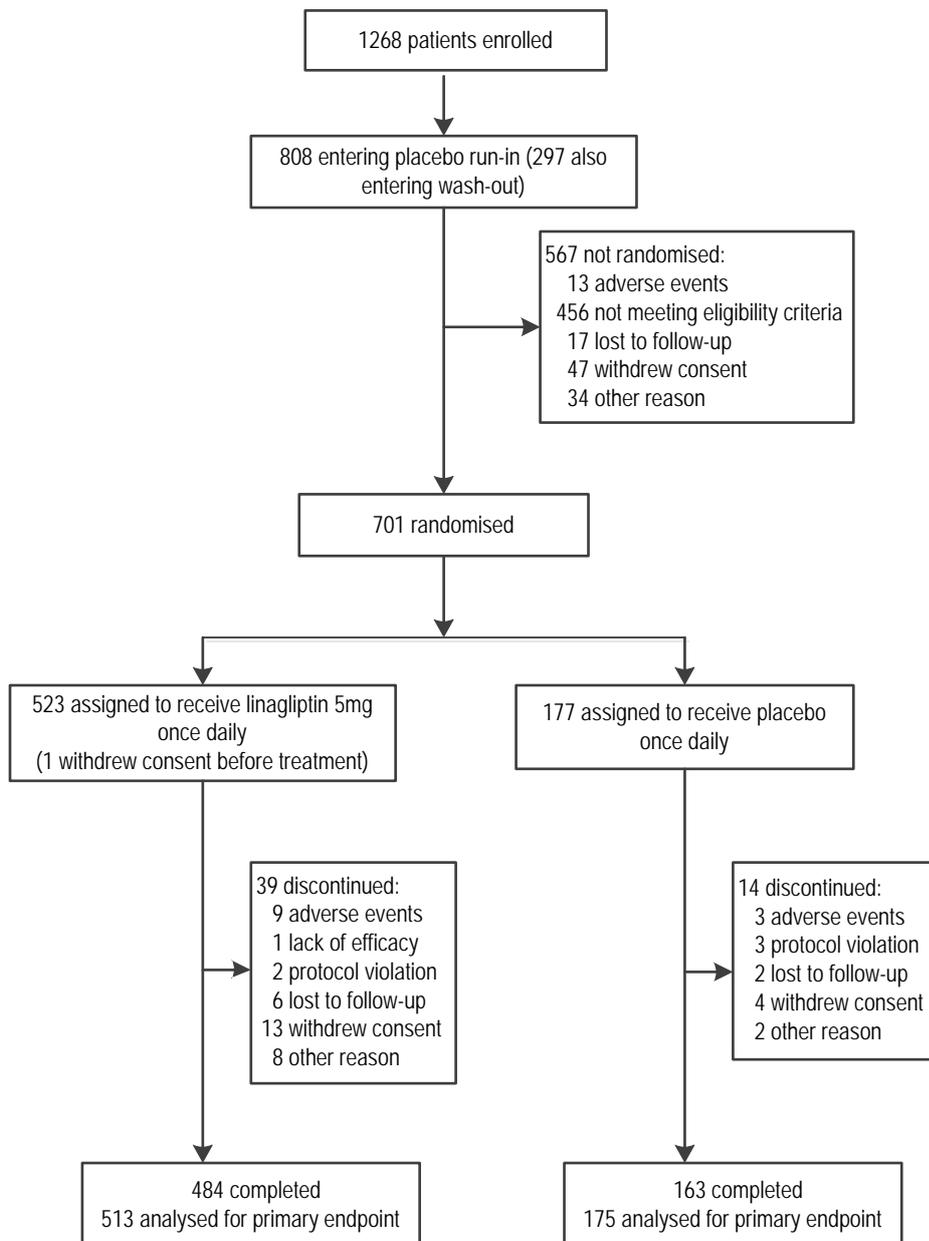


Abbildung 24: Flussdiagramm Studiendesign und -methodik für die Studie von Taskinen, 2011

Tabelle 4-166 (Anhang): Studiendesign und -methodik für die Studie von Bergenstal, 2010a

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Die Studie untersuchte die Wirksamkeit, Sicherheit und Verträglichkeit von Taspoglutid im Vergleich zu Sitagliptin und Placebo bei Patienten mit Typ-2-Diabetes, die durch Metformin allein nicht ausreichend kontrolliert waren.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multizentrische, randomisierte, double-dummy, placebo- und aktivkontrollierte parallel-geführte (im Verhältnis 2:1) ^b Sicherheits- und Wirksamkeitsstudie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderungen der Methodik wurden nach dem Studienbeginn vorgenommen
4	Probanden / Patienten	Typ-2-Diabetiker, die durch Metformin allein unzureichend kontrolliert waren
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Erwachsene im Alter von 18 bis 75 Jahren mit Typ-2-Diabetes, therapiert mit Metformin (≥ 1500 mg/Tag) für mindestens 12 Wochen, HbA1c 7-10% bei der Eingangsuntersuchung, BMI ≥ 25 (>23 für Asiaten) und ≤ 45 kg/m² bei der Eingangsuntersuchung, stabiles Körpergewicht +/- 5% für mindestens 12 Wochen vor der Eingangsuntersuchung.</p> <p>Ausschlusskriterien: Typ-1-Diabetes oder akute metabolische diabetische Komplikationen wie Ketoazidose oder hyperosmolares Koma in den vergangenen 6 Monaten; Hinweise auf klinisch relevante diabetische Komplikationen; klinisch symptomatische gastrointestinale Erkrankung; Herzinfarkt, koronare Bypassoperation, Kardiomyopathie nach Transplantation oder Schlaganfall in den vergangenen 6 Monaten; bekannte Hämoglobinopathie oder chronische Anämie.</p>
4b	Studienorganisation und Ort der Studiendurchführung	148 Studienorte
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Sitagliptin 100 mg oral täglich oder Placebo, zusätzlich zur gewohnten Dosis des nicht-verblindeten Metformins
6	Zielkriterien	Veränderung des HbA1c-Wertes zwischen der Eingangsuntersuchung und nach 24 Wochen in Prozenten
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden	Primäres Zielkriterium: Veränderung des HbA1c-Wertes zwischen der Eingangsuntersuchung und nach

Item ^a	Charakteristikum	Studieninformation
	(z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	24 Wochen. Sekundäre Zielkriterien: Veränderungen zwischen dem Ausgangswert und den Werten nach 24 Wochen bei Nüchternplasmaglukose, Körpergewicht und Lipidprofil; Ansprechraten für HbA1c (Ziel $\leq 7,0\%$, $\leq 6,5\%$) und Körpergewicht; β -Zellfunktion.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderungen wurden nach dem Studienbeginn vorgenommen
7	Fallzahl	Randomisierungszahlen wurden nicht berichtet
7a	Wie wurden die Fallzahlen bestimmt?	Nicht berichtet
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht berichtet
8	Randomisierung, Erzeugung der Behandlungsfolge	Nicht berichtet
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Patienten und Prüfer waren verblindet
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Nicht berichtet
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 25)
13a	Anzahl der Studienteilnehmer für jede durch	Siehe <i>Flow-Chart</i> (Abbildung 25)

Item ^a	Charakteristikum	Studieninformation
	Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 25)
14	Aufnahme / Rekrutierung	Die Studie begann im Oktober 2008 und endete am 17. Oktober 2011
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Patienten wurden über 24 Wochen beobachtet. Die Untersuchungen nach Randomisierung erfolgten in Woche 0, 4, 8, 12, 16 und 24.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

^b, Randomisierung erfolgte zu 4 Armen, d. h. Taspoglutid 10 mg, Taspoglutid 20 mg, Sitagliptin 100 mg und Placebo im Verhältnis 2:2:2:1, aber nur Sitagliptin- und Placeboarme wurden in die Nutzenbewertung einbezogen.

Abkürzungen: RCT, Randomisierte kontrollierte Studie; T2DM, Typ-2-Diabetes mellitus; BMI, *Body-Mass-Index*; HbA1c, Glykiertes Hämoglobin

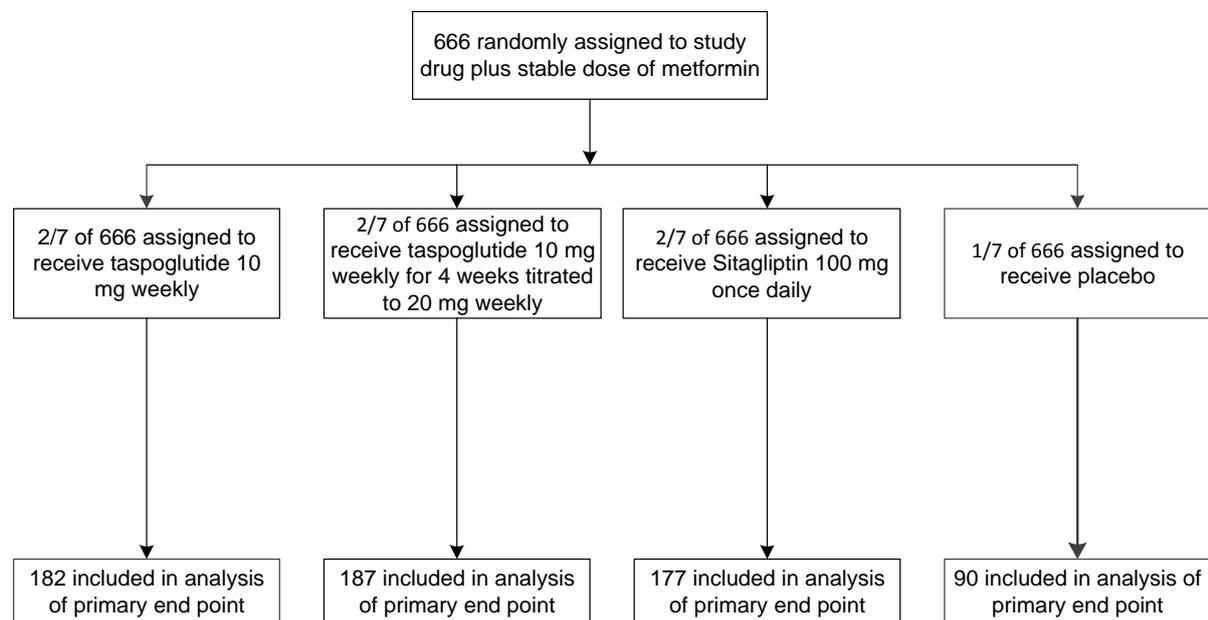


Abbildung 25: Flussdiagramm Studiendesign und -methodik für die Studie von Bergenstal 2010a.

Anmerkung: Randomisierung erfolgte zu 4 Armen, d. h. Taspoglutid 10 mg, Taspoglutid 20 mg, Sitagliptin 100 mg und Placebo im Verhältnis 2:2:2:1, aber nur Sitagliptin- und Placeboarme wurden in die Nutzenbewertung einbezogen.

Tabelle 4-167 (Anhang): Studiendesign und -methodik für die Studie von Raz, 2008

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Das Studienziel war die Wirksamkeit und Sicherheit von Sitagliptin als Zusatztherapie zu Metformin bei Patienten mit mittelschwerem (HbA1c $\geq 8,0\%$ und $\leq 11,0\%$) Typ-2-Diabetes zu bewerten.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, parallel angelegte (im Verhältnis 1:1) placebokontrollierte Wirksamkeits- und Sicherheitsstudie ^b
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderungen der Methodik wurden nach dem Studienbeginn vorgenommen
4	Probanden / Patienten	Typ-2-Diabetiker mit unzureichender glykämischer Kontrolle unter Metformin-Behandlung
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Ambulante Typ-2-Diabetes-Patienten, Männer und Frauen, im Alter von 18-78 Jahren, die zur Zeit Monotherapie mit Metformin oder einem anderen OHA erhielten, oder mit einer Kombination von Metformin und einem anderen OHA behandelt wurden, konnten zur Teilnahme an der Studie ausgewählt werden, wenn ihre HbA1c-Werte den Auswahlkriterien entsprachen – am Ende einer <i>run-in</i>-Periode, Patienten, die einen HbA1c-Wert von 8,0-11,0% hatten, waren wählbar für eine 2-wöchige, einfachblinde Placebo-<i>run-in</i>-Periode. Nach Abschluß dieser Periode, wurden Patienten randomisiert, die adäquate Therapietreue gezeigt hatten (hatten $\geq 85\%$ von den zur Verfügung gestellten Tabletten eingenommen) und Nüchtern-<i>fingerstick</i>-Glukose $\geq 7,2$ mmol/L und $\leq 15,6$ mmol/L hatten.</p> <p>Ausschlusskriterien: Behandlung mit Insulin während 8 Wochen vor dem Screening, Behandlung mit einem PPARγ-Medikament (z. B. Pioglitazon oder Rosiglitazon) oder mit GLP-1-Agonisten (z. B. Exenatide) während 12 Wochen, Typ-1-Diabetes, BMI < 20 kg/m² oder > 43 kg/m², oder Nüchternplasmaglukose (FPG), das konsequent < 7.2 mmol/L oder > 15.6 mmol/L während der <i>run-in</i>-Periode war. Gebrauch von einem anderen OHA neben Metformin war nicht während der Studie erlaubt. Es war den Patienten erlaubt, stabile Dosen von lipidsenkenden und antihypertensiven</p>

Item ^a	Charakteristikum	Studieninformation
		Medikamenten, Schilddrüsenhormonen und hormonellen Kontrazeptiva zu erhalten. Schwangere und stillende Patienten waren von der Studie ausgeschlossen.
4b	Studienorganisation und Ort der Studiendurchführung	Multinational
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Sitagliptin 100 mg einmal täglich; Placebo. Beide mit Metformin (1500-2550 mg) einmal täglich. Beim Eintritt in die <i>run-in</i> -Periode, wurden die Patienten, die ein anderes OHA erhielten, auf Monotherapie mit Metformin gesetzt, das dann auf eine Dosis von mindestens 1500 mg/Tag auftitriert wurde (Maximum, 255 mg/Tag). Diese Patienten traten dann für mindestens 6 Wochen in eine Diät- und Sportübungsperiode mit einer stabilen Metformin-Dosis ein. Patienten, die schon Monotherapie mit Metformin mit einer stabilen Dosis von mindestens 1500 mg/Tag erhielten, traten direkt in die 6-wöchige Diät- und Sportübungsperiode ein.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäres Zielkriterium: Änderung des HbA1c-Wertes zwischen Eingangsuntersuchung und Untersuchung nach 18 Wochen in Prozenten.</p> <p>Sekundäre Zielkriterien: Änderung von FPG und von 2-h postprandialer Plasmaglukose (gemessen nach einer Standardmahlzeit) zwischen Ausgangswert und dem Wert nach 18-wöchiger Behandlung, und die Änderung des HbA1c-Wertes zwischen Eingangsuntersuchung und Untersuchung nach 30-wöchiger Behandlung. Weitere Endpunkte waren Nüchtern- und Postprandialblutspiegel von Insulin, C-Peptid, Proinsulin, und das Proinsulin-Insulin-Verhältnis in Woche 18, <i>area under curve</i> (AUC) für Glukose, Insulin, C-Peptid und das Insulin-Glukose-AUC-Verhältnis.</p> <p>Von diesen, die β-Zellfunktion und Proinsulin-Insulin-AUC-Verhältnis wurden mittels Homöostase-Modell berechnet, um die β-Zellfunktion zu bewerten, Insulinresistenz mittels Homöostase-Modell (HOMA-IR) und <i>quantitative insulin sensitivity check index</i> wurden berechnet, um Insulinsensitivität zu bewerten. Zusätzlich, die Anzahl der Patienten, die Notfalltherapie erforderten und die Zeiten, an denen die Notfalltherapie erfolgte, wurden bewertet, wie auch die Prozentzahl der Patienten, die das therapeutische Ziel HbA1c <7%</p>

Item ^a	Charakteristikum	Studieninformation
		erreichten.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurden keine Änderungen der Zielkriterien nach dem Studienbeginn vorgenommen.
7	Fallzahl	Randomisierungszahlen: Sitagliptin 100 mg einmal täglich: n=96; Placebo einmal täglich: n=94
7a	Wie wurden die Fallzahlen bestimmt?	Mit Annahme einer gruppeninternen Standardabweichung von 1% für die HbA1c-Messungen (Charbonnel 2006), und einer mit einem zweiseitigen Test berechneten Signifikanz $\alpha=0.05$, würden ca. 86 Patienten pro Behandlungsgruppe mit 90% statistischer <i>Power</i> den Beweis für einen wahren Zwischen-Gruppenunterschied von 0,5% in der Mittelwertänderung vom Ausgangswert zeigen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Computergenerierter Plan
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar: die Autoren beschreiben die Studie als verblindet, beschreiben aber nicht ausdrücklich wer verblindet wurde und wie die Verblindung durchgeführt wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Wirksamkeit wurde mittels <i>full-analysis set</i> (FAS) analysiert. Diese Patientenpopulation umfasste die Patienten, die nach Randomisierung mindestens eine Dosis Sitagliptin oder Placebo erhalten und sowohl an der Eingangsuntersuchung als auch an mindestens einer weiteren Untersuchung teilgenommen hatten. Die erste Messung von HbA1c nach

Item ^a	Charakteristikum	Studieninformation
		<p>Randomisierung erfolgte in Woche 6; deshalb hatten alle Patienten in der FAS-Population die doppelblinde Behandlung mindestens für 6 Wochen erhalten.</p> <p>Kontinuierliche Wirksamkeitsparameter wurden mittels Modell zur Kovarianzanalyse (ANCOVA) analysiert, wobei in den Behandlungsgruppen Veränderungen zwischen Ausgangswerten und Werten in Wochen 18 und 30 verglichen wurden, mit Therapiezuordnung, Ausgangswerte und vor-OHA-Status als Kovariaten. Die Zwischen-Gruppenunterschiede für Wirksamkeitsendpunkte wurden durch das Testen der Differenz bei den kleinsten Quadraten (<i>least-squares mean change</i>) bestimmt. Die fehlenden Daten wurden nach der Methode der letzten Beobachtung (<i>last-observation-carried-forward</i>) behandelt.</p> <p>Mit sekundären Hypothesen verbundene Endpunkte (Überlegenheit von FPG und 2h-PPG-Senkung in Woche 18 und von HbA1c-Senkung in Woche 30) wurden bedingt mittels <i>step-down</i>-Verfahren getestet, das die Multiplizität korrigierte (Bauer 1991), und sie wurden bevorzugt in der Reihenfolge 2h-PPG in Woche 18, FPG in Woche 18, und dann HbA1c in Woche 30 getestet, vorausgesetzt, dass der primäre Wirksamkeitsendpunkt für HbA1c in Woche 18 erfüllt war. Der Anteil der Patienten, die das primäre HbA1c-Ziel <7,0% in Woche 18 und Woche 30 erreichten, wurde zwischen den Gruppen mit dem logistischen Regressionsmodell verglichen. Der Zwischen-Gruppenunterschied in den Anteilen der Patienten, die glykämische Notfalltherapie benötigten, wurde mittels Kaplan-Meier-Schätzung und log-rank-Test ermittelt. Die Ergebnisse wurden als Mittelwert±SD präsentiert, wenn nicht anders angegeben.</p> <p>Sicherheit und Verträglichkeit wurden analysiert mit Population bestehend von allen Patienten, die mindestens eine Dosis der doppelblinden Studienmedikation erhalten hatten. In die Sicherheitsanalyse wurden alle Nebenwirkungen, Laborsicherheitsmessungen, Gewicht und Vitalfunktionen eingeschlossen. Interferenztest wurde durchgeführt, um statistische Signifikanz für die Zwischen-Gruppenunterschiede für Hypoglykämie und vorher festgelegte spezielle gastrointestinale unerwünschte Ereignisse (Übelkeit, Erbrechen und abdominale Schmerzen) zu bestimmen.</p> <p>Die Bestimmungsmethode war exakter Test</p>

Item ^a	Charakteristikum	Studieninformation
		von Fisher, die Konfidenzintervalle wurden mit Wilsons Methode ermittelt (Newcombe 1998).
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Prädefinierte Subgruppen wurden bezüglich Änderungen des HbA1c-Wertes zwischen Eingangsuntersuchung und Untersuchung nach 18 und 30 Wochen analysiert, um die bereits bestehenden Risikofaktoren einzuschätzen, die potenziell das Behandlungsergebnis beeinflussen könnte. Diese Subgruppen waren definiert durch Unterschiede in Alter, Geschlecht, ethnischer Zugehörigkeit, Dauer des Diabetes, Therapie vor OHA, und Ausgangswerten von BMI, HbA1c, HOMA-β, HOMA-IR, und Proinsulin/Insulin-Verhältnis.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 26)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 26)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 26)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Rekrutierungsdaten wurden nicht berichtet; die Studiendauer war 30 Wochen; Messungen wurden alle sechs Wochen vorgenommen.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend – die Studie endete nach 30 Wochen gemäß Protokoll
a: nach CONSORT 2010.		

Anmerkungen: ^b Randomisierung war Sitagliptin (100 mg/Tag) plus Metformin oder Placebo+Metformin;

Abkürzungen: ANCOVA, Kovarianzanalyse; AUC *area under the curve* - Fläche unter der Kurve; CI, Konfidenzintervall; FAS, *full analysis set*; FPG, Nüchternplasmaglukose; HbA1c, Glykiertes Hämoglobin; OHA, antihyperglykämischer Wirkstoff; RCT, Randomisierte kontrollierte Studie; SD, Standardabweichung

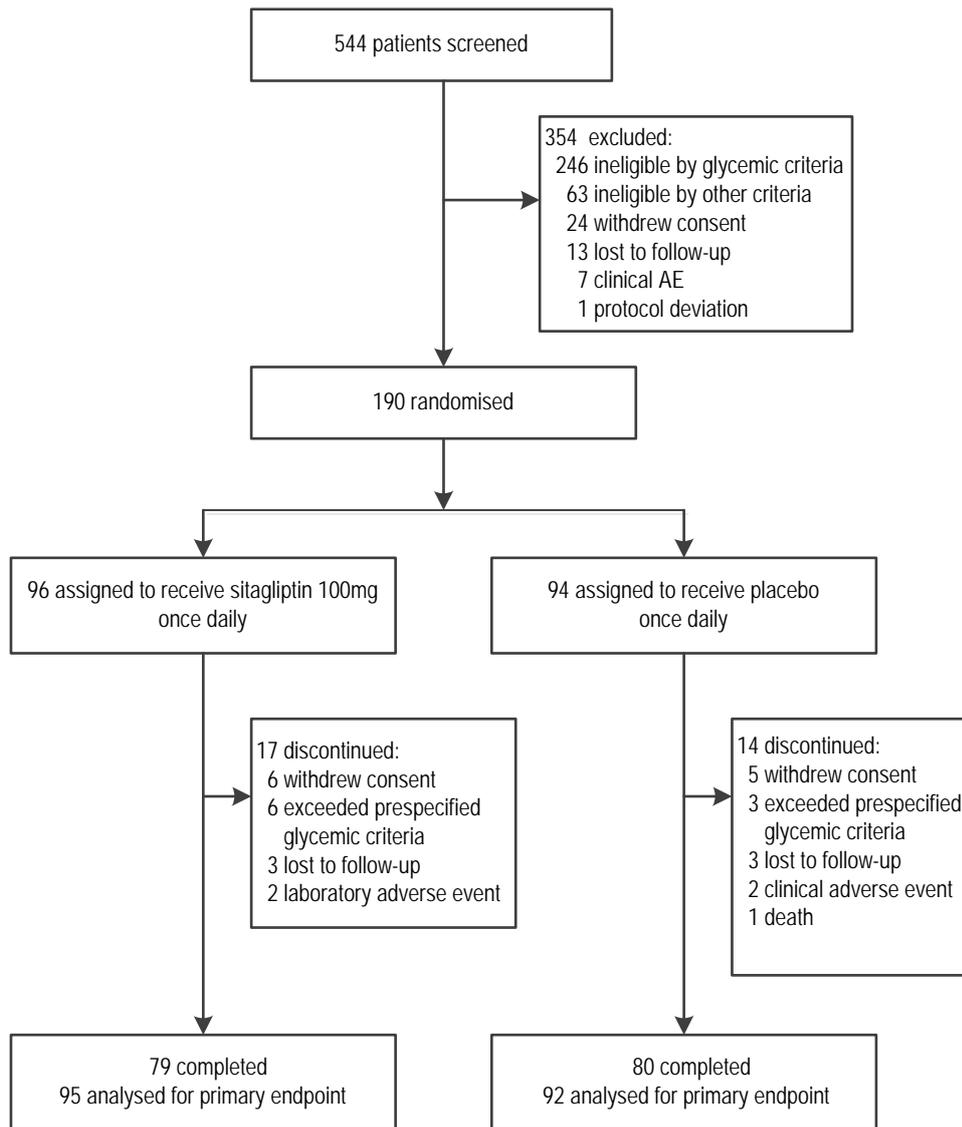


Abbildung 26: Flussdiagramm Studiendesign und -methodik für die Studie von Raz, 2008

Tabelle 4-168 (Anhang): Studiendesign und -methodik für die Studie von Bolli, 2008

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Vergleich der Wirksamkeit und Verträglichkeit von Vildagliptin versus Pioglitazon als Zusatztherapie bei Patienten mit Typ-2-Diabetes, die nicht mit Metformin allein ausreichend kontrolliert waren
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Die Studie war eine 24-wöchige randomisierte, doppelblinde, aktivkontrollierte, parallelgruppen-geführte (Verhältnis 1:1) Wirksamkeits- und Sicherheitsstudie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderungen wurden nach dem Studienbeginn vorgenommen
4	Probanden / Patienten	Mit Metformin allein nicht ausreichend kontrollierte Typ-2-Diabetiker
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Typ-2-Diabetes-Patienten mit HbA1c-Wert von 7,5–11,0%, die beim Studieneintritt eine stabile Dosis (≥ 1500 mg/Tag) Metformin erhielten. Männer und Frauen (nicht fruchtbar oder im gebärfähigen Alter mit einer medizinisch anerkannten Verhütungsmethode); im Alter von 18–77 Jahren, mit einem BMI von 22–45 kg/m², und mit Nüchternplasmaglukose (FPG) von <15 mmol/L konnten zur Teilnahme ausgewählt werden.</p> <p>Ausschlusskriterien: Patienten mit Typ-1- oder anderen Diabetesformen, akuten metabolischen diabetischen Komplikationen, Herzinfarkt, instabiler Angina oder koronarer Bypassoperation innerhalb der letzten 6 Monate. Kongestive Herzinsuffizienz (gemäß <i>New York Heart Association [NYHA] classes I–IV</i>) und Lebererkrankung wie Cirrhosis oder chronische aktive Hepatitis schlossen die Teilnahme ebenfalls aus. Patienten mit einem von den folgenden abnormalen Laborwerten wurden ebenfalls ausgeschlossen: Alanin-Aminotransferase oder Aspartat-Aminotransferase größer als das 2,5-Fache der Obergrenze der Normalwerte, direktes Bilirubin größer als das 1,3-Fache der Obergrenze des Normalwertes, Serumkreatininspiegel ≥ 132 $\mu\text{mol/L}$ (Männer) oder ≥ 125 $\mu\text{mol/L}$ (Frauen), klinisch relevantes abnormes schilddrüsen-stimulierendes Hormon oder Nüchterntriglyceride (TG) >7.9 mmol/L.</p>

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Internationale Studie in 118 Zentren; 26 in Deutschland, 25 in UK, 24 in USA, 16 in Spanien, 12 in Italien, 5 in der Schweiz, 4 in Österreich, 3 in Süd-Afrika und 3 in Australien
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Vildagliptin 100 mg täglich (in zwei identischen Dosen verabreicht) oder Pioglitazon 30 mg täglich (als eine Dosis verabreicht) zusätzlich zu der gewohnten Dosis des nicht-verblindeten Metformins.
6	Zielkriterien	Veränderung des HbA1c-Wertes in Woche 24 im Vergleich zu Studienbeginn
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Veränderung des HbA1c-Wertes in Woche 24 im Vergleich zu Studienbeginn in der <i>per protocol</i> (PP) Population. Die letzte Beobachtung (<i>last observation carried forward</i> , LOCF) wurde für Patienten verwendet, die die Studie vorzeitig abbrachen. Sekundäre Zielkriterien: FPG, Nüchternlipide und Körpergewicht.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderungen wurden nach dem Studienbeginn vorgenommen.
7	Fallzahl	Randomisierungszahlen: Vildagliptin 100 mg/Tag: n=295; Pioglitazon 30 mg/Tag: n=281
7a	Wie wurden die Fallzahlen bestimmt?	Nicht-Unterlegenheit von Vildagliptin gegenüber Pioglitazon bezüglich abnehmenden HbA1c-Wertes konnte belegt werden, wenn die Obergrenze des 95% Konfidenzintervalls für die Zwischen-Gruppendifferenz in Δ (korrigierte Mittelwerte der Änderungen vom Ausgangswert zum Endpunkt) für HbA1c 0,4% nicht überschritt. Wenn die Obergrenze 0,3% nicht überschritt, konnte die Nicht-Unterlegenheit von Vildagliptin gegenüber Pioglitazon auch mit einer 0,3%-Marge mittels sequentiellen Inferenzen belegt werden. 470 Patienten waren erforderlich in der PP-Population, um eine 95% statistische <i>Power</i> für die Nicht-Unterlegenheit von Vildagliptin gegenüber Pioglitazon mit einem einseitigen Signifikanzniveau von 0,025 zu belegen, wenn eine Nicht-Unterlegenheitsmarge von 0,4%, mit der Annahme des wahren Unterschieds 0 zwischen den Therapien und der Standardabweichung 1,2%, verwendet wurde. Entsprechend war die Studie mit der Populationsgröße von 510 Patienten ausreichend gepowert. Die Nicht-Unterlegenheitsmarge von 0,6 mmol/L war präspezifiziert, um den Einfluß von Vildagliptin und Pioglitazon auf die

Item ^a	Charakteristikum	Studieninformation
		Reduktion von FPG zu vergleichen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Computerbasierte Generierung
8a	Methode zur Generierung der zufälligen Zuteilung	Die Zuordnung in die Behandlungsgruppen erfolgte durch ein automatisches, zentrales Telefonsystem.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Jeder Patient bekam eine Randomisierungsnummer zugeteilt, die eingesetzt wurde, um die Patienten zu einem der zwei Behandlungsarmen zuzuweisen, und eine eindeutige Medikationsnummer für die erste Packung des Studienmedikaments spezifiziert, die den Patienten verteilt wurde. Es gab keine Stratifizierung.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Probanden erhielten eine eindeutige Patientenummer durch ein automatisches zentrales Telefonsystems.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Prüfpersonal
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Patienten und Prüfarzte waren voll verblindet. Sponsoren waren nicht verblindet.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht zutreffend
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die korrigierten Mittelwerte der Änderungen vom Ausgangswert zum Endpunkt (AMΔ) für primäre und sekundäre Endpunkte wurden mittels des Modells der Kovarianzanalyse analysiert, mit Therapie und gepoolte Studienzentren (<i>pooled centre</i>) als Klassifikationsvariablen und Eingangswert als Kovariate. Die Nicht-Unterlegenheits-Betrachtungen wurden mit einseitigen Tests und einem Signifikanzniveau von 0,025 durchgeführt. Die Überlegenheits-Betrachtungen wurden mit zweiseitigen Tests und einem Signifikanzniveau von 0,05 durchgeführt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Subgruppenanalysen für Patienten mit Basis-HbA1c-Wert $\leq 8.0\%$, für Patienten mit Basiswert $\geq 9.0\%$, und für adipöse und nicht-

Item ^a	Charakteristikum	Studieninformation
		übergewichtige Patienten. Die Änderung im Gewicht wurde in der adipösen (BMI >30) und in der schwer adipösen (BMI >35) Subgruppe analysiert.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 27)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 27)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 27)
14	Aufnahme / Rekrutierung	Die erste Patientensite fand am 29. September 2005 statt, und die letzte Patientensite am 22. Februar 2007.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Patienten wurden in den Behandlungswochen 4, 12, 16 und 24 beobachtet. Der 24-wöchigen Periode folgte eine 28-wöchige Periode.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

Abkürzungen: T2DM, Type-2-Diabetes; HbA1c, Glykiertes Hämoglobin; RCT, Randomisierte kontrollierte Studie; PP, Per protocol; FPG, Nüchternplasmaglukose; TG, Triglyceride; NYHA, New York Heart Association; BMI, *Body-Mass-Index*; AMΔ, Die korrigierten Mittelwerte der Änderungen vom Eingangswert zum Endpunkt (*Adjusted mean changes from baseline to endpoint*); LOCF, *last observation carried forward*; UK, United Kingdom; USA, United States of America

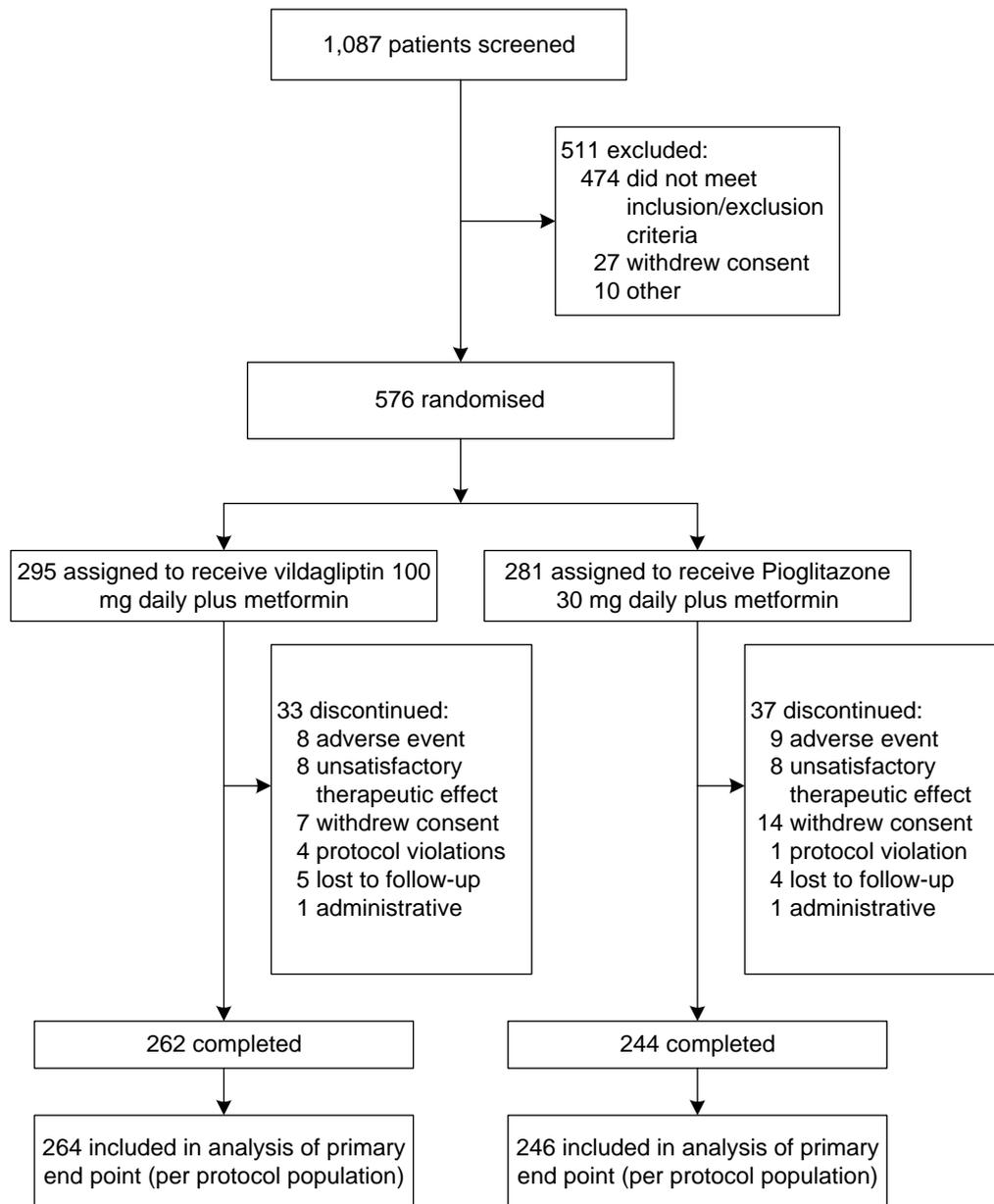


Abbildung 27: Flussdiagramm Studiendesign und -methodik für die Studie von Bolli 2008

Tabelle 4-169 (Anhang): Studiendesign und -methodik für die Studie von Bergenstal, 2010b

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Diese Studie untersucht die Sicherheit und Wirksamkeit des einmal wöchentlich verabreichten Exenatide, eines glukagon-ähnlichen Peptid-1-Rezeptorantagonisten im Vergleich zu der höchsten zugelassenen Dosis des Dipeptidyl-Peptidase-4-Inhibitors, Sitagliptin oder des Thiazolidindion, Prioglitazon, bei Patienten, die mit Metformin behandelt werden.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Bei der Studie handelt es sich um eine 26-wöchige, randomisierte, doppelblinde, double-dummy parallelgruppen-geführte (im Verhältnis 1:1) ^b Sicherheits- und Wirksamkeitsstudie.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderungen der Methodik nach dem Beginn der Studie
4	Probanden / Patienten	Mit Metformin allein nicht ausreichend kontrollierte Typ-2-Diabetiker
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>In die Studie aufgenommen wurden Patienten, die 18 Jahre alt oder älter waren, die Typ-2-Diabetes hatten, aber ansonsten gesund waren, und die Metformin in einer stabilen Dosis für mindestens 2 Monate vor der Eingangsuntersuchung erhielten.</p> <p>Zusätzliche Einschlusskriterien:</p> <p>HbA1c-Wert von 7,1%-11,0%; BMI-Wert von 25-45 kg/m²; stabiles Körpergewicht (d. h. keine Änderung um mehr als 3% in mindestens den letzten 3 Monaten vor der ersten Untersuchung); Nüchternplasmaglukose <15.5 mmol/L; entweder keine Behandlung oder eine stabile Therapie mit einer von den folgenden Medikationen für mindestens 2 Monate vor der Eingangsuntersuchung: Hormonersatztherapie (Frauen), orales Kontrazeptivum (Frauen), antihypertensive Mittel, lipidsenkende Mittel, Schilddrüsenhormonersatztherapie, Antidepressiva oder Mittel, die bekanntlich das Körpergewicht beeinflussen, einschließlich verschreibungspflichtige Medikamente (z. B. Orlistat [XENICAL[®]], Sibutramin [MERIDIA[®]], Topiramate [TOPAMAX[®]]) and <i>over-the-counter</i>-Medikamente gegen Fettsucht.</p> <p>Der Patient ist männlich oder weiblich und erfüllt alle folgenden Kriterien: nichtstillend,</p>

Item ^a	Charakteristikum	Studieninformation
		<p>negatives Schwangerschaftstestresultat (<i>human chorionic gonadotropin</i>-Test). Im Falle die Patientin hat keine Hysterektomie gehabt oder ist im gebärfähigen Alter (einschließlich perimenopausal Frauen, die die Menstruation innerhalb des letzten Jahres hatten), muss sie eine angemessene Empfängnisverhütung anwenden und bereit sein, diese während der gesamten Studiendauer und 90 Tage nach der letzten Injektion der Studienmedikation fortzusetzen. Die klinischen Laborwerte des Patienten (klinische Chemie, Hämatologie und Urinanalyse) sind nicht als klinisch relevant bewertet worden. Die Ergebnisse der physikalischen Untersuchung und des Elektrokardiogramms des Patienten sind als nicht klinisch relevant gestuft worden.</p> <p>Ausschlusskriterien:</p> <p>Schwangere Frauen wurden ausgeschlossen.</p> <p>Patienten mit einer klinisch signifikanten Erkrankung, die potenziell die Studienteilnahme oder das persönliche Wohlbefinden beeinflussen könnte, wurden ausgeschlossen. Dies beinhaltete folgendes: Lebererkrankung oder Serumwerte von Alanin-Aminotransferase oder Aspartat-Aminotransferase mehr als das Dreifache der Obergrenze der Normalwerte; Nierenerkrankung (Serumkreatinin $\geq 1,5$ mg/dL bei Männern und $>1,4$ mg/dL bei Frauen), kardiovaskuläre Erkrankung, einschließlich signifikantes Ödem, kongestive Herzinsuffizienz, oder kardiologischer Status gemäß <i>New York Heart Association Class III / IV</i>; Gastroparese; klinisch signifikante maligne Erkrankung (mit Ausnahme von Basal- und Plattenepithelkarzinom der Haut) in den 5 Jahren vor der Eingangsvisite; Makulaödem oder Hinweise auf bekannte oder vermutete Infektionen (z. B. <i>Human Immunodeficiency Virus</i> oder Tuberkulose).</p> <p>Patienten mit Drogen- oder Alkoholmissbrauch wurden ausgeschlossen, oder wenn sie eine Vorgeschichte vom Missbrauch hatten, die nach der Meinung des Prüfarztes die Studientreue dieser Person gefährden könnte. Patienten wurden ausgeschlossen, wenn sie einen Nüchterntriglycerid-Wert ≥ 600 g/dL bei der Eingangsvisite hatten; wenn sie zuvor Exenatide LAR ausgesetzt waren; wenn sie Blut in den letzten 60 Tagen vor der Eingangsvisite gespendet hatten oder geplant hatten, Blut während der Studie zu spenden. Patienten wurden ausgeschlossen, wenn sie eine größere Operation oder eine</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Bluttransfusion innerhalb von 2 Monaten vor der Eingangsuntersuchung hatten; oder wenn sie zur Zeit behandelt wurden oder es wurde erwartet, dass sie behandelt werden oder eine Behandlung benötigen, mit einem von den folgenden von der Behandlung ausgeschlossenen Medikamenten: Exenatide (BYETTA[®]) oder irgendein DPP-4-Inhibitor, SU, TZD, oder GLP-1-Analog innerhalb von 3 Monaten vor der ersten Visite; alpha-Glukosidase-Inhibitor, Meglitinid, Nateglinid, oder Pramlintid (SYMLIN[®]) innerhalb von 30 Tagen von der ersten Visite; Insulin innerhalb von 2 Wochen von der ersten Visite oder für länger als 1 Woche innerhalb von 3 Monaten von der ersten Visite; orale, intravenöse, oder intramuskuläre systemische Kortikosteroide; oder starke, inhalierte oder intrapulmonäre (einschließlich ADVAIR[®]) Steroide, die bekanntlich eine hohe Systemabsorptionsrate haben; Medikamente, die Wechselwirkungen mit dem CYP2C8-Enzymsystem haben, einschließlich Gemfibrozil (LOPID[®]) und Rifampin. Patienten wurden ausgeschlossen, wenn sie ein anderes Studienmedikament innerhalb von einem Monat von der ersten Visite erhielten; wenn sie bekannte Allergien oder Überempfindlichkeit gegenüber einer Komponente der Studienbehandlung (einschließlich PLG und Microsphere Diluent) hatten oder wenn sie zuvor eine klinisch signifikante Nebenwirkung (z. B. signifikantes Ödema) im Zusammenhang mit TZD oder DPP-4-Inhibitor hatten.</p>
4b	Studienorganisation und Ort der Studiendurchführung	72 Kliniken in USA, Indien und Mexiko.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Sitagliptin (100 mg) oral einmal täglich plus Placebo als Injektion einmal wöchentlich und Pioglitazon (45 mg) oral einmal täglich plus Placebo als Injektion einmal wöchentlich.
6	Zielkriterien	Änderung des HbA1c-Wertes zwischen Ausgangswert und Untersuchung nach 26 Wochen.
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäres Ziel: Änderung des HbA1c-Wertes zwischen Ausgangswert und Untersuchung nach 26 Wochen.</p> <p>Sekundäre Ziele: Prädefinierte Endpunkte aufgezeichnet von der Eingangsuntersuchung bis Woche 26: Anteil der Patienten, die das HbA1c-Ziel $\leq 6,5\%$ oder $\leq 7,0\%$ erreichten; Nüchternplasmaglukose ≤ 7 mmol/L; 6-Punkt Selbstmonitoring des Blutglukoseprofils;</p>

Item ^a	Charakteristikum	Studieninformation
		Körpergewicht; Nüchternlipidprofil; Nüchterninsulinprofil; systolischer und diastolischer Blutdruck; kardiovaskuläre Risikomarker (Urinalbumin/Kreatinin-Verhältnis, Serumadiponectin, B-Typ Natriuretisches Peptid, hochsensitives C-reaktives Protein und Plasminogen-Aktivator-Inhibitor-1); patientenberichtete Ergebnisse (<i>patient-reported outcomes</i>) aus den Fragebögen <i>the Impact of Weight on Quality of Life Questionnaire-Lite (IWQOL)</i> , <i>Psychological General Well-being (PGWB) index</i> , <i>the Diabetes Treatment Satisfaction Questionnaire (DTSQ)</i> , and <i>EuroQol-5 dimensions (EQ-5D)</i> , Sicherheit; und Verträglichkeit. Ein zusätzlicher sekundärer Endpunkt, der Anteil der Patienten, die HbA1c <7,0% erreichten, war nicht prädefiniert.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderungen wurden in den Zielen nach dem Studienbeginn vorgenommen.
7	Fallzahl	Randomisierungszahlen: Sitagliptin 100 mg/Tag: n=172; Pioglitazon 45 mg/Tag: n=172
7a	Wie wurden die Fallzahlen bestimmt?	Wir schätzten, dass 500 Patienten erforderlich wären, um mit einer statistischen <i>Power</i> von mindestens 90% einen statistisch signifikanten Unterschied ($\alpha=0,05$, zweiseitiger Test) zwischen Exenatide (einmal wöchentlich) und Sitagliptin oder Pioglitazon nach Hochbergs Multiple-Test-Verfahren und Annahme einer Differenz von 0,5% zwischen den Gruppen bei einer gängigen Standardabweichung von 1,2%, und einer Abbruchrate von 10%, nachzuweisen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Randomisierung erfolgte zentral mittels <i>interactive voice response</i> -System, um die Zuteilung zu verdecken, und war von Sponsor, Prüfarzten, Prüfzentrumspersonal und Patienten unabhängig.
8a	Methode zur Generierung der zufälligen Zuteilung	Computergenerierte Reihenfolge mit implementiertem <i>interactive voice response</i> -System
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung war stratifiziert durch das Land und den HbA1c-Wert beim Screening (d. h. <0,9% versus $\geq 9,0\%$)
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die oralen Kapseln und Injektionslösungen hatten identisches Aussehen, um die Medikamentenverteilung zu maskieren.

Item ^a	Charakteristikum	Studieninformation
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Alle Patienten, Prüfzentrumspersonal, Prüfarzte und der Sponsor waren bezüglich der Zuteilung zur Therapie während der doppelblinden Behandlungsperiode verblindet. Nach Fertigstellung des statistischen Analysenplans und anschließender Datenbanksperrung, wurde der Sponsor bezüglich der Behandlung entblindet.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht zutreffend
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Multiplizität für die Vergleiche von Exenatide versus Sitagliptin wurde mittels Hochbergs Verfahren ⁸ eingestellt, um die gesamte Typ-1-Fehlerrate bei 5% für HbA1c, Nüchternplasmaglukose, Körpergewicht, Nüchternlipidprofil, und Blutdruck zu prüfen. Wenn nicht anders spezifiziert, sind deskriptive statistische Methoden für demografische Indikatoren und Analysen für die <i>Intention-to-treat</i> -Population, alle randomisierten Patienten umfassend, die mindestens eine Dosis des Studienmedikaments erhielten. Die auswertbare Population bestand aus allen <i>Intention-to-treat</i> -Teilnehmern, die Studienhandlungen bis Woche 22 beendeten und eine adäquate Wirkung erhielten. Analysen der HbA1c-Änderungen bei jeder Visite basierten auf einem allgemeinen linearen Modell, das Behandlungs-, Land- und Basis-HbA1c-Strata (<9.0% vs ≥9.0%) einschloss. Zusätzlich wurde die Wechselwirkung zwischen den Behandlungs- und Basis-HbA1c-Strata bewertet. Analysen der Änderungen der anderen Parameter bei jeder Visite basierten auf einem allgemeinen linearen Modell, das Behandlungs-, Land-, und Basis HbA1c-Strata einschloss, mit dem entsprechenden Basiswert des Parameters als Kovariate. Log Transformation wurde bei den Triglyceriden und kardiovaskulären Markern vor dem Einsatz des Modells angewendet. Die Werte >10 mg/L des hochsensitiven C-reaktiven Proteins wurden von dem Vergleich ausgeschlossen, da ein erhöhter Wert des hochsensitiven C-reaktiven Proteins auf nicht-kardiovaskuläre Ursachen einer Inflammation hinweist. Vergleich der Behandlungsgruppen bezüglich der Anteile

Item ^a	Charakteristikum	Studieninformation
		der Patienten, die die Ziele von HbA1c und Nüchternplasmaglukose erreichten, wurde mit einem <i>Cochran-Mantel-Haenszel-Test</i> ausgeführt, stratifiziert mit Basis-HbA1c-Stratum und dem Land. Der Unterschied der Anteile der Patienten in den Behandlungsgruppen, die die Studie beendeten, wurde mit exaktem Test von Fisher bewertet. Für die fehlenden Daten wurde die letzte Beobachtung übertragen (<i>last observation carried forward</i>). Die Änderungen vom Ausgangswert zu Woche 26 für die Wirksamkeitsendpunkte wurden in Kleinste-Quadrate-Mittelwerten (<i>least square means</i>) ausgedrückt. Eine Post-hoc-Analyse für die Zwischen-Gruppenunterschiede für alle drei Behandlungen wurde auch ausgeführt. Statistische Analyse wurde mit SAS (Version 8.2) ausgeführt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht zutreffend
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 28)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 28)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 28)
14	Aufnahme / Rekrutierung	Patienten wurden vom 22. Januar, 2008 bis zum 6. August, 2008 rekrutiert.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Patienten wurden über 26 Wochen beobachtet. Nach Randomisierung erfolgten Untersuchungen in Wochen 0, 4, 6, 10, 14, 18, 22 und 26.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

^b Randomisierung erfolgte zu drei Armen, d. h. Exenatide 2 mg, Sitagliptin 100 mg und Pioglitazon 45 mg im Verhältnis 1:1:1, aber nur Sitagliptin und Pioglitazon wurden in die Nutzenbewertung einbezogen, da Exenatide zur Zeit der Ausführung nicht zugelassen war

Abkürzungen: T2DM, Type-2-Diabetes; HbA1c, Glykiertes Hämoglobin; RCT, Randomisierte kontrollierte Studie; BMI, *Body-Mass-Index*; LAR, *Long-lasting-release*; PLG, Poly lactide-co-Glycolid; SU, Sulfonylurea; TZD, Thiazolidindion; GLP-1, Glukagon-ähnliches Peptid-1; DPP-4, Dipeptidyl-Peptidase-4; IWQOL, *Impact of Weight on Quality of Life Questionnaire-Lite*; PGWB, *Psychological General Well-Being*; EQ-5D, *EuroQoL-5*; DTSQ, *Diabetes Treatment Satisfaction Questionnaire*; SD, Standardabweichung; USA, United States of America.

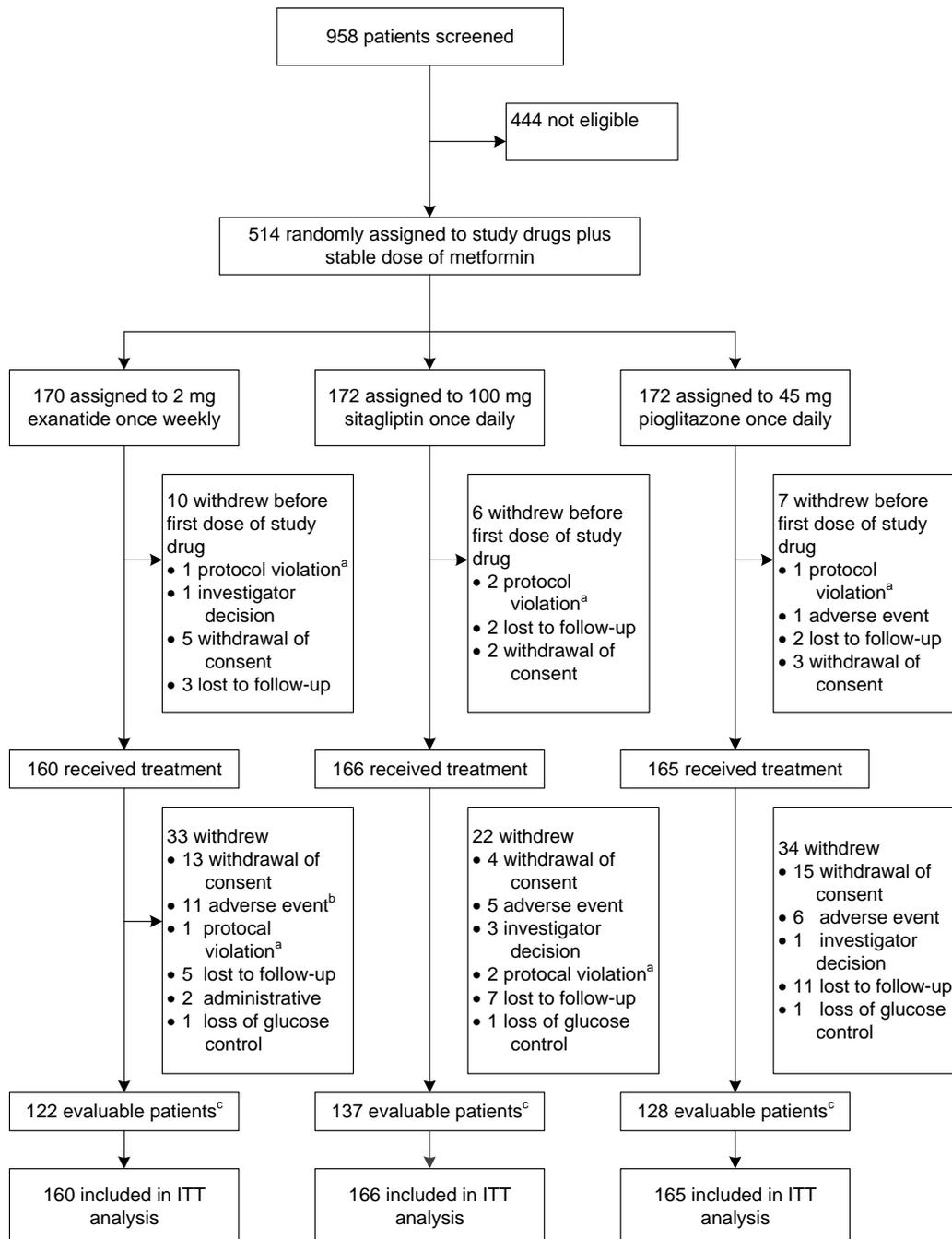


Abbildung 28: Flussdiagramm Studiendesign und -methodik für die Studie von Bergenstal 2010b.

Anmerkung: ^a, Alle Protokollverletzungen erfolgten durch Nichteinhaltung der Untersuchungstermine oder der Behandlung, oder beides; ^b, Ein unerwünschtes Ereignis war nicht durch die Behandlung bedingt; ^c, Die Patienten beendeten die Studie bis Woche 22 in Übereinstimmung mit dem Protokoll und erhielten eine angemessene Dosis der Studienmedikation während der Behandlungsperiode. Randomisierung erfolgte zu drei Armen, d. h. Exenatide 2 mg, Sitagliptin 100 mg und Pioglitazon 45 mg im Verhältnis 1:1:1, aber nur Sitagliptin und Pioglitazon wurden in die Nutzenbewertung einbezogen, da Exenatide zur Zeit der Ausführung nicht zugelassen.

Abkürzung: ITT, Intention to treat analysis.

Tabelle 4-170 (Anhang): Studiendesign und -methodik für die Studie von DeFronzo, 2005

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Diese Studie evaluiert die Fähigkeit des GLP-1-Agonisten Exenatide (Exendin-4) die Blutzuckerkontrolle von Patienten mit Typ-2-Diabetes, welche mit maximal wirksamer Metformin-Dosis eine mangelhafte Blutzuckerkontrolle erreichten, zu verbessern.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Eine dreifach-blinde, placebokontrollierte Wirksamkeits-/Sicherheitsstudie mit paralleler Zuordnung (Verhältnis 1:1:1) ¹
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nach Beginn der Studie wurden keine Änderungen in der Methodik durchgeführt.
4	Probanden / Patienten	Typ-2-Diabetiker die mit alleiniger Metformin Behandlung unzureichend kontrolliert waren
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Patienten im Alter von 19 bis 78 Jahren; Screening Nüchtern-Plasma-Blutzuckerspiegel von <13,3 mmol/L (<240 mg/dL), BMI von 27-45 kg/m², und ein HbA1c-Wert von 7,1-11,0%; Metformin-Dosis ≥1500 mg/Tag für 3 Monate vor Zeitpunkt des Screenings mit keinem (für eine Bevölkerung mit Typ-2-Diabetes) klinisch signifikanten abnormalen Laborwert (>25% außerhalb des Normbereichs); postmenopausale Frauen, sterilisierte Personen, oder Patientinnen, die Kontrazeptiva für 3 Monate vor Zeitpunkt des Screenings und fortlaufende während der Studie einnahmen.</p> <p>Ausschlusskriterien:</p> <p>Einnahme von Sulfonylharnstoffen, Meglitinide, Thiazolidindionen, Glukosidaseinhibitoren, exogene Insulintherapie, Medikamente zur Gewichtsabnahme, Kortikosteroide, Medikamente, die bekannt dafür sind die gastrointestinale Motilität zu beeinflussen, Transplantationsmedikationen, oder irgendwelche Prüfpräparate, oder der Nachweis von klinisch signifikanten Komorbiditäten mehr als 3 Monate vor Zeitpunkt des Screenings.</p>

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	82 Standorte in den USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	zweimal täglich 5 mg Exenatide, zweimal täglich 10 mg Exenatide, Placebo. Das Studiendesign sieht eine Eingewöhnungszeit (4 Wochen) mit einer geringeren Exenatide-Dosis (5 mg zweimal täglich) in den Behandlungsarmen A und B mit vor, bevor die festgelegte Exenatide-Dosis entweder auf 10 mg zweimal täglich (Arm B) erhöht wurde oder bei 5 mg zweimal täglich für die gesamte Studiendauer beibehalten wurde.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäre Zielmessungen waren: Prozentuale Änderung zum HbA1c-Ausgangswert in Woche 30. Sekundäre Zielmessungen waren: Prozentualer Anteil der Patienten, die in Woche 30 einen HbA1c-Wert $\leq 7\%$ erreichten; Nüchtern- und postprandiale <i>meal-cohort</i> , Plasma-Blutzuckerspiegel, Körpergewicht, Nüchtern- und postprandiale Insulinblutspiegel, Nüchtern-Proinsulin und Lipide
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurden keine Änderungen der Zielkriterien nach Studienbeginn vorgenommen.
7	Fallzahl	Zahl der randomisierten Patienten: Exenatide 5 mg einmal täglich: n=110, Exenatide 10 mg einmal täglich: n=113, Placebo: n=113
7a	Wie wurden die Fallzahlen bestimmt?	Zur Bereitstellung einer statistischen <i>Power</i> von 90% wurde eine minimale Stichprobengröße von 300 Personen mit mindestens einem HbA1c-Messwert nach dem Ausgangswert geschätzt, um mindestens zwischen einem Exenatide-Behandlungs-Arm und Placebo eine 0,6% Differenz als Änderung vom HbA1c-Ausgangswert erfassen zu können ($\alpha=0,05$ Fisher's <i>protected testing</i> Prozedur).
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung wurde anhand der HbA1c-Werte, die zum Zeitpunkt des

Item ^a	Charakteristikum	Studieninformation
		Screenings gemessen wurden, stratifiziert (<9,0% und ≥9,0%), um eine balancierte Verteilung der Personen in den Behandlungsarmen zu erreichen.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar. Die Autoren bezeichnen die Studie als „dreifach-blind“, berichten aber nicht explizit darüber, wer verblindet wurde oder wie diese Verblindung vorgenommen wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Zur Testung der Unterschiede in der Veränderung der Ausgangs-HbA1c-Werte bis zu jeder Untersuchung und des Gewichts über alle Behandlungsgruppen wurde ein allgemeines lineares Modell benutzt. Ins Modell wurden die Behandlung (Placebo und zwei aktive Behandlungsarme), Strata des Ausgangs-HbA1c-Werts (<9,0% und ≥9,0%), und die Studienstandorte als <i>fixed effects</i> miteingeschlossen. Die ITT-Population galten alle randomisierten Probanden, die mindestens eine Injektion der Medikation erhielten, begonnen am Abend von Tag 1. Alle Wirksamkeits- und Sicherheitsanalysen wurden mit der ITT-Population durchgeführt mit Ausnahme des prozentualen Anteils jener Probanden, die einen HbA1c-Wert von ≤7% (evaluierbare Population) erreichten und der <i>meal tolerance cohort</i> . Fehlende Werte der ITT-Probanden, (einschließlich fehlender Werte an Zwischenuntersuchungstagen) wurden durch Daten von vorgesehenen Untersuchungstagen unter Anwendung der LOCF Methode aufgefüllt. Die

Item ^a	Charakteristikum	Studieninformation
		<p>Quadratmittelwerte und Standardfehler wurden aus dem allgemeinen linearen Modell für jede Behandlung ermittelt. Paarweise Vergleiche der Behandlungseffekte wurden mittels der <i>Fisher's protected testing</i> Prozedur ermittelt, um die Typ-1-Fehler aufgrund multipler Vergleiche zu kontrollieren. Ähnliche Analysen wurden für jeden metabolischen Nüchtern-Parameter und für postprandiale Blutzuckerkonzentrationen ohne Adjustierung für die multiplen Vergleiche durchgeführt. Mittels Regressionsanalyse wurde die Post-hoc-Evaluation der Änderung des Körpergewichts vs. der Dauer der Übelkeit ermittelt. Zwischen den Behandlungsgruppen wurde der Anteil der Probanden mit einem HbA1c $\leq 7\%$ mit dem <i>Cochran-Mantel-Haenszel-Test</i> verglichen, wobei die Strata der Ausgangs-HbA1c-Werte als Stratifizierungsfaktor dienten.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Die evaluierbare Population bestand aus alle randomisierten Probanden, die die Behandlung über 30 Wochen mit mindestens 80% der Injektionen der Studienmedikation vollständig erfüllt haben. Probanden, die 7 aufeinanderfolgende Injektionstage während der letzten 2 Monate der Studie versäumt haben, wurden ausgeschlossen.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 29)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 29)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 29)
14	Aufnahme / Rekrutierung	Die Studie wurde zwischen Januar 2002 und Juni 2003 durchgeführt.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Nicht klar – es scheint, als ob die Follow-up-Besuche in den Wochen 2, 4, 6, 12, 16, 24 und 30 stattfinden.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht anwendbar – endete wie nach Angaben im Protokoll nach 30 Wochen
a: nach CONSORT 2010.		

^a Randomisiert wurde zwischen 2 Dosen von Exendin-4 (5 mg und 10 mg) und einem Placebo in einem 1:1:1 Verhältnis

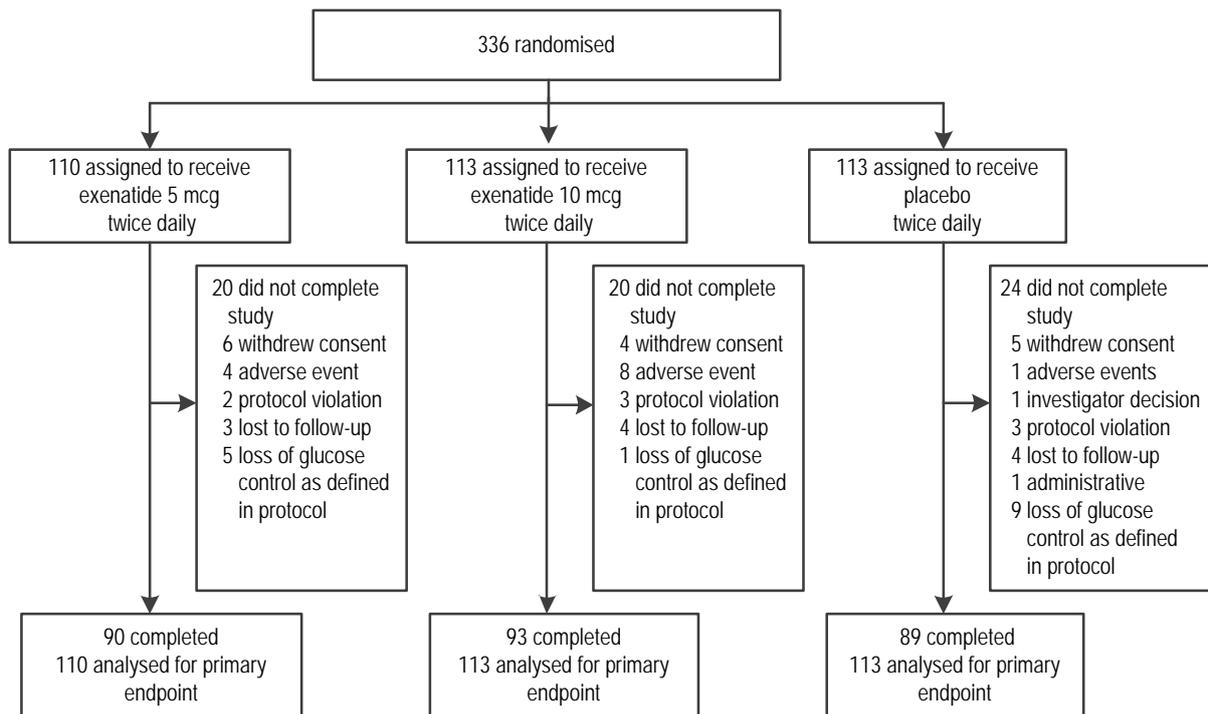


Abbildung 29: Flussdiagramm Studiendesign und -methodik für die Studie von DeFronzo, 2005

Tabelle 4-171 (Anhang): Studiendesign und -methodik für die Studie von Nauck, 2009

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Die Wirksamkeit und Sicherheit von Liraglutide (einem Glucagon-like Peptid-1 Rezeptor Agonist) in der Kombination mit Metformin wurde mit Placebo oder Glimperid in der Kombination mit Metformin verglichen bei Probanden die zuvor OAD behandelt wurden.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, double-dummy, placebo- und aktivkontrollierte Wirksamkeits- und Sicherheitsstudie ¹ mit Paralleldesign (Verhältnis 2:2:2:1:2)
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nach Beginn der Studie wurden keine Änderungen in der Methodik durchgeführt.
4	Probanden / Patienten	Typ-2-Diabetiker, die mit alleiniger Metformin-Behandlung unzureichend kontrolliert waren.
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: Patienten im Alter von 18 bis 80 Jahren; HbA1c-Wert zwischen 7 und 11% (Vorstudie OAD Monotherapie für ≥3 Monate) oder zwischen 7 und 10% (Vorstudie OAD Kombinationstherapie für ≥3 Monate), und BMI ≤40 kg/m ² Ausschlusskriterien: Behandlung mit Insulin während der letzten 3 Monate (ausgenommen kurzzeitige Behandlung)
4b	Studienorganisation und Ort der Studiendurchführung	170 Studienstandorte in 21 Ländern
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Liraglutide 0,6 mg ² Liraglutide 1,2 mg, Liraglutide 1,8 mg, Placebo, Glimperid 4 mg ³ . Alle Interventionen wurden einmal täglich zusammen mit der Injektion von Metformin verabreicht. Das Double-dummy Design erfordert, dass Probanden der Liraglutide- und Placebo-Gruppe ein Glimperid-Placebo erhalten, wobei Probanden in der Glimperid- und Placebo-Gruppe einen Injektion von Liraglutide-Placebo erhalten. Nach einer 3-wöchigen Metformin-Titrationsperiode fand die Randomisierung

Item ^a	Charakteristikum	Studieninformation
		<p>der Behandlung statt (Anstieg der Dosis bis zu 2000 mg/Tag: 1000 mg morgens und 1000 mg abends), danach folgte eine 3-wöchige Metformin-Erhaltungsperiode. Probanden, die vor Studienbeginn Metformin einnahmen, wurden erhielten entweder mittels Titration eine Dosisoptimierung oder wurden direkt in die Metformin-Erhaltungsphase zugeordnet.</p> <p>Nach der Randomisierung unterzogen sich die Probanden einer 2- bis 3-wöchigen Titrationsperiode für Liraglutide (bis zu 0,6, 1,2, oder 1,8 mg, wie randomisiert, 0,6 mg-Anstiege pro Woche) und Glimepirid (bis zu 4 mg, mit 1-, 2-, und 4 mg-Dosen in Wochen 1, 2, und 3). Glimepirid (aktiv und Placebo) wurde oral einmal täglich morgens zugeführt. Liraglutide (aktiv oder Placebo) wurde subkutan einmal täglich zu einer beliebigen Tageszeit im Oberarm, Bauch oder Oberschenkel gespritzt mittels eines Pen-Injektionsgeräts. Probanden wurden angehalten Liraglutide täglich zur gleichen Zeit zu spritzen.</p> <p>Nach der Titrationsperiode erfolgte eine 23- bis 24-wöchige Erhaltungsperiode, währenddessen die Dosen der Studienmedikationen erhalten blieben. Dennoch konnte die Metformin-Dosis bis zu einem Minimum von 1500 mg/Tag im Falle einer unzumutbaren Hypoglykämie oder anderen unerwünschten Ereignissen gesenkt werden, musste aber zwischen 1500 und 2000 mg/Tag während der Erhaltungsperiode konstant gehalten werden.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäre Zielmessungen waren:</p> <p>Prozentuelle Veränderungen zum HbA1c-Ausgangswert in Woche 26.</p> <p>Sekundäre Zielmessungen waren:</p> <p>Änderung des Körpergewichts, des Nüchtern Plasma-Blutzuckerspiegel, des 7-Punkte-Blutzuckerprofil (vor jeder Mahlzeit, 90 Minuten nach dem Frühstück, Mittagessen und Abendessen, und vor dem Schlafengehen), und der β-Zellfunktion basierend auf dem Nüchtern-Insulinwert, Nüchtern-C-Peptid, Nüchtern-Proinsulin-Insulin Verhältnis, und den Homöostase-Model-Bewertungsindex der β-Zellfunktion (Matthews 1985).</p>

Item ^a	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurden keine Änderungen der Zielkriterien nach Studienbeginn vorgenommen.
7	Fallzahl	Anzahl der randomisierten Probanden: Liraglutide 0,6 mg einmal täglich: n=242; Liraglutide 1,2 mg einmal täglich: n=241; Liraglutide 1,8 mg einmal täglich: n=242; Glimepirid einmal täglich: n=244; Placebo einmal täglich: n=122
7a	Wie wurden die Fallzahlen bestimmt?	Den Stichprobenberechnungen liegen Unterschiede im HbA1c-Wert und Körpergewicht von 0,5 und 3%, bzw., Unterschieden nach 6 Monaten Behandlung zu Grunde. Die zu erwartende Standardabweichung für den HbA1c-Wert und den Varianzkoeffizient für das Körpergewicht lagen bei 1,2 bzw. 3%. Die kombinierte statistische <i>Power</i> (berechnet als Produkt der marginalen statistischen <i>Power</i> für HbA1c und Körpergewicht) betrug mindestens 85%.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung erfolgte mit Hilfe eines Telefon-basierten oder Web-basierten Randomisierungssystem. Probanden wurden zufällig mit der kleinsten verfügbaren Randomisierungsnummer markiert.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Probanden wurden unter Berücksichtigung ihrer bisherigen Anwendung der OAD Mono- oder Kombinationstherapie stratifiziert.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar. Die Autoren bezeichneten diese Studie als „doppelblind, <i>double-dummy</i> “, berichten aber nicht explizit darüber, wer verblindet wurde oder wie diese Verblindung vorgenommen wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von	Die Autoren bezeichneten diese Studie als

Item ^a	Charakteristikum	Studieninformation
	Interventionen	„doppelblind, <i>double-dummy</i> “.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Analyse der Wirksamkeit hinsichtlich der Endpunkte wurde mit der ITT-Population durchgeführt, welche aus Probanden besteht, die mindestens eine Dosis der Studienmedikation erhielten und eine Messung des Parameters nach der Ausgangsmessung erfolgte. Jeder Endpunkt wurde mittels ANCOVA-Modell analysiert, wobei Behandlung, Land und vorherige antidiabetische Behandlung als <i>fixed effects</i> und Ausgangswerte als Kovariable definiert wurden. Fehlende Werte wurden mit der LOCF Methode ergänzt. Überlegenheit oder Nichtunterlegenheit der Blutzuckerkontrolle mit Liraglutide vs. Komparatoren wurden gefolgert, falls die obere Grenze des zweiseitigen 95%-Konfidenzintervalls für die Behandlungsunterschiede <0% war; Nichtunterlegenheit wurde gefolgert, falls sie <0,4% war. Mittels logistischer Regression wurde der Anteil der Probanden, der die HbA1c-Zielwerte erreichte (Zielwert der <i>American Diabetic Association</i> <7%; Zielwert der <i>American Association of Clinical Endocrinologists</i> ≤6,5%), verglichen, wobei Behandlung und Ausgangs-HbA1c-Wert Kovariablen darstellten.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Die Konfidenzintervalle für die sekundären Endpunkte wurden mit dem <i>Dunnnett's</i> Test korrigiert. Hypoglykämische Episoden wurden mit einem allgemeinen linearen Modell analysiert, einschließlich der Behandlung als ein <i>fixed effect</i> . Die Werte wurden als Mittelwerte ± Standardabweichung dargesellt, sofern nicht anders angegeben. Das Signifikanzniveau wurde auf $p < 0,05$ festgesetzt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 30)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 30)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 30)
14	Aufnahme / Rekrutierung	

Item ^a	Charakteristikum	Studieninformation
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Der Zeitraum der Probandenrekrutierung wurde nicht berichtet. Die Studie dauerte 26 Wochen. Die Messungen erfolgten alle 4 Wochen.
14b	Informationen, warum die Studie endete oder beendet wurde	Keine Angabe – endet wie nach Angaben im Protokoll nach 26 Wochen.
a: nach CONSORT 2010.		

¹ Randomisiert wurde mit einer von 3 täglichen Dosen mit Liraglutide (0,6, 1,2 und 1,8 mg/Tag) gespritzt mit Metformin, Placebo mit Metformin, oder einer Kombinationstherapie mit Glimperid (4 mg einmal täglich) mit Metformin; ² Dieser Studienarm wurde nicht in die Analyse miteingeschlossen; ³ Dieser Studienarm wurde nur in die Analyse des systolischen Blutdrucks eingeschlossen;

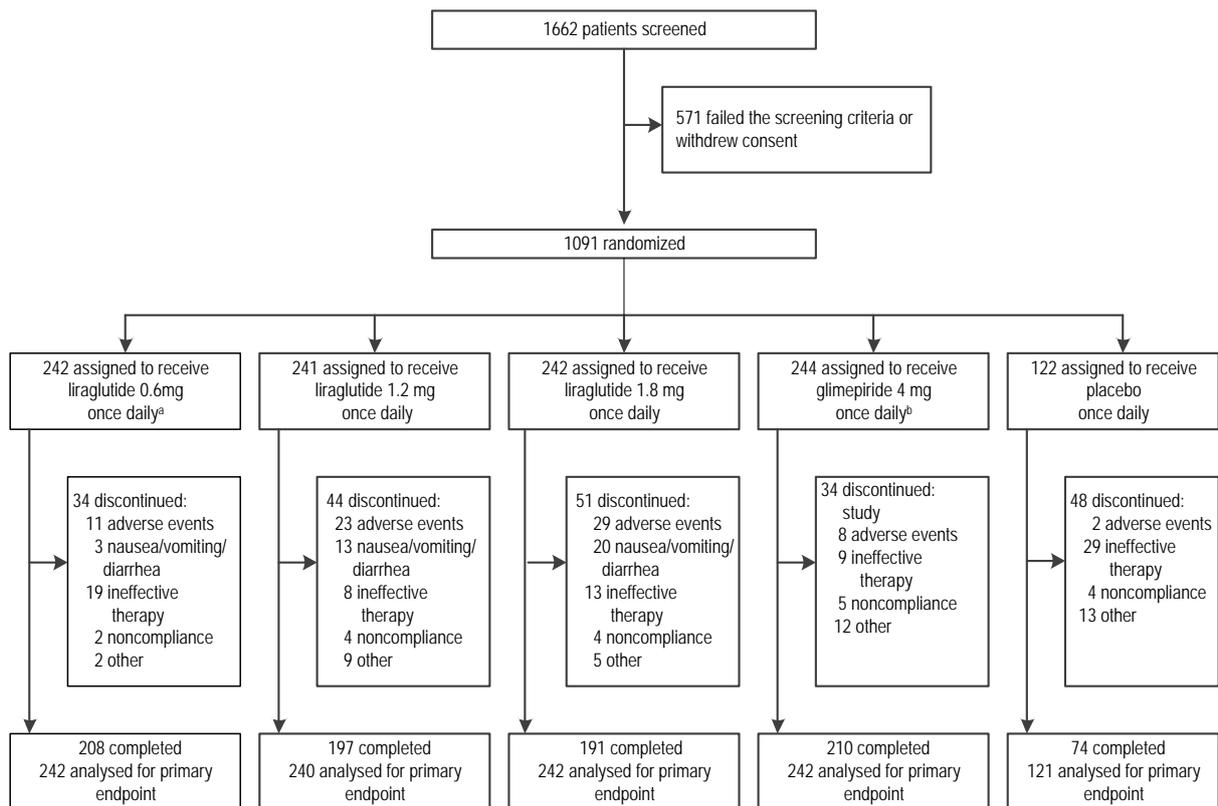


Abbildung 30: Flussdiagramm Studiendesign und -methodik für die Studie von Nauck, 2009

Notes: ^a Dieser Arm wurde nicht in die Analyse eingeschlossen; ^b Dieser Arm war nicht in die Analyse des systolischen Blutdruckes eingeschlossen

Tabelle 4-172 (Anhang): Studiendesign und -methodik für die Studie von Pratley, 2010

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Das Ziel dieser Studie war die Wirksamkeit und Sicherheit von der Liraglutide- und Sitagliptin-Therapie für 26 Wochen bei Typ-2-Diabetes-Patienten zu vergleichen, die keine ausreichende glykämische Kontrolle mit Metformin erreichten.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Aktivkontrollierte Parallelstudie mit einem Zuteilungsverhältnis von 1:1:1.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nach Studienbeginn wurden keine Änderungen der Methodik vorgenommen
4	Probanden / Patienten	Mit Metformin allein nicht ausreichend kontrollierte Typ-2-Diabetiker
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <p>Patienten, die in die Studie aufgenommen wurden, waren im Alter von 18-80 Jahren, hatten Typ-2-Diabetes, hatten glykiertes Hämoglobin (HbA1c)-Wert von 7,5-10,0% und BMI von $\leq 45,0$ kg/m², und wurden mit Metformin (≥ 1500 mg täglich) seit ≥ 3 Monaten behandelt.</p> <p>Ausschlusskriterien:</p> <p>Behandlung mit einem anderen antihyperglykämischen Medikament als Metformin in den letzten 3 Monaten vor der Studie; wiederholte schwere Hypoglykämie oder Hypoglykämie-Wahrnehmungsstörung; gegenwärtiger Gebrauch eines Medikaments außer Metformin, das Glukose beeinflussen könnte; Kontraindikation zu Studienmedikamenten; verschlechterte Nieren- oder Leberfunktion, klinisch signifikante kardiovaskuläre Erkrankung; oder Tumor.</p>
4b	Studienorganisation und Ort der Studiendurchführung	Studie wurde in 158 niedergelassenen Praxen (<i>office-based-sites</i>) in 11 europäischen Ländern (Kroatien, Deutschland, Irland, Italien, Niederlande, Rumänien, Serbien, Slowakien, Slowenien, Spanien und England), in USA und Kanada durchgeführt.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	1,2 mg oder 1,8 mg subkutanes Liraglutide einmal täglich, oder 100 mg orales Sitagliptin einmal täglich; alle zusätzlich zu Metformin (Metformin-Behandlung blieb stabil). Liraglutide wurde mit 0,6 mg/Tag begonnen und mit 0,6 mg/Woche auf die zugeordnete Dosis gesteigert; Injektion wurde subkutan mit einem Stiftgerät (<i>pen device</i>) durchgeführt. Sitagliptin wurde begonnen

Item ^a	Charakteristikum	Studieninformation
		und blieb bei einer Dosis von 100 mg/Tag. Teilnehmer wurden aufgefordert, die Medikation jeden Tag zur gleichen Zeit einzunehmen.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäres Zielkriterium: Änderung in den HbA1c-Werten zwischen Eingangsuntersuchung und Untersuchung nach 26 Wochen.</p> <p>Sekundäre Zielkriterien erfasst am Studienbeginn und in Woche 26 waren: Anteil der Teilnehmer, die HbA1c-Ziele <7% oder ≤6,5% erreichten; Nüchternplasmaglukose; postprandiale Plasmaglukose; Körpergewicht; β-Zellfunktion; Nüchternlipidprofil; kardiovaskuläre Risikomarker (hochsensitives C-reaktives Protein, Plasminogen-Aktivator-Inhibitor-1, N-terminales Pro-B-Typ natriuretisches Peptid, Adiponectin, Interleukin 6, <i>tumor necrosis factor-α</i>, und von Willebrand-Faktor); Blutdruck; Herzfrequenz; physikalische Messungen (Taillenumfang, Taille-Hüfte-Verhältnis); Behandlungszufriedenheit; und ein zusammengesetzter Endpunkt bestehend von Anteil der Patienten mit HbA1c <7.0%, mit keiner Hypoglykämie und mit Gewichtsänderung von 0 kg oder weniger.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	Randomisierungszahlen: 1,2 mg Liraglutide (subkutan einmal täglich): n=225; 1,8 mg Liraglutide (subkutan einmal täglich): n=221; 100 mg Sitagliptin (oral einmal täglich): n=219
7a	Wie wurden die Fallzahlen bestimmt?	Aufgrund von primärem Endpunkt und Nicht-Unterlegenheitsmarge, rechneten die RCT-Prüfärzte, dass für 85% statistische <i>Power</i> 163 Teilnehmer pro Gruppe, und mit einer vorhergesagten Abbruchrate von 25%, 217 Teilnehmer pro Gruppe erforderlich sind.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Computerbasierte Generierung
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierungsreihenfolge wurde von Novo Nordisk mit Computer generiert.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Teilnehmer waren randomisiert im Verhältnis 1:1:1, stratifiziert durch das Land, zu 1,2 mg oder 1,8 mg subkutanem Liraglutide einmal täglich, oder zu 100 mg oralem Sitagliptin einmal täglich.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment)	Fortlaufende Zuteilung des Randomisierungscodes an die einzelnen Teilnehmer wurde unter Verwendung von

Item ^a	Charakteristikum	Studieninformation
	Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	telefonbasiertem (<i>interactive voice response system</i>) oder webbasiertem (<i>interactive web response system</i>) Randomisierungssystem verdeckt.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Prüfarzte verteilten die Patienten randomisiert zu der niedrigsten verfügbaren Nummer aus dem Zahlenbereich, der dem Zentrum zugeteilt worden war. Die Patientennummern hatten sechs Stellen (XXXYYY), in denen XXX das Zentrum und YYY die Patientenummer in diesem Zentrum bezeichneten.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Dies war eine offene Studie. Niemand war verblindet.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht zutreffend
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Der primäre Endpunkt wurde mittels ANCOVA-Modell mit Therapie und Land als feste Effekten (<i>fix effect</i>) und HbA1c-Ausgangswert als Kovariate analysiert. Um den Gesamt Typ-1-Fehler von 5% zu schützen, wurde eine hierarchische Annäherung verwendet, wie im Protokoll und in dem statistischen Analysenplan prädefiniert: die Nicht-Unterlegenheit wurde zuerst ermittelt und, wenn bestätigt, wurde danach Überlegenheit getestet, beide mit einem Signifikanzniveau von 2.5%. Nicht-Überlegenheit wurde festgestellt, wenn die obere Grenze des zweiseitigen 95% Konfidenzintervalls für den Therapieunterschied $<0,4\%$ war; Überlegenheit wurde festgestellt, wenn die obere Grenze <0 war. Die meisten sekundären Endpunkte wurden mittels ANCOVA beurteilt, mit Therapie und Land als feste Effekte (<i>fixed effects</i>) und Ausgangswertmessung als Kovariate. Logistische Regression wurde verwendet, um die Anteile der Teilnehmer, die HbA1c-Ziele und den zusammengesetzten Endpunkt erreichten, zu vergleichen, und um die Quotenverhältnisse (<i>odds ratios</i>) zu generieren. Therapie war ein fixer Effekt und HbA1c-Ausgangswert eine Kovariate, und Ausgangsgewicht eine zusätzliche Kovariate nur für den zusammengesetzten Endpunkt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Die Behandlungszufriedenheit wurde mit dem Acht-Punkt <i>Diabetes Treatment Satisfaction Questionnaire</i> -Fragebogen

Item ^a	Charakteristikum	Studieninformation
		(Bradley, 1994) in einer Subgruppe von 505 Teilnehmern bewertet; Gesamtbehandlungszufriedenheit war die Summe von sechs von den acht <i>Diabetes Treatment Satisfaction Questionnaire</i> Punkten. Die Änderung in HbA1c wurde für die Subgruppe mit dem HbA1c-Ausgangswert $\geq 9,0\%$ analysiert.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 31)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 31)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 31) Die primären Wirksamkeitsanalysen wurden mit <i>full analysis set</i> (randomisierte Patienten, die mindestens eine Dosis des Studienmedikaments erhielten und mindestens eine HbA1c-Messung nach dem Ausgangswert hatten), die fehlenden Daten der letzten Beobachtung übertragen (<i>last observation carried forward</i>), und mit <i>per-protocol set</i> , durchgeführt; die sekundären Wirksamkeitsanalysen wurden mit <i>full analysis set</i> , mit Ausnahme der Behandlungszufriedenheitsanalysen, in denen die fehlenden Daten nicht berücksichtigt wurden, da die Daten nur für Screening und Randomisierungsvisiten verfügbar waren, durchgeführt.
14	Aufnahme / Rekrutierung	Diese Studie wurde vom 16. Juni, 2008 bis zum 11. Juni, 2009 durchgeführt.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Nach der 26-wöchigen Studie konnten die Teilnehmer in eine noch laufende 12 Monate dauernde Follow-up-Studie wechseln. Sicherheits- und Wirksamkeitsparameter wurden am Beginn und nach 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 und 26 Wochen gemessen.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

Abkürzungen: ANCOVA, Analysis of covariance; HbA1c, Glykiertes Hämoglobin

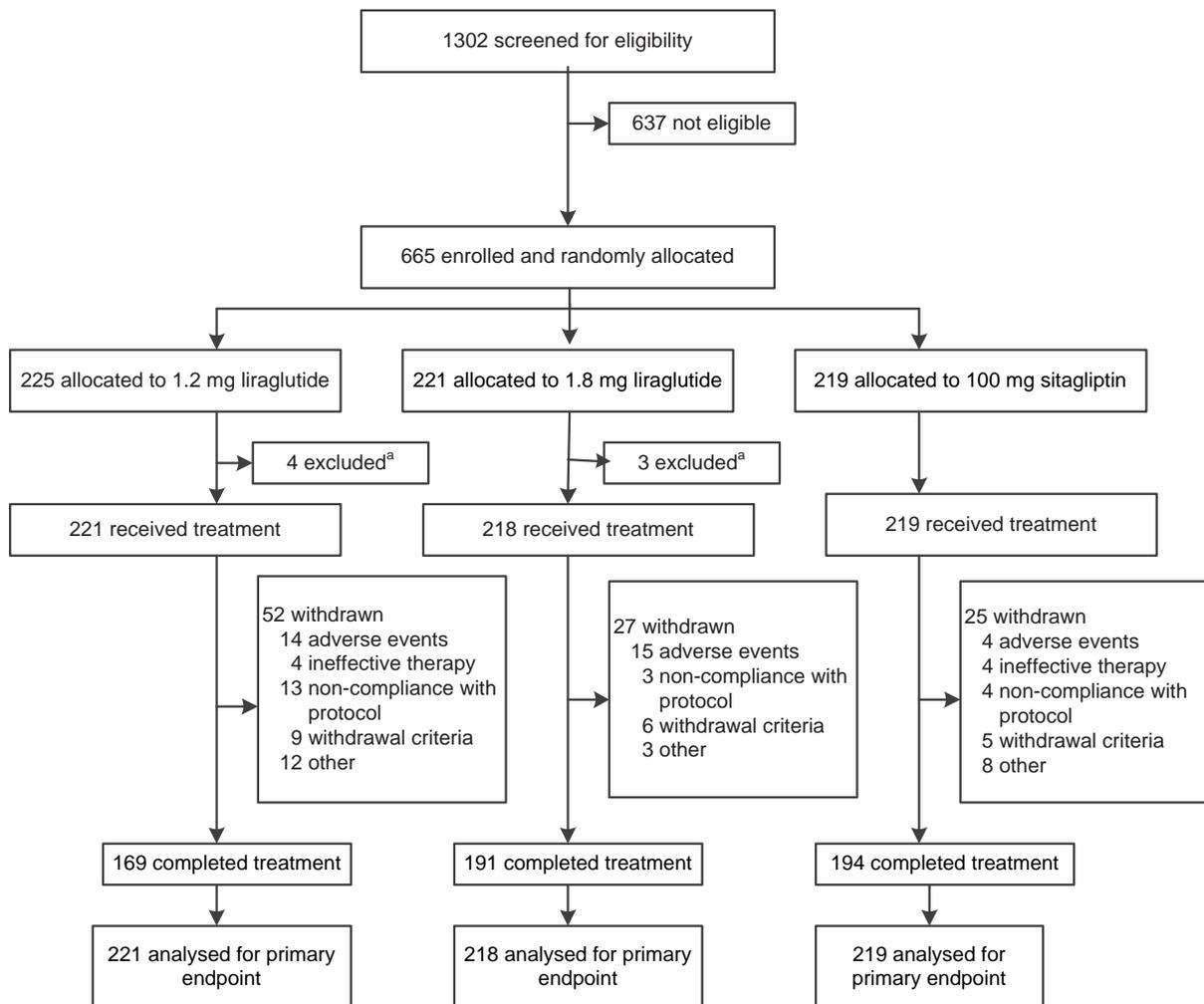


Abbildung 31: Flussdiagramm Studiendesign und -methodik für die Studie von Pratley, 2010

^aDie Patienten wurden ausgeschlossen, wenn sie die Ausschlusskriterien erfüllten, wenn sie ihre Einverständniserklärung zurückzogen oder wenn sie nach der Randomisierung nicht mehr zur Untersuchung erschienen.

Tabelle 4-173 (Anhang): Studiendesign und -methodik für die Studie von Kaku, 2009

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Die Studie untersuchte die Wirksamkeit und Sicherheit der Kombinationstherapie aus Pioglitazon und Metformin
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Bei der Studie handelt es sich um eine doppelblinde, placebokontrollierte, parallelgruppen-geführte (im Verhältnis 1:1) Sicherheits- und Wirksamkeitsstudie.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderungen der Methodik nach dem Studienbeginn
4	Probanden / Patienten	Mit Metformin allein nicht ausreichend kontrollierte Typ-2-Diabetiker
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Weibliche oder männliche, ambulant behandelte Patienten, im Alter von 20-64 Jahren, die mit Diät und Körperübungen, aber mit keinen anderen antidiabetischen Medikamenten als Metformin behandelt wurden</p> <p>Ausschlusskriterien: Patienten mit Diabetes Typ 1; Patienten mit eingeschränkter Leberfunktion, Herzfehler oder anderen ernsthaften Herzerkrankungen; Hinweise auf andere klinisch relevante Komplikationen wie Krebs, schwere Lungenerkrankung, gastrointestinale pankreatische oder hämatologische Erkrankungen Krankheitsgeschichte mit Laktatazidose, Ketoazidose oder diabetischem Koma oder Präkoma in den vorausgehenden 26 Wochen Patienten mit Drogen- oder Alkoholabhängigkeit Medikamente, die die glykämische Kontrolle beeinflussen könnten, waren nicht erlaubt</p>
4b	Studienorganisation und Ort der Studiendurchführung	25 Zentren innerhalb Japans
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Pioglitazon (15 mg oder 30mg) verglichen mit Placebo; beide einmal täglich versetzt mit Metformin (500 mg oder 750 mg)</p> <p>Während des 12-wöchigen Beobachtungszeitraums erhielten die Patienten Metformin 500 mg oder 750 mg pro Tag. Diese Dosis wurde während der Studie beibehalten. Patienten mit einem HbA1c-Wert von ≥ 6.5 und $< 10\%$ wurden</p>

Item ^a	Charakteristikum	Studieninformation
		zufällig ausgewählt und erhielten einmal täglich Pioglitazon 15 mg vor oder nach dem Frühstück oder Placebo für einen Zeitraum von 12 Wochen. Wurde die Medikamentendosierung gut vertragen, wurde die Dosis in den Wochen 13 bis 28 mg auf 30 mg erhöht.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Veränderung des HbA1c-Wertes zwischen der Eingangsuntersuchung und nach 28 Wochen Sekundäres Zielkriterium: Zeitlicher Verlauf des HbA1c-Werts und der Nüchternblutglukose, Prozentzahl an Patienten, die einen HbA1c <6.5% haben
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurden keine Änderungen vorgenommen.
7	Fallzahl	Pioglitazon 15 mg einmal täglich: n=83; Placebo einmal täglich: n=86
7a	Wie wurden die Fallzahlen bestimmt?	Auf bestehende Daten (Einhorn 2000) gestützt wurde kalkuliert (mit Hilfe des 'nQuery Advisor 5.0'; Statistical Solutions Ltd), dass 72 evaluierbare Patienten pro Behandlungsarm eine 90%ige Wahrscheinlichkeit ergeben, mit der Verwendung eines zweiseitigen T-Tests einen signifikanten Unterschied von 5% zwischen den beiden Behandlungsgruppen zu erkennen. Unter Berücksichtigung von 10% Rücknahmen und Abbrüchen wurde mit 160 Patienten die Behandlungsphase geplant.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung wurde durch ein unabhängiges Kontrollgerät stratifiziert, um eine gleiche Verteilung zwischen den Gruppen mit Metformin 500 mg und 750 mg zu garantieren.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die	Nicht berichtet

Item ^a	Charakteristikum	Studieninformation
	Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar: Die Autoren bezeichneten diese Studie als doppelblind, aber berichteten nicht explizit, wer verblindet wurde oder wie die Verblindung durchgeführt wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die letzte Messung vor der Beendigung oder Vervollständigung des Protokolls wurde als behandlungsabschließende Messung betrachtet.</p> <p>Für die Primär- und Sekundärendpunkte und andere Effektivitätskriterien wurden Mittelwert, Standardabweichung und zweiseitige 95% KI, sowie ein zweiseitiger Einstichproben-t-Test vorgenommen. Für alle statistischen Analysen wurde ein Wert von $p < 0,05$ als statistisch signifikant eingestuft.</p> <p>Abhängig von den Daten wurden im Berichtszeitraum die demografischen Informationen mit Hilfe von Chi-Square, zweiseitigem Wilcoxon und Zweistichproben-t-Test auf Homogenität analysiert. Ein p-Wert von < 0.2 wurde indikativ eingestuft für ein Ungleichgewicht zwischen den beiden Gruppen bei Beginn der Behandlung.</p> <p>Eine zweifaktorielle Varianzanalyse (ANOVA; Kombinations- versus Monotherapie) wurde durchgeführt, um den unabhängigen Effekt der Metformin-Dosis festzustellen. Eine Kovarianzanalyse wurde benutzt, um die Auswirkung von demografischen Faktoren wie Geschlecht, BMI und dem vor der Behandlung festgestellten HbA1c auf den Studienendpunkt festzustellen. Fisher's Test und zweiseitige KI wurden benutzt, um unerwünschte Ereignisse zwischen der Gruppe mit Kombinationstherapie und der Gruppe mit Monotherapie zu vergleichen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht zutreffend
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 32)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die	Siehe <i>Flow-Chart</i> (Abbildung 32)

Item ^a	Charakteristikum	Studieninformation
	a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 32)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Studie wurde 2005 initiiert. Sie dauerte 40 Wochen (12 Wochen Beobachtungszeitraum und 28 Wochen Behandlungszeitraum); Messungen wurden alle vier Wochen vorgenommen
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht feststellbar - endet nach Protokoll nach Woche 40
a: nach CONSORT 2010.		

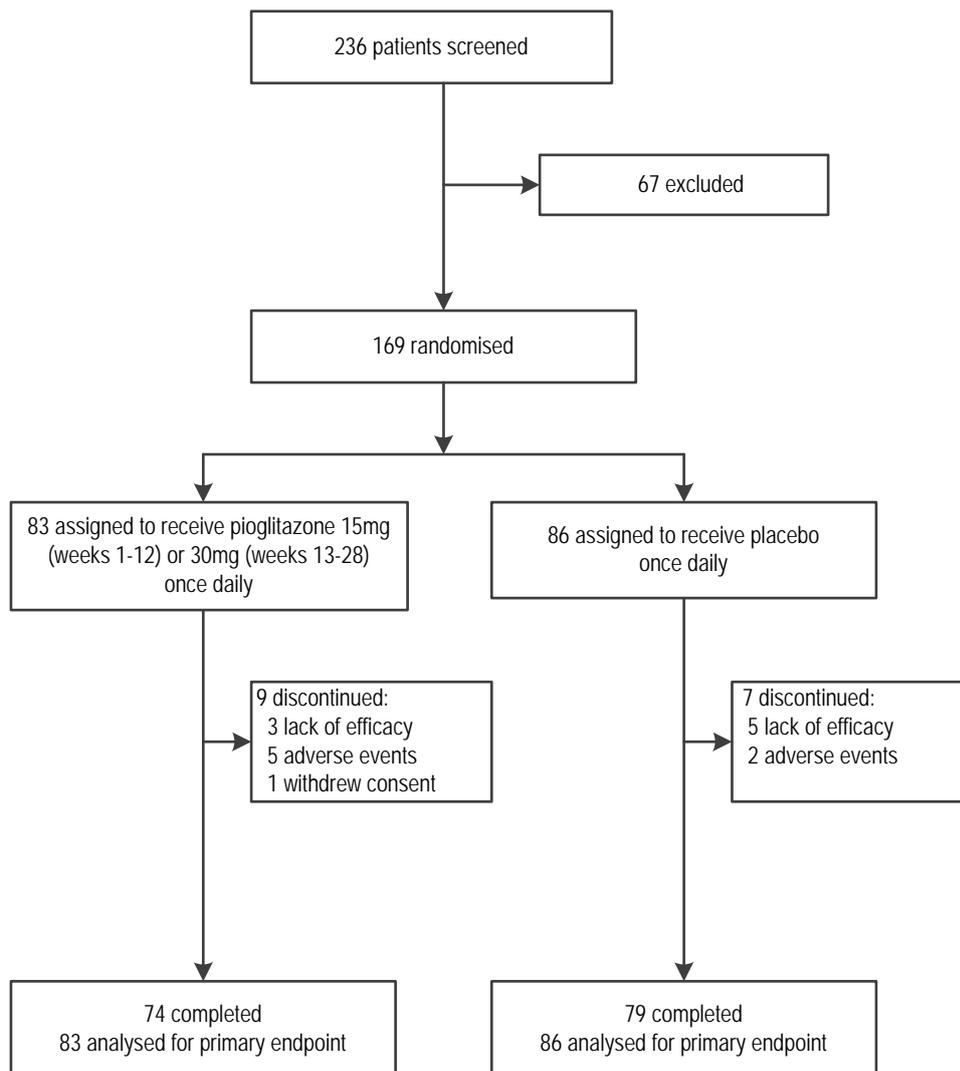


Abbildung 32: Flussdiagramm Studiendesign und -methodik für die Studie von Kaku, 2009

Tabelle 4-174 (Anhang): Studiendesign und -methodik für die Studie von Charpentier 2001

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Diese Studie wurde entwickelt, um die glykämische Kontrolle nach einem Wechsel zu einer Behandlung mit Glimepirid oder einer Kombinationstherapie aus Glimepirid und Metformin bei Patienten festzustellen, die durch Metformin allein nicht ausreichend kontrolliert waren.
-	Methoden	
3	Studiendesign	Randomisierte, kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, doppelblinde, parallelgruppen-geführte (im Verhältnis 1:2), multizentrische Studie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	Mit Metformin allein nicht ausreichend kontrollierte Typ-2-Diabetiker
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Patienten im Alter zwischen 35 und 70 Jahren mit Typ-2-Diabetes, die mit Metformin Monotherapie 255 mg täglich für mindestens 4 Wochen nicht ausreichend kontrolliert waren, d. h. Nüchternblutzucker ($7,8 \text{ mmol/L} < \text{FBG} \leq 13,9 \text{ mmol/L}$) und einem Serumkreatininwert von $< 110 \text{ } \mu\text{mol/L}$.</p> <p>Neu diagnostizierte Patienten (< 1 Jahr) wurden eingeschlossen, vorausgesetzt, der BMI war $\geq 23,0 \text{ kg/m}^2$ bei Frauen und bei Männern $\geq 25,0 \text{ kg/m}^2$ und es gab keine Hinweise auf spontanen Gewichtsverlust oder Ketonurie in Verbindung mit Glykosurie.</p> <p>Ausschlusskriterien:</p> <p>Sekundärer oder insulin-abhängiger Diabetes, jegliche schwere Krankheiten, Übergewicht vom Grad 3 ($\text{BMI} \geq 40,0 \text{ kg/m}^2$), Krankheitsgeschichte mit erheblichen kardiovaskulären Ereignissen innerhalb der letzten 6 Monate, Allergie auf Sulfonylharnstoffe oder Drogen- / Alkoholabhängigkeit.</p>
4b	Studienorganisation und Ort der Studiendurchführung	34 Zentren in Frankreich
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Kombination von Glimepirid und Metformin (Glimepirid verabreicht in einer Dosis von 1 mg täglich, die sich stufenweise zu 2,4 mg oder 6 mg täglich erhöht, um eine Nüchternblutzucker innerhalb der Zielkriterien zu erreichen [$3,9 \text{ mmol/L} \leq \text{FBG} \leq 7,8 \text{ mmol/L}$], verglichen mit Metformin plus Placebo (850 mg Metformin dreimal täglich)

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	Veränderung des HbA1c-Wertes zwischen der Eingangsuntersuchung und nach 20 Wochen
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Veränderung des HbA1c-Wertes zwischen der Eingangsuntersuchung und nach 20 Wochen Sekundäres Zielkriterium: Blutzucker (Nüchternblutzucker und post-prandialer Blutzucker, der 90 Minuten nach Ende des Frühstücks gemessen wurde), Insulin, C-Peptide, Triglyceride, Gesamtcholesterin, <i>High-Density</i> -Lipoprotein (HDL)-Cholesterin, Apolipoprotein B, diastolischer und systolischer Blutdruck im Sitzen, Gewicht und BMI
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	Randomisierte Fallzahl: Glimepirid kombiniert mit Metformin n=147 Metformin plus Placebo n=75
7a	Wie wurden die Fallzahlen bestimmt?	Nicht berichtet
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht feststellbar – Studie wurde wie geplant ausgeführt
8	Randomisierung, Erzeugung der Behandlungsfolge	Behandlungssequenz wurde zufällig erzeugt von der Abteilung für Biometrie, Hoechst Marion Roussel, Frankreich
8a	Methode zur Generierung der zufälligen Zuteilung	Zeitplan wurde erstellt mit Hilfe einer Minimierungsprozedur für Nüchternblutzucker (7.8 mmol/L to < 11.1 mmol/L or ≥11.1 mmol/L to ≤13.9 mmol/L), BMI (<28.0 kg/m ² oder ≥28.0 kg/m ² je nach Geschlecht) und einem Zentrum, welches fortlaufende Zahlen zu Behandlungscodes verband, die zufällig vergeben wurden.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Der Plan wurde für jedes Zentrum erstellt. Für die Studienbehandlung (Metformin, Glimepirid, Kombinationsbehandlung) wurde ein Zuteilungsverhältnis von 1:2:2 verwendet.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Der zufällig erzeugte Plan verband fortlaufende Zahlen zu Behandlungscodes, die zufällig verteilt wurden.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Randomisierung wurde von einem zentralen Randomisierungszentrum via Minitel zu jedem, Studienstandort durchgeführt. Die Patienten wurden randomisiert nach der Bestätigung der

Item ^a	Charakteristikum	Studieninformation
		demografischen und anderer grundlegenden Daten durch den Prüfarzt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar: Die Autoren bezeichneten diese Studie als doppelblind, aber berichteten nicht explizit, wer verblindet wurde oder wie die Verblindung durchgeführt wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Autoren bezeichneten diese Studie als „doppelblind“, bzw. „double-dummy“.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Alle randomisierten Patienten, die mindestens eine Dosis des untersuchten Medikaments und sowohl einen Ausgangswert als auch mindestens einen Behandlungswert hatten, wurden ausgewählt für eine ITT-Wirksamkeitsanalyse. Alle Patienten, die mindestens eine Dosis des untersuchten Medikaments nahmen, wurden in eine <i>Safety</i> Analyse eingeschlossen.</p> <p>Alle Analysen wurden mit einer SAS <i>Software</i> (Version 6.08) durchgeführt. Die ITT-Analyse wurde mit einer Kovarianzanalyse des Behandlungseffekts durchgeführt, adjustiert an die Ausgangswerte, mit Angabe der fehlenden Werte des primären (Angaben berücksichtigen den letzten Nüchternblutzucker-Wert). Der primäre Endpunkt wurde unter Ausschluss der Studienabbrecher analysiert, um die Robustheit der Ergebnisse und den Einfluss auf die Bewertungen einschätzen zu können.</p> <p>Wenn der zweiseitige p-Wert, in Verbindung mit der Veränderung des HbA1c-Wertes am Endpunkt in den mit Glimepirid oder mit Kombinationstherapie behandelten Gruppen, ≤ 0.025 war, wurde Glimepirid allein oder die Kombination aus Glimepirid und Metformin als besser bewertet als Metformin allein.</p> <p>Außerdem wurde, wenn die zweiseitigen p-Werte, in Verbindung mit der Veränderung des HbA1c-Wertes am Endpunkt in den mit Glimepirid oder mit Kombinationstherapie behandelten Gruppen, ≤ 0.05 betragen, die Behandlung mit Glimepirid oder die Kombinationstherapie mit Glimepirid als wesentlich wirksamer bewertet als die Behandlung Metformin alleine. Zusätzlich wurden 90% und 95% KI für jeden Vergleich berechnet. Wenn die obere Grenze des 90% KI ≤ 0.3 war, wurden beide Behandlungen als klinisch gleichwertig eingestuft. Enthielt der 95% KI keine 0, wurde damit die Überlegenheit einer Behandlung gegenüber der anderen festgestellt.</p>

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Ein zusätzliches Modell zur Ausrichtung nach Zentrum, BMI und Geschlecht war ebenfalls eingeschlossen. Wenn es erforderlich war, wurden angemessene Umwandlungen der Variablen verwendet.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 33)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 33)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 33)
14	Aufnahme / Rekrutierung	Daten bezüglich der Patientenrekrutierung wurden nicht berichtet.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Patienten wurden 20 Wochen nach der Randomisierung beobachtet, mit Untersuchungen in Wochen 0, 3, 6, 9, 12 und 20
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht feststellbar - endet nach Protokoll nach Woche 20
a: nach CONSORT 2010.		

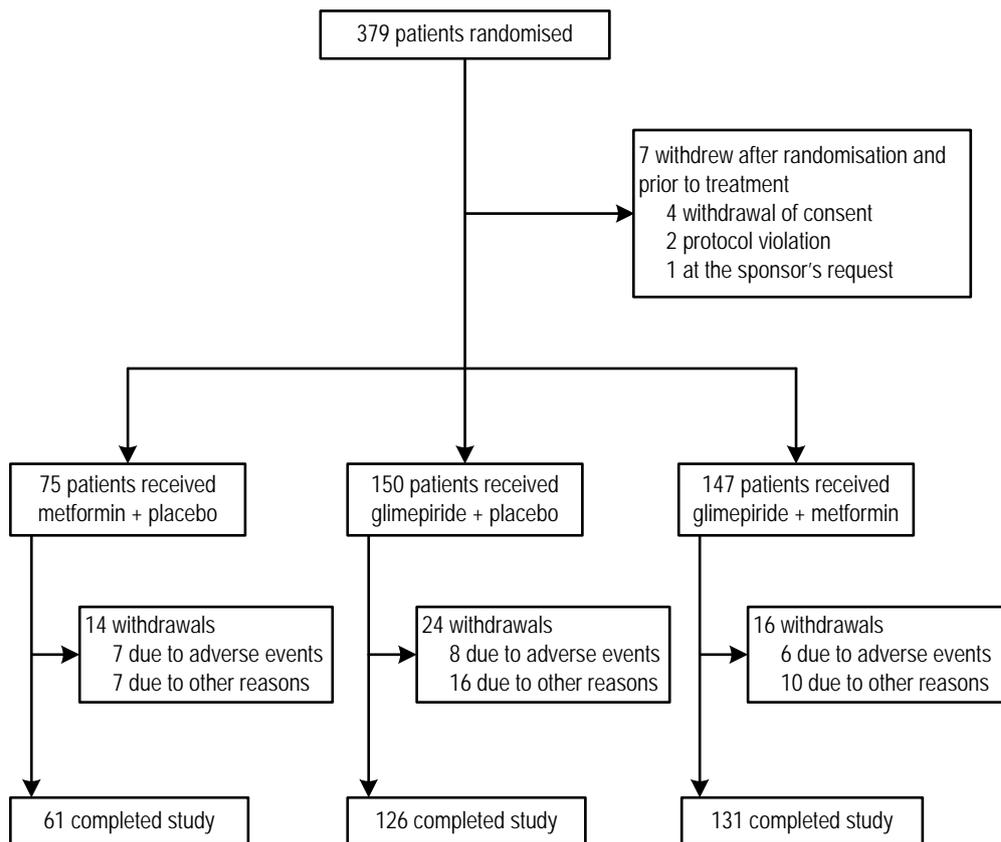


Abbildung 33: Flussdiagramm Studiendesign und -methodik für die Studie von Charpentier 2001

Anmerkung: Randomisierung in die drei Arme Metformin+Placebo, Glimperid+Placebo und Glimperid+Metformin (im Verhältnis 1:2:2), aber nur der Metformin+Placebo-Arm und der Glimperid+Metformin-Arm wurden für die Nutzenbewertung verwendet

Tabelle 4-175 (Anhang): Studiendesign und -methodik für die Studie von Papathanassiou, 2009

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Das Ziel der Studie war, den Effekt von Glimperid und Pioglitazon auf die endotheliale Funktion bei Patienten mit Metformin-behandeltem Typ-2-Diabetes zu vergleichen.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Offene, aktivkontrollierte randomisierte Studie mit Paralleldesign (Verhältnis 1:1)
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet.
4	Probanden / Patienten	Typ-2-Diabetiker, die unzureichend mit Metformin behandelt wurden.
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: Typ-2-Diabetiker, die 6 Monate vor Studienbeginn mit Metformin behandelt wurden; Glykiertes Hämoglobin (HbA1c) >6,5% und normale Leberenzymwerte und Nierenfunktion Ausschlusskriterien: Koronare Herzerkrankung, zerebrovaskuläre oder periphere vaskuläre Erkrankung in der Vergangenheit, chronische Herzinsuffizienz, Leber- oder Nierenerkrankung, Anämie, Störung der Schilddrüsenfunktion und neue Einnahme einer beliebigen Medikation innerhalb der vorherigen 8 Wochen
4b	Studienorganisation und Ort der Studiendurchführung	Griechenland
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Glimperid 4 mg einmal täglich (n=14) wurde verglichen mit Pioglitazon 30 mg einmal täglich (n=14), beide zusätzlich zur Metformin-Behandlung
6	Zielkriterien	Beurteilung der endothelialen Funktion durch flussvermittelte Vasodilatation
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäre Zielmessungen waren: Beurteilung der endothelialen Funktion durch flussvermittelte Vasodilatation Sekundäre Zielmessungen waren:

Item ^a	Charakteristikum	Studieninformation
		Glukose, Insulin, Glykiertes Hämoglobin, Lipidkonzentrationen, Leber- und Nierenfunktion, großes Blutbild, BMI, minimale Taillenumfangmessung zwischen Beckenrand und Rippenbogen
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	Anzahl der randomisierten Probanden: Glimepirid 4 mg einmal täglich: n=14; Pioglitazon 30 mg einmal täglich: n=14
7a	Wie wurden die Fallzahlen bestimmt?	Diese Stichprobengröße ermöglicht eine statistische <i>Power</i> von annähernd 85% um einen 70%-Unterschied in den Effekten von Pioglitazon vs. Glimepirid in der flussvermittelten Vasodilatation aufzudecken.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Keine Angabe – alle randomisierten Probanden haben die Studie abgeschlossen.
8	Randomisierung, Erzeugung der Behandlungsfolge	Basiert einfach auf der Reihenfolge der Vorstellung der geeigneten Patienten in der Ambulanz der Klinik.
8a	Methode zur Generierung der zufälligen Zuteilung	Für die Studie geeignete Patienten, die zur Teilnahme einwilligten, wurden randomisiert zugeteilt, einfach basierend auf der Reihenfolge der Vorstellung in der Ambulanz der Klinik.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Patienten wurden in 2 Gruppen randomisiert (Blockrandomisierung 1:1)
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Für die Studie geeignete Patienten, die zur Teilnahme einwilligten, wurden randomisiert in 2 Gruppen (Blockrandomisierung 1:1) zugeteilt, einfach basierend auf der Reihenfolge der Vorstellung in der Ambulanz der Klinik.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Unklar; Der Bericht enthält unzureichende Informationen angesichts der Person, die die zufällige Zuweisung durchführt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Es handelt sich um eine offene Studie. Keiner wurde verblindet, außer den Auswertern der endothelialen Funktion und des Durchmessers der Oberarmarterie, die für die Behandlung der Patienten verblindet wurden.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Ergebnisse für stetige Variablen wurden als Mittelwert \pm Standardabweichung dargestellt. Der Kolmogorov-Smirnov Z-Test

Item ^a	Charakteristikum	Studieninformation
		<p>wurde zur Bestimmung der Normalverteilung der stetigen Variablen benutzt; alle Variablen wurden als normalverteilt eingestuft. Der χ^2-Test und der ungepaarte <i>Student's</i> T-Test wurden benutzt, um kategoriale und stetige Variablen zu vergleichen, insbesondere zwischen den zwei Gruppen am Ausgangspunkt.</p> <p>Um Veränderungen zum Ausgangswert innerhalb jeder Gruppe abschätzen zu können wurde der gepaarte <i>Student's</i> T-Test benutzt. Die Datenanalyse beruht auf ITT.</p> <p>Um Veränderungen in den physikalischen, vaskulären Messungen und Labormessungen zwischen den 2 Behandlungsgruppen bewerten zu können (zum Zeitpunkt des Follow-ups im Vergleich zum Ausgangswert) wurde die Methode der <i>Repeated Measures Analysis of Variance (RM ANOVA)</i> benutzt, welche für <i>Confounder</i> adjustiert wurde. Einfache lineare Regression und partielle Korrelationsanalyse wurden durchgeführt, um die Beziehung zwischen Veränderungen in der flussvermittelten Vasodilatation und der betrachteten klinischen Variablen, Labormessungen und der Behandlungsart zu bewerten.</p> <p>Ein p-Wert von <0,05 wurde als signifikant betrachtet. Zur Analyse wurde das SPSS Statistik-Paket (Version 15.0 für Windows, SPSS Inc. Chicago, IL, USA) benutzt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 34)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 34)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 34)
14	Aufnahme / Rekrutierung	Die Angaben zum Rekrutierungszeitraum wurden nicht berichtet.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Patienten wurden nach der Randomisierung für 6 Monate beobachtet, und die klinischen Beurteilungen erfolgten in den Monaten 1, 3 und 6.
14b	Informationen, warum die Studie endete oder	Keine Angabe – Die Studie endete wie nach

Item ^a	Charakteristikum	Studieninformation
	beendet wurde	Angaben im Protokoll nach 6 Monaten
a: nach CONSORT 2010.		

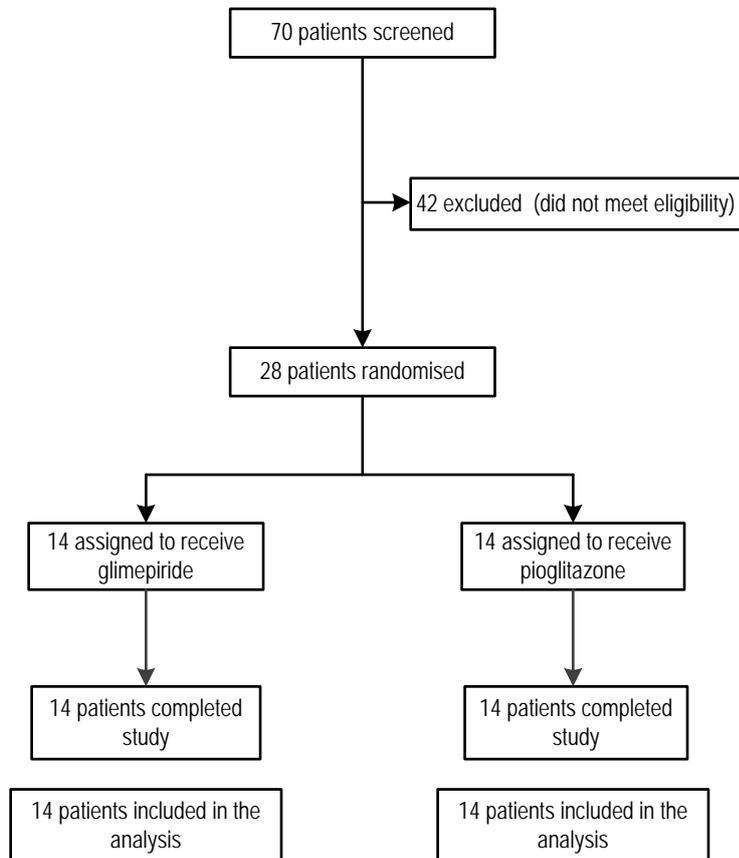


Abbildung 34: Flussdiagramm Studiendesign und -methodik für die Studie von Papathanassiou 2009

Tabelle 4-176 (Anhang): Studiendesign und -methodik für die Studie von Matthews, 2005

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Die laufende Studie untersucht die Langzeitwirksamkeit, Sicherheit und Verträglichkeit der <i>Add-on</i> -Therapie von Pioglitazon, kombiniert mit Sulphonylharnstoff unter weitergeführter Metforminbehandlung bei Patienten mit Typ-2-Diabetes, die mit Metformin allein unzureichend kontrolliert waren.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, parallelgruppen-geführte (im Verhältnis 1:1), <i>double-dummy</i> Studie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Es wurden keine methodischen Veränderungen aufgeführt.
4	Probanden / Patienten	Typ-2-Diabetiker, die mit Metformin allein unzureichend kontrolliert waren (mit $\geq 50\%$ der empfohlenen maximalen Dosis oder mit der maximal erlaubten Dosis für ≥ 3 Monate)
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Alter von 35 bis 75 Jahren; HbA1c von $\geq 7,5\%$ oder $\leq 11,0\%$; Nüchtern-C-Peptid von $\geq 1,5$ ng/mL (0,50 nmol/L) und stabile oder schlechter werdende glykämische Kontrolle für ≥ 3 Monate vor dem <i>Screening</i>.</p> <p>Ausschlusskriterien: Patienten mit Diabetes Typ 1; Ketoazidose, Herzinfarkt, transitorische ischämischen Attacken oder Schlaganfall in den vorangegangenen 6 Monaten; symptomatischer Herzfehler; akute Malabsorption oder chronische Pankreatitis; familiäre adenomatöse Polyposis; maligne Erkrankungen in den letzten 10 Jahren oder Drogenmissbrauch. Weibliche Patienten mussten postmenopausal oder sterilisiert sein oder ausreichende Empfängnisverhütung verwenden, schwangere oder stillende Frauen wurden ausgeschlossen. Vorherige Behandlungen mit Insulin, Gliclazid, Pioglitazon oder anderen Sulphonylharnstoffen oder Glitazonen waren nicht erlaubt. Während der Studie waren Thiazide zur Behandlung von Ödemen erlaubt; falls eine antihypertensive Behandlung indiziert war, wurden ACE-Inhibitoren, Angiotensin-II-Antagonisten oder Calcium-Antagonisten gegeben.</p>
4b	Studienorganisation und Ort der Studiendurchführung	75 Zentren in 9 europäischen Ländern und Australien

Item ^a	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Pioglitazon 15 mg pro Tag (n=317) oder Gliclazid 80 mg pro Tag (n=313), plus Metformin in der Dosis, die vor der Studie bereits eingenommen wurde; die Dosis an Pioglitazon wurde zu 30 mg und 45 mg verabreicht, die Dosis an Gliclazid zu 160 mg, 240 mg (160 mg und 80 mg) und 320 mg (160 mg zweimal täglich). Die Unterbrechung der Titration bzw. eine Heruntertitrierung war nur erlaubt bei Unverträglichkeit, aktueller Hypoglykämie oder einem erhöhten Risiko für Hypoglykämie. Die Patienten führen mit der nächsthöheren Dosierung fort, außer der Prüfarzt stellte fest, dass eine Erhöhung der Dosis zum Risiko einer Hypoglykämie führe (Dosiserhöhung wurde für eine Untersuchung von Woche 4 auf Woche 8 verschoben oder die Dosis von Woche 8 wurde für den Rest der Studie beibehalten) oder der Patient berichtete symptomatische Hypoglykämie (einstufige Reduktion) oder der Patient erlebte ein unerwünschtes Ereignis, das eine Dosisreduzierung erforderte (einstufige Reduktion in Woche 8, 12 oder 16 mit keiner weiteren Heruntertitrierung). Die Dosis, die in Woche 16 erreicht wurde, wurde für die verbleibenden 36 Wochen beibehalten. Es wurde keine Reduzierung der Metformindosis, die vor Studienbeginn eingenommen wurde, erlaubt.
6	Zielkriterien	Der primäre Endpunkt war die Veränderung der HbA1c-Werte zum Ausgangswert bis Woche 52
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Veränderung der HbA1c-Werte zum Ausgangswert bis Woche 52; Messungen des HbA1c wurde in Wochen 0, 4, 8, 12, 16, 24, 32, 42 und 52 durchgeführt. Sekundäres Zielkriterium: Veränderungen des Nüchternblutzuckers, Insulin, Lipide, C-Peptide, das Spaltprodukt des 32/33- Proinsulin, Albumin im Urin und Kreatinin (um das Albumin-Kreatinin-Verhältnis zu bestimmen); Messungen der Nüchternblutzuckers und des Insulins wurden zu Beginn der Studie und in Wochen 8, 16, 24, 32, 42 und 52 durchgeführt; der atherogene Index des Plasma wurde als Logarithmus berechnet (Triglycerid: HDL-Cholesterin Verhältnis) als ein Index der LDL-Partikelgröße (Dobiasova 2001); Urinalbumin und Kreatinin wurden gemessen in Wochen 0, 24, 32, 42 und 52; in ausgewählten Zentren wurden C-Peptide und das Spaltprodukt des 32/3-Proinsulin gemessen in Wochen 0, 24, 42 und 52.

Item ^a	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurden nach Beginn der Studie keine Veränderungen an den Zielkriterien vorgenommen.
7	Fallzahl	Randomisierte Fallzahl: Prioglitazon 15 mg täglich: n=317, Gliclazid 80 mg täglich: n=313
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl wurde festgelegt durch die Demonstration einer Zwischengruppen-Differenz von 0,35% bei der Veränderung des HbA1c-Wertes vom Ausgangswert zu Woche 52 (der primären Effektivitätsvariablen) durch Gebrauch eines zweiseitigen T-Tests. Insgesamt wurden auf Grundlage eines Levels von 95% Power und 5% Signifikanz 225 Patienten pro Gruppe benötigt, die 24 Wochen der Studie absolvieren.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht feststellbar
8	Randomisierung, Erzeugung der Behandlungsfolge	Computerbasierte Generierung (aus Charbonnel 2005)
8a	Methode zur Generierung der zufälligen Zuteilung	Computerbasierte Generierung (aus Charbonnel 2005)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Blockrandomisierung (aus Charbonnel 2005)
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Patienten wurden via ein zentrales Telefonsystem randomisiert (QTONE) (aus Charbonnel 2005).
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar: Die Autoren bezeichneten diese Studie als doppelblind, aber berichteten nicht explizit, wer verblindet wurde oder wie die Verblindung durchgeführt wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Veränderung des HbA1c wurde mit Hilfe des ANCOVA Modells mit dem Faktor „Behandlung“ und dem Ausgangswert als Kovariate festgestellt. Der Unterschied zwischen den adjustierten Mittelwerten

Item ^a	Charakteristikum	Studieninformation
		<p>wurde berechnet mit 95% KI und einem zweiseitigen T-Test ($\alpha = 0.05$).</p> <p>Andere Effektivitätsvariablen wurden auf ähnliche Weise analysiert. Ergebnisse zu den Lipiden und dem Urinalbumin-Kreatinin-Verhältnis wurden ebenso vor der Analyse logartithmiert, da bestimmte Daten nicht normal-verteilt waren. LDL-Cholesterin wurde mit der Friedewald-Formel berechnet.</p> <p>Beschreibende Statistiken wurden verwendet, um Änderungen in Körpergewicht, Demographien, Laborwerte, <i>Baseline</i> Charakteristika und unerwünschte Ereignisse zusammenzufassen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 35)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 35)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Siehe <i>Flow-Chart</i> (Abbildung 35)</p> <p>Die Analyse wurde an der <i>ITT</i>-Population durchgeführt, die alle Patienten einschloss, die mindestens eine Dosis der Studienmedikation eingenommen haben, und deren HbA1c als Ausgangswert und mindestens einmal nach dem Ausgangswert aufgezeichnet wurde. Andere Effektivitätsvariablen wurden auf ähnliche Weise analysiert.</p>
14	Aufnahme / Rekrutierung	Daten bezüglich der Patientenrekrutierung wurden nicht berichtet.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	<p>Die 52-wöchige Studienphase wurde eingeteilt in eine 16-wöchige Titrationsphase mit festgelegter Dosis und eine 36-wöchige Erhaltungsphase. Messungen des HbA1c, FPG und Insulin wurden zu Beginn der Studie, in Wochen 4, 8, 12, 16, 24, 32, 42 und 52 gemacht. Lipide (Triglycerid, HDL-Cholesterin, LDL-Cholesterin und Gesamtcholesterin) wurden zu Beginn der Studie und in Wochen 8, 16, 24, 32, 42 und 52 gemessen. Urinalbumin und Kreatinin wurden zu Beginn der Studie und in Wochen 24, 32, 42 und 52 gemessen. In ausgewählten Zentren wurden C-Peptide und das Spaltprodukt des 32/33- Proinsulin gemessen.</p>

Item ^a	Charakteristikum	Studieninformation
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht feststellbar
a: nach CONSORT 2010.		

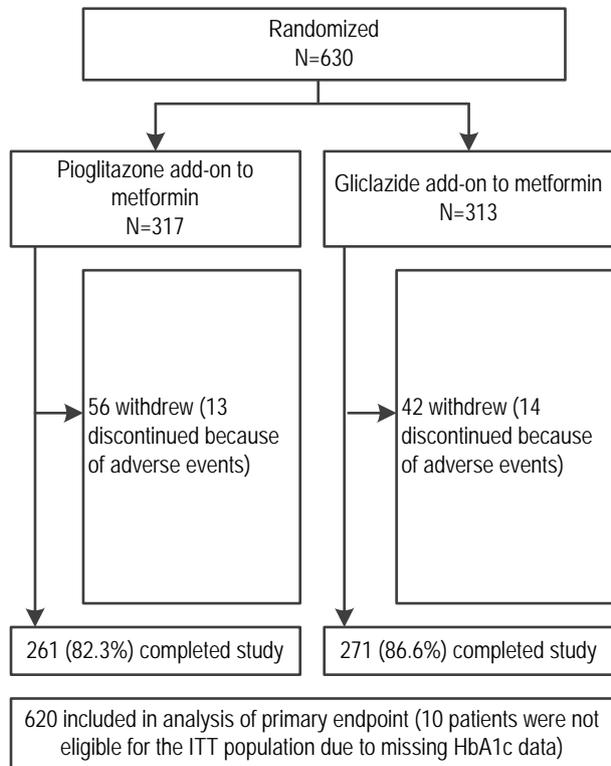


Abbildung 35: Flussdiagramm Studiendesign und -methodik für die Studie von Matthews, 2005

Tabelle 4-177 (Anhang): Studiendesign und -methodik für die Studie von Nauck, 2007

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Diese Studie untersuchte die Wirksamkeit und Sicherheit von Sitagliptin vs. Glipizid bei Patienten mit Typ-2-Diabetes, die eine nicht ausreichende glykämische Kontrolle (HbA1c ≥ 6.5 und $\leq 10\%$) unter Metformin-Monotherapie aufwiesen.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, aktivkontrollierte, parallelgeführte (Verhältnis 1:1) Nicht-Unterlegenheitsstudie ^b
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderung der Methodik nach dem Studienbeginn
4	Probanden / Patienten	Typ-2-Diabetiker durch Metformin allein unzureichend therapiert
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Männer und Frauen im Alter von 18 bis 80 Jahren, diagnostiziert mit Typ-2-Diabetes, die zur Zeit keine orale antihyperglykämische (OHA) Therapie erhielten, mit einer OHA-Monotherapie oder mit Metformin in Kombination mit einer OHA behandelt wurden, waren potenziell wählbar für die Studie, wenn sie alle Studienkriterien erfüllten.</p> <p>Patienten, die schon unter Metformintherapie ≥ 1500 mg/Tag waren und einen HbA1c-Wert ≥ 6.5 und $\leq 10\%$ hatten, traten direkt in eine 2-wöchige Placebo-Anlaufphase und waren wählbar für die Randomisierung. Patienten, die nicht zur Zeit eine OHA-Therapie erhielten, Patienten, die eine andere OHA-Monotherapie als Metformin ≥ 1500 mg/Tag erhielten oder Patienten, die Metformin in Kombination mit einer anderen OHA erhielten, traten in eine Titrierungs- und dosisstabile Phase mit Metformin-Monotherapie für mindestens 8 Wochen ein.</p> <p>Ausschlusskriterien:</p> <p>Typ-1-Diabetes, Insulingebrauch innerhalb 8 Wochen vor Eingangsuntersuchung, Verschlechterung der Nierenfunktion unvereinbar mit Metformin oder Nüchternplasmaglukose (FPG) (oder Nüchtern-fingerstick-Glucose) > 15.0 mmol/L (270 mg/dL) während oder kurz vor der Randomisierung. Andere hyperglykämische Behandlungen waren nicht während der Studie zugelassen. Gleichzeitige lipidsenkende und antihypertensive</p>

Item ^a	Charakteristikum	Studieninformation
		Medikationen, Hormonersatztherapie und Medikamente zur Empfängnisverhütung waren erlaubt, ohne dabei ihre Dosierung zu verändern.
4b	Studienorganisation und Ort der Studiendurchführung	Multinational
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Sitagliptin 100 mg einmal täglich und Glipizid 5-20 mg einmal täglich; beide zusammen mit Metformin (≥ 1500 mg einmal täglich) verabreicht. Nach der Anfangsdosis von 5 mg/Tag, wurde Glipizid gemäß den im Protokoll definierten Kriterien auf eine potentielle Maximaldosis von 20 mg/Tag titriert. Jeweils in 3-wöchigen Abständen während der ersten 18 Wochen, wurde Glipizid auftitriert, wenn <i>fingerstick</i> -Glucose-Werte vor den Mahlzeiten $>6,1$ mmol/L (110 mg/dL) waren. Das Auftitrieren von Glipizid wurde abgebrochen, wenn nach der Meinung des Prüfers das Auftitrieren dem Patienten das Risiko einer Hypoglykämie bedeutete. Die Glipizid-Dosis konnte jede Zeit während der Studie verringert werden, um wiederholte hypoglykämische Ereignisse zu vermeiden.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Veränderung des HbA1c-Wertes zwischen der Eingangsuntersuchung und nach 52 Wochen. Sekundäre Zielkriterien: Beständigkeit der HbA1c-Reduktion, gemessen mit dem Vergleich der Anstiegsrate von HbA1c von Woche 24 mit Woche 52; Anteil der Patienten, die HbA1c $<7\%$ oder $<6,5\%$ erreichten, FPG, Insulin-, Proinsulin- und Lipidparameter [Gesamtcholesterin, <i>low-density</i> Lipoprotein-Cholesterin, Triglyceride, <i>high-density</i> Lipoprotein-Cholesterin (HDL-C) und <i>non</i> -HDL-C] wurden am Studienanfang und zu verschiedenen Zeitpunkten während der Studie verglichen. β -Zellfunktion gemäß <i>Homeostasis model assessment</i> (HOMA) und Proinsulin/Insulin-Verhältnis wurden verwendet, um β -Zellaspekte zu beurteilen (Tura 2003, Matthews 1985). HOMA-Insulinresistenz und <i>quantitative insulin sensitivity check index</i> wurden berechnet, um die Änderungen in Insulinresistenz zu beurteilen (Matthews 1985, Katz 2000).
6b	Änderungen der Zielkriterien nach Studienbeginn,	Keine Änderungen der Zielkriterien wurden

Item ^a	Charakteristikum	Studieninformation
	mit Begründung	nach dem Studienbeginn vorgenommen
7	Fallzahl	Zahl der randomisierten Patienten: Sitagliptin 100 mg einmal täglich: n=588; Glipizid 5-20 mg einmal täglich: n=584
7a	Wie wurden die Fallzahlen bestimmt?	Nicht berichtet
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar; die Autoren beschreiben die Studie als doppelblind, aber beschreiben nicht explizit wer verblindet war und wie die Verblindung durchgeführt wurde
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Für die HbA1c-Änderung vom Ausgangswert, wurde Sitagliptin gegenüber Glipizid nicht-unterlegen betrachtet, wenn die obere Grenze des zweiseitigen 95% Konfidenzintervalles (CI) für die durchschnittliche Differenz zwischen Sitagliptin und Glipizid kleiner als die Marge $\delta=0.3\%$ war. Diese Marge wurde so ausgewählt, dass für die aufzuzeigende Nicht-Unterlegenheit (d. h. die Obergrenze des Konfidenzintervalles kleiner als die gewählte Marge) der beobachtete Zwischen-Gruppenunterschied klein sein würde. Das Modell zur Kovarianzanalyse wurde verwendet, um die Behandlungsgruppen bezüglich der Wirksamkeitsendpunkte zu vergleichen, fokussierend auf die Änderung zwischen den Ausgangswerten und den Werten in Woche 52, mit Ausgangswert und

Item ^a	Charakteristikum	Studieninformation
		<p>Status vor der OHA-Behandlung als Kovariaten. Die Differenz zwischen Sitagliptin und Glipizid bezüglich Wirksamkeitsendpunkte wurde durch das Testen der Differenz bei den kleinsten Quadraten (<i>least squares mean change or mean per cent change</i>) zwischen Ausgangswerten und Werten in Woche 52 bestimmt. Zusätzliche Wirksamkeitsanalysen basierten auf der Population aller randomisierten Patienten, die mindestens eine Dosis des Studienmedikaments erhalten hatten und die zusätzlich zu der Eingangsuntersuchung an mindestens einer weiteren Untersuchung teilgenommen hatten; fehlende Daten in dieser Analyse wurden gemäß LOCF (<i>last observation carried forward</i>) behandelt.</p> <p>Die Beständigkeit der HbA1c-Reduktion wurde zwischen den Behandlungsgruppen mittels Beständigkeitskoeffizienten verglichen, definiert als Anstiegsrate von Woche 24 zu Woche 52. Der Anteil der Patienten, die einen HbA1c-Wert <7 oder <6,5 erreichten, wurde zwischen den Behandlungsgruppen mit Hilfe der logistischen Regressionsanalyse verglichen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Subgruppenanalysen für den Wirksamkeitsendpunkt (d. h. Veränderung des HbA1c-Wertes zwischen der Eingangsuntersuchung und nach 52 Wochen) wurden in Subgruppen durchgeführt, die mit Ausgangs-HbA1c-Kategorien (<7, 7% bis <8, 8% bis <9, ≥9%) definiert waren.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 35)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 35)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 35)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Daten für Probandenrekrutierung wurden nicht mitgeteilt; die Studiendauer war 52 Wochen; Messungen wurden alle sechs oder acht Wochen vorgenommen
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend – die Studie endete protokollgemäß nach 52 Wochen

Item ^a	Charakteristikum	Studieninformation
a: nach CONSORT 2010.		

Anmerkungen:); ^b Randomisierungsverhältnis 1:1 für Sitagliptin 100 mg einmal täglich oder Glipizid 5-20 mg einmal täglich, beide in Kombination mit ≥ 1500 mg Metformin/Tag verabreicht;

Abkürzungen: CI, Konfidenzintervall; FPG, Nüchternplasmaglukose; HbA1c, Glykiertes Hämoglobin; OHA, oraler antihyperglykämischer Wirkstoff; RCT, Randomisierte kontrollierte Studie

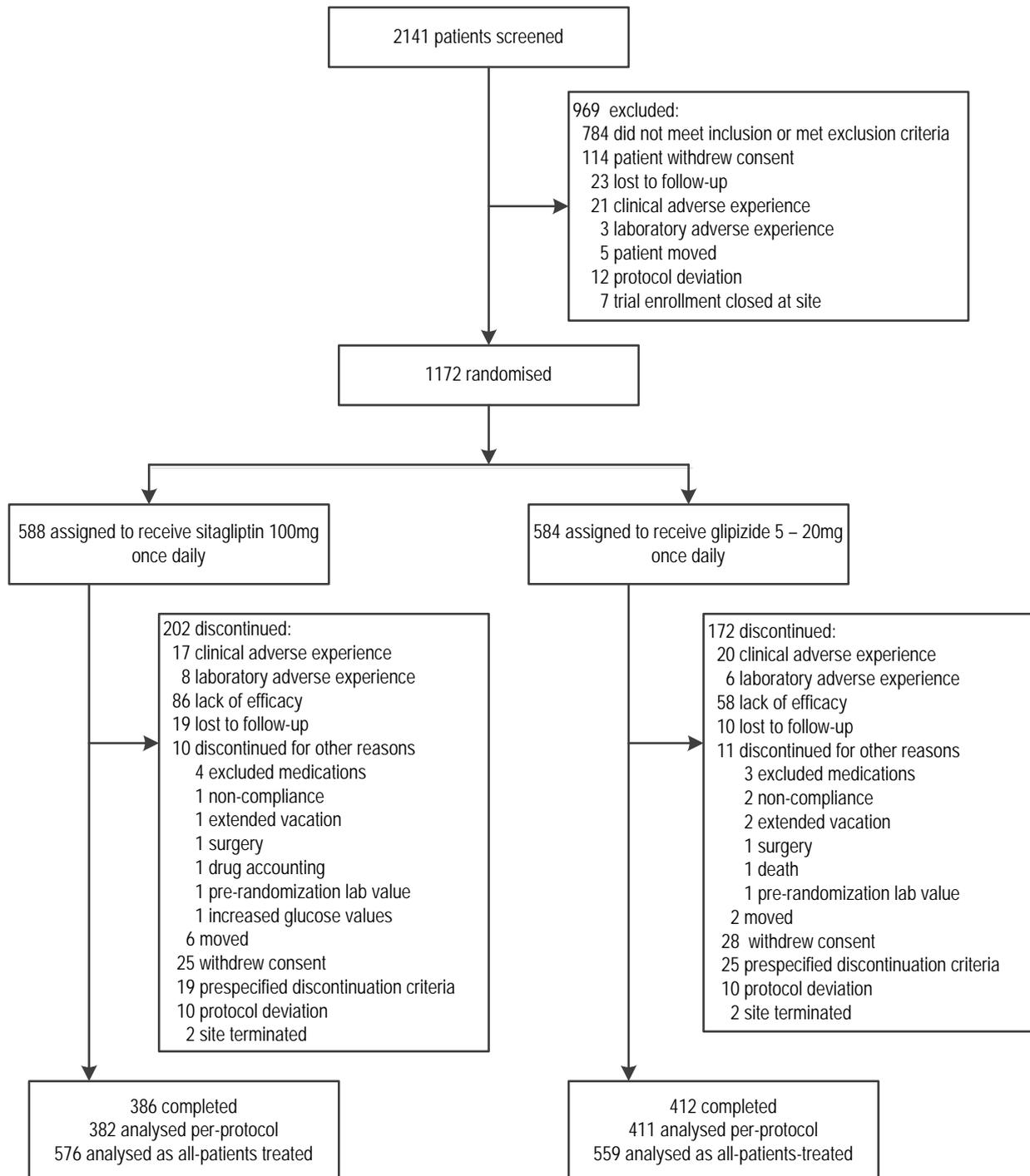


Abbildung 36: Flussdiagramm Studiendesign und -methodik für die Studie von Nauck, 2007

Tabelle 4-178 (Anhang): Studiendesign und -methodik für die Studie von Göke, 2010

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Diese 52-wöchige Studie untersuchte die Wirksamkeit und Sicherheit von Saxagliptin vs. Sulphonylurea Glipizid zusätzlich zu Metformin verabreicht bei Patienten mit Typ-2-Diabetes, die durch Metformin allein nicht ausreichend kontrolliert waren.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Die Studie war eine 52-wöchige, Phase-IIIb-, internationale, multizentrische, aktivkontrollierte, randomisierte, doppelblinde, Nicht-Unterlegenheits- (<i>non-inferiority</i>), Parallelstudie (im Verhältnis 1:1) mit einer noch laufenden 52-wöchigen Verlängerungsperiode.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	Typ-2-Diabetiker durch Metformin allein unzureichend kontrolliert
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Geeignete Studienteilnehmer waren Männer und Frauen im Alter von ≥ 18 Jahren mit Typ-2-Diabetes und HbA1c-Wert $> 6,5-10,0\%$, unter Metformin-Monotherapie mit einer stabilen Dosis ≥ 1500 mg/Tag für mindestens 8 Wochen vor Eintritt in die Studie.</p> <p>Ausschlusskriterien:</p> <p>Hauptausschlusskriterien umfassten Typ-1-Diabetes; diabetische Ketoazidose oder hyperosmolares nichtketonisches Koma; Insulin-Therapie innerhalb des vergangenen Jahres; Behandlung mit einem Thiazolidindion innerhalb von 12 Wochen vor Eintritt in die Studie; Behandlung mit systemischen Kortikoiden abgesehen von Ersatztherapie; vorherige Behandlung mit einem DPP-4-Inhibitor; Spenden von Blut, Plasma oder Blutplättchen innerhalb von 3 Monaten vor Eintritt in die Studie; kongestive Herzinsuffizienz (gemäß <i>New York Heart Association stage III/IV</i>) und/oder eine bekannte linksventrikuläre Ejektionsfraktion $< 40\%$; kardiovaskuläre Erkrankung in den vergangenen 6 Monaten definiert als Herzinfarkt, Koronarangioplastie oder Bypass, Herzklappenerkrankung oder – eingriff, instabile Angina pectoris, transitorische ischämische Attacke oder Schlaganfall; Hämoglobinopathien; signifikanter Alkohol- oder Drogenmissbrauch innerhalb eines Jahres vor</p>

Item ^a	Charakteristikum	Studieninformation
		Eintritt in die Studie; Behandlung mit <i>human immunodeficiency virus</i> - / antiviralen Medikamenten oder mit Cytochrom P450 3A4-Auslösern; Serumkreatinin $\geq 1,5$ mg/dL (≥ 133 $\mu\text{mol/L}$) für Männer oder $\geq 1,4$ mg/dL (≥ 124 $\mu\text{mol/L}$) für Frauen; aktuelle Lebererkrankung und/oder signifikante anomale Leberfunktion [Aspartat-Aminotransferase mehr als das Zweifache der Obergrenze der Normalwerte und/oder Alanin-Aminotransferase mehr als das Zweifache der Obergrenze der Normalwerte und/oder Totalbilirubin $>2,0$ mg/dL (>34 μmol)] oder eine andere klinisch relevante Anormalität beim Screening.
4b	Studienorganisation und Ort der Studiendurchführung	Nicht berichtet (internationale Studie mit mehr als einem Zentrum)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Saxagliptin 5 mg/Tag zusätzlich zur gewohnten Dosis Metformin beim Studienbeginn (1500, 2000, 2500 or 3000 mg täglich) oder Glipizid, nach Bedarf auf 5-20 mg/Tag auftitriert, zusätzlich zu dem nicht-verblindeten Metformin; Studienmedikation wurde oral unmittelbar vor oder mit einer Mahlzeit eingenommen. Glipizid wurde mit 5 mg/Tag (morgens) begonnen und in dreiwöchigen Abständen zu einem Maximum von 20 mg/Tag auftitriert. Um die Verblindung zu gewährleisten, wurde die Double-dummy-Technik verwendet. Die Titrierungsschritte waren 10 mg/Tag (morgens), gefolgt von 15 mg/Tag (10 mg morgens, 5 mg abends) und 20 mg/Tag (10 mg morgens, 10 mg abends). Erste Dosierungsbeurteilung war in Woche 3; weitere Beurteilungen erfolgten in Wochen 6, 9, 12, 15 und 18. Während der Eskalationsperiode konnte Glipizid reduziert werden, falls hypoglykämische Nebenwirkungen eintraten, und konnte danach einmal erhöht werden.
6	Zielkriterien	Veränderung des HbA1c-Wertes zwischen dem Ausgangswert und dem Wert nach 52 Wochen.
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primärer Wirksamkeitsendpunkt war Veränderung des HbA1c-Wertes zwischen dem Ausgangswert und dem Wert nach 52 Wochen und wurde zur Beurteilung herangezogen, ob Saxagliptin plus Metformin nicht unterlegen gegenüber Glipizid plus Metformin ist. Sekundäre Hauptendpunkte waren der Anteil der Patienten, die über ≥ 1 hypoglykämisches Ereignis in 52 Wochen berichteten, die Änderung des Körpergewichts in Woche 52 und die mittlere Steigung der

Item ^a	Charakteristikum	Studieninformation
		Regressionsgerade der Änderung von Woche 24 zu Woche 52 in HbA1c. Andere sekundäre Endpunkte beinhalteten die Änderung vom Eingangswert von Nüchternplasmaglukose (FPG), Nüchterninsulin, C-Peptid, Glukagon und Proinsulin; Anteil der Patienten, die therapeutisches glykämisches Ansprechen (definiert als HbA1c \leq 6,5%) erreichten; Änderung in HbA1c-Wert in Patienten mit HbA1c-Ausgangswert \geq 7,0%; Anteil der Patienten mit HbA1c-Ausgangswert \geq 7,0%, die therapeutisches glykämisches Ansprechen (definiert als HbA1c $<$ 7,0%) erreichten, und Änderung in β -Zellfunktion (bestimmt mit Homöostase-Modell). Änderung der 120-min Postprandialglukose zwischen dem Ausgangswert und nach 52 Wochen wurde in einer Subgruppe ebenfalls bewertet.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	Anzahl der randomisierten Teilnehmer: Saxagliptin+Metformin: n=428; Glipizid+Metformin: n=430
7a	Wie wurden die Fallzahlen bestimmt?	Ein Populationsgröße von 419 Patienten pro Behandlungsarm war erforderlich um mit einer statistischen <i>Power</i> von 95% den Nicht-Unterlegenheitsvergleich bei der Änderung zwischen HbA1c-Ausgangswert und dem Wert in Woche 52 auf dem 5%-Niveau zu belegen, angenommen dass die Standardabweichung der Änderung 1,1% war, mit einer auf 0,35% gesetzten Nicht-Unterlegenheitsgrenze und einem wahren Unterschied von 0 zwischen den zwei randomisierten Behandlungen. Bei der Stichprobengröße wurde davon ausgegangen, dass 35% der randomisierten Patienten von der <i>per protocol</i> -Analyse ausgeschlossen würden.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Beim Eintritt in die Studie wurde den Patienten eine eindeutige Nummer zugeordnet, die während der gesamten Studie angewandt wurde.
8a	Methode zur Generierung der zufälligen Zuteilung	Nach der Einführungsperiode wurden die in Frage kommenden Patienten mittels eines interaktiven <i>Web-response</i> -Systems randomisiert.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Patienten wurden in eine von den zwei Behandlungsgruppen nach einem ausgewogenen Blockrandomisierungsverfahren randomisiert

Item ^a	Charakteristikum	Studieninformation
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Beim Eintritt in die Studie wurde den Patienten eine eindeutige Nummer zugeordnet, die während der gesamten Studie angewandt wurde; nach der Einführungsperiode wurden die in Frage kommenden Patienten mittels eines interaktiven <i>Web-response</i> -Systems mit Hilfe eines ausgewogenen Blockrandomisierungssystems in eine von den zwei Behandlungsgruppen randomisiert
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Unklar; es ist keine ausreichende Information über die Person vorhanden, die Randomisierungszuordnung durchführte.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar. Die Autoren beschreiben die Studie als doppelblind, und geben an, dass die Patienten randomisiert waren, und mit einer doppelblinden Therapie behandelt wurden. Die Autoren liefern aber keine explizite Beschreibung.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Double-dummy-Technik wurde verwendet, um die Verblindung zu gewährleisten. Passende Placebo-Tabletten für Saxagliptin und passende Placebo-Kapseln für Glipizid wurden verwendet.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Wirksamkeitsanalysen für kontinuierliche Variablen wurden mittels ANCOVA-Modell durchgeführt, mit Behandlung als Effekt und Ausgangswert als Kovariate. Im Rahmen des ANCOVA-Modells wurden Punktschätzungen und 95% Konfidenzintervalle für die Mittelwertveränderung innerhalb jeder Behandlungsgruppe, wie auch für die Unterschiede in der Mittelwertveränderungen zwischen Saxagliptin+Metformin- und Glipizid+Metformin-Gruppe ermittelt. Saxagliptin wurde als nicht-unterlegen gegenüber Glipizid erachtet, wenn die obere Grenze des 95% Konfidenzintervalles für die Differenz der HbA1c-Wertänderung vom Ausgangswert zu Woche 52 $<0,35\%$ zwischen den zwei Behandlungsgruppen war. Ein <i>fixed-sequence</i> -Test wurde für die primären und sekundären Hauptwirksamkeitsendpunkte verwendet, um die Gesamt-Typ-1-Fehlerrate der Studie zu prüfen. Alle Vergleiche waren zweiseitig mit einem 5% Signifikanzniveau. Der Anteil der Patienten, die ≥ 1 hyperhyperglykämische Ereignis berichteten, und der Prozentsatz der Patienten, die einen therapeutischen

Item ^a	Charakteristikum	Studieninformation
		glykämischen Response erreichten, wurden mittels des exakten Tests von Fisher analysiert. Ein gemischtes Modell mit patientenspezifischen beobachteten Werten bewertete die Beständigkeit des HbA1c-Effektes. Demografische und andere Basismerkmale wurden mit deskriptiven statistischen Methoden zusammengefasst. Die Analyse der sekundären Endpunkte und die Analyse der nach HbA1c-Ausgangswert kategorisierten Subgruppen waren prädefiniert.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Eine Subgruppenanalyse wurde für den Anteil der Patienten durchgeführt, die einen HbA1c-Wert <7,0% erreichten, für Patienten mit dem Ausgangs-HbA1c-Wert <7,0%, und für Patienten, die hypoglykämische Ereignisse hatten. Korrigierte Mittelwertveränderung gegenüber dem HbA1c-Ausgangswert wurde mit den Subgruppen mit Ausgangs-HbA1c-Werten <7,0; 7,0-8,0; 8,0-9,0; und >9,0% analysiert.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> Abbildung 37)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 37) Die Wirksamkeitsanalyse des primären Endpunkts und der sekundären Hauptendpunkte, HbA1c einbeziehend, wurde mit <i>per-protocol-analysis-set</i> durchgeführt; allerdings zogen wir die <i>Intention-to-treat</i> -Zahlen aus dem klinischen Studienbericht für die Netzwerk-Meta-Analyse. Das <i>per-protocol-analysis-set</i> umfasste Patienten, die die 52-wöchige randomisierte Behandlungsperiode ohne signifikante Protokollabweichungen beendeten. Das <i>full-analysis-set</i> , das alle Patienten umfasste, die mindestens eine Dosis der randomisierten Studienmedikation erhielten und nach der Eingangsuntersuchung an mindestens einer weiteren Wirksamkeitsuntersuchung teilgenommen hatten, wurde zu der bestätigenden Analyse des primären Endpunkts und zu allen anderen Wirksamkeitsanalysen zugezogen.
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> Abbildung 37)
14	Aufnahme / Rekrutierung	Rekrutierung begann am 11. Dezember 2007 und endete am 5. August 2008
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die letzte Visite in der 52-wöchigen Behandlungsperiode war am 28. August 2009. Ein Nachbeobachtungsschema war

Item ^a	Charakteristikum	Studieninformation
		nicht spezifiziert; aber HbA1c wurde in Wochen 3, 6, 9, 12, 15, 18, 24, 30, 39 und 52 gemessen. Diese Studie hatte eine 52-wöchige Verlängerungsperiode.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend

a: nach CONSORT 2010.

Abkürzungen: ANCOVA, Analysis of covariance; HbA1c, Glykiertes Hämoglobin; DPP-4-Inhibitor, Dipeptidyl-Peptidase-4-Inhibitor; FPG, Nüchternplasmaglukose

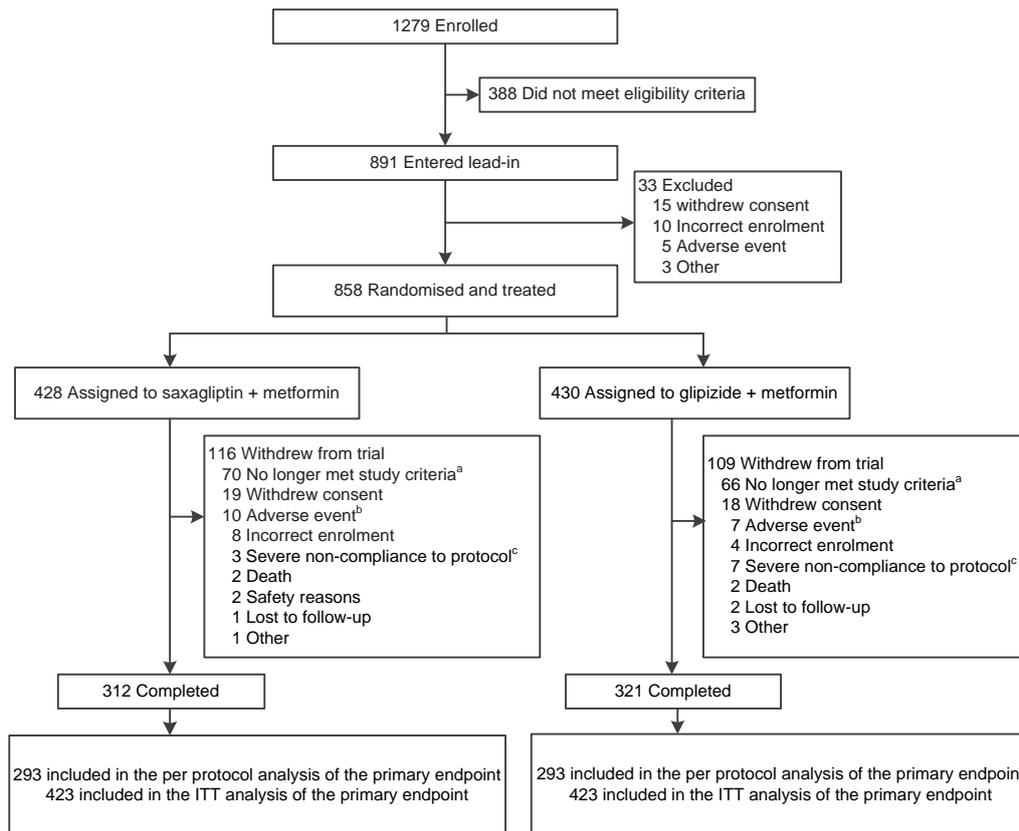


Abbildung 37: Flussdiagramm Studiendesign und -methodik für die Studie von Göke, 2010

^aDie Kategorie 'No longer met study criteria' entspricht der Kategorie 'Development of study specific discontinuation criteria'; ^bDie Patienten wurden nur in diese Kategorie eingeschlossen, wenn der Prüfarzt ein unerwünschtes Ereignis als primären Grund für einen Abbruch festgestellt hat. Ein Patient kann ein unerwünschtes Ereignis gehabt haben, dass zu einem Abbruch geführt hat, und hier nicht aufgelistet sein, wenn der Prüfarzt einen anderen primären Grund für einen Abbruch festgestellt hat; ^cNach Meinung des Prüfarztes. **ITT** = Intention to treat.

Tabelle 4-179 (Anhang): Studiendesign und -methodik für die Studie von Filozof, 2010

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Ziel dieser 52-wöchigen, randomisierten, doppelblinden, aktivkontrollierten Studie war, die Nicht-Unterlegenheit von Vildagliptin im Vergleich zu Gliclazid zu zeigen, als Zusatztherapie bei Patienten mit Typ-2-Diabetes, die durch Metformin allein nicht ausreichend kontrolliert waren.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multizentrische, randomisierte, doppelblinde, aktivkontrollierte Studie im Verhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	Typ-2-Diabetiker durch Metformin allein unzureichend therapiert
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Männer und Frauen (nicht fruchtbar oder mit einer medizinisch zugelassenen Verhütungskontrolle) im Alter von 18-78 Jahren, mit Typ-2-Diabetes und HbA1c-Wert von 7,5-11,0%, die Metformin mindestens für 3 Monate, und mit einer stabilen Dosis ≥ 1500 mg täglich mindestens für vier Wochen vor der ersten Visite erhalten hatten.</p> <p>Ausschlusskriterien: Patienten mit Typ-1-Diabetes, mit durch Bauchspeicheldrüsenverletzungen verursachtes Diabetes oder weitere Formen von Diabetes (<i>Cushing's Syndrom</i> und Akromegalie) und Patienten mit akuten metabolischen diabetischen Komplikationen (Ketoazidose oder hyperosmolarer Zustand) innerhalb der vergangenen 6 Monate; Patienten mit schwerwiegenden Herzerkrankungen (<i>torsades de pointes</i>, anhaltende und klinisch relevante ventrikuläre Tachykardie oder Kammerflimmern, percutane koronare Intervention innerhalb der letzten 3 Monate, Herzinfarkt, Bypassoperation, instabile Angina; oder Schlaganfall in den vergangenen 6 Monaten und behandlungsbedürftige kongestive Herzinsuffizienz, Grad II oder III atrioventrikulärer Block oder Verlängerung der QTc-Dauer) oder klinisch signifikante Nieren- oder Lebererkrankung; Aspartat-Aminotransferase oder Alanin-Aminotransferase mehr als das Zweifache der</p>

Item ^a	Charakteristikum	Studieninformation
		Obergrenze der Normalwerte, Totalbilirubin mehr als das Zweifache der Obergrenze der Normalwerte, positives Hepatitis-B-Oberflächenantigen und/oder Hepatitis-C-Antikörper, Serumkreatinin $\geq 132 \mu\text{mol/L}$ bei Männern und $\geq 123 \mu\text{mol/L}$ bei Frauen, oder anomale Kreatininclearance, klinisch relevante Schilddrüsenhormonwerte außerhalb des Normalbereichs oder Nüchterntriglyceride $>7,9 \text{ mmol/L}$ beim Screening.
4b	Studienorganisation und Ort der Studiendurchführung	Internationale Studie mit mehr als einem Studienzentrum
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Ausgewählte Patienten wurden bei Studienbeginn 1:1 zu Vildagliptin (50 mg zweimal täglich) oder Gliclazid (80 mg/Tag) zusätzlich zu einer stabilen Dosis Metformin (1500 mg täglich) randomisiert. Gliclazid-Dosis wurde von einer Anfangsdosis von 80 mg/Tag auf eine Maximumdosis von 320 mg/Tag auftitriert, wenn Nüchternplasmaglukose (FPG) $>7.0 \text{ mmol/L}$ war oder wenn Nüchternblutglukose $>6.3 \text{ mmol/L}$ war (basierend auf eine Nüchtern- <i>finger-stick capillary glucose</i> -Messung, die in dem Studienzentrum durchgeführt wurde). Die Patienten bekamen die nächsthöhere Dosis in Woche 4 (160 mg), dann in Woche 8 (240 mg) und in Woche 12 (320 mg).
6	Zielkriterien	Primärer Wirksamkeitsendpunkt war Veränderung des HbA1c-Wertes zwischen dem Ausgangswert und nach 52 Wochen
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primärer Wirksamkeitsendpunkt war Veränderung des HbA1c-Wertes zwischen dem Ausgangswert und nach 52 Wochen. Sekundäre Wirksamkeitsbeurteilungen beinhalteten den Anteil der Patienten, die einen HbA1c-Wert $<7\%$ und $\leq 6,5\%$ erreichten, Änderungen in FPG, Körpergewicht, β -Zellfunktion und Insulinresistenz.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	Anzahl der randomisierten Teilnehmer: Vildagliptin 50 mg zweimal täglich: n=513; Gliclazid bis auf 320 mg/Tag: n=494
7a	Wie wurden die Fallzahlen bestimmt?	800 Patienten (400 pro Gruppe) waren erforderlich, um die Nicht-Unterlegenheit von Vildagliptin gegenüber Gliclazid für die Reduzierung des HbA1c-Wertes zu zeigen, mit einem einseitigen Niveau von 0,025 am Ende der Studie mit 92% statistischer Power (ein wahrer Unterschied von 0,1% zugunsten von Gliclazid, Standardabweichung von 1,25

Item ^a	Charakteristikum	Studieninformation
		für HbA1c-Reduktion in Woche 52 und eine Studienabbruchrate von 20% über die Periode von 52 Wochen angenommen).
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Nicht berichtet
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar. Die Autoren beschreiben die Studie als doppelblind, und geben an, daß alle randomisierten Patienten verblindet waren, wobei Double-dummy-Design verwendet wurde. Die Autoren liefern aber keine explizite Beschreibung.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Verblindung aller randomisierten Patienten erfolgte mittels Double-dummy-Design.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Nicht-Unterlegenheit von Vildagliptin 50 mg zweimal täglich gegenüber Gliclazid bis auf 320 mg täglich, beide zusammen mit Metformin, wurde mittels Konfidenzintervall (95% CI)-Ansatz [Nicht-Unterlegenheitsmarge <0,4% von der Kovarianzanalyse (ANCOVA)], ermittelt. Die letzte verfügbare Beobachtung nach der Basismessung (<i>last observation carried forward</i>) wurde für die Analyse verwendet. Die Analysen von primären und sekundären Wirksamkeitsvariablen basierten auf der <i>per protocol</i> -Population. Die Sensitivitätsanalyse basierte auf die <i>Intention-to-treat</i> -Population und wurde durchgeführt, um die Robustheit der Schlußfolgerung zu bewerten. Die primären und sekundären Wirksamkeitsvariablen wurden mittels ANCOVA-Modell durchgeführt, mit Therapie und gepoolte Studienzentren (<i>pooled centre</i>) als Variablen und Ausgangswert als Kovariate. Für FPG war die Nicht-Unterlegenheitsmarge <0,6.

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Die Subgruppenanalysen für Veränderung des HbA1c-Wertes vom Ausgangswert zum Endpunkt wurden entsprechend dem Ausgangswert von HbA1c (HbA1c ≤8,0% und >8,0%; ≤9,0% und >9,0%), dem Alter (≤65 Jahre und >65 Jahre), dem Geschlecht und Basiswert von BMI (<30,0, ≥30,0 und ≥35,0 kg/m ²) durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 38)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 38) Die Analysen von primären und sekundären Wirksamkeitsvariablen basierten auf der <i>per protocol</i> -Population. Die Sensitivitätsanalyse basierte auf die <i>Intention-to-treat</i> -Population und wurde durchgeführt, um die Robustheit der Schlußfolgerung zu bewerten. <i>Intention-to-treat</i> -Population: Randomisierte Patienten, die mindestens eine Dosis des Studienmedikaments erhalten hatten, und sowohl an der Eingangsuntersuchung als auch an mindestens einer weiteren Untersuchung teilgenommen hatten. Ergebnisse wurden nicht berichtet, aber sie wurden als vergleichbar mit den Ergebnissen der <i>per protocol</i> -Population berichtet. <i>Per protocol</i> -Population: Umfasste Patienten in der <i>Intention-to-treat</i> -Population mit mehr als 24-wöchiger Behandlung, mit keinen gravierenden Protokollverletzungen, und Patienten, die sich der letzten gültigen Beurteilung der primären Wirksamkeitsvariable HbA1c innerhalb von 7 Tagen nach der letzten Dosis des Studienmedikaments unterzogen und entweder (i) mehr als 48 Wochen der Behandlung vollendeten oder (ii) <48 Wochen Behandlung hatten, aber die Studie wegen des unbefriedigenden Ansprechens abbrachen.
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 38)
14	Aufnahme / Rekrutierung	Rekrutierungsdaten wurden nicht berichtet
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studienbeginn wurde nicht berichtet. Patientenbeobachtung umfasste neun geplante Besuche und einen Telefonanruf über die 52 Wochen der Behandlung; Telefonmonitoring der verdächtigen Symptome von potenzieller Lebertoxizität (z. B. Anorexie, Übelkeit, Erbrechen,

Item ^a	Charakteristikum	Studieninformation
		Müdigkeit, Oberbauchbeschwerden rechts, Gelbsucht); einen ungeplanten Besuch, um vorherige Bewertung im Falle von verdächtigen Symptomen innerhalb von 3 Tagen zu verifizieren.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht berichtet

a: nach CONSORT 2010.

Abkürzungen: ANCOVA, Analysis of covariance; HbA1c, Glykiertes Hämoglobin; FPG, Nüchternplasmaglukose; CI Konfidenzintervall

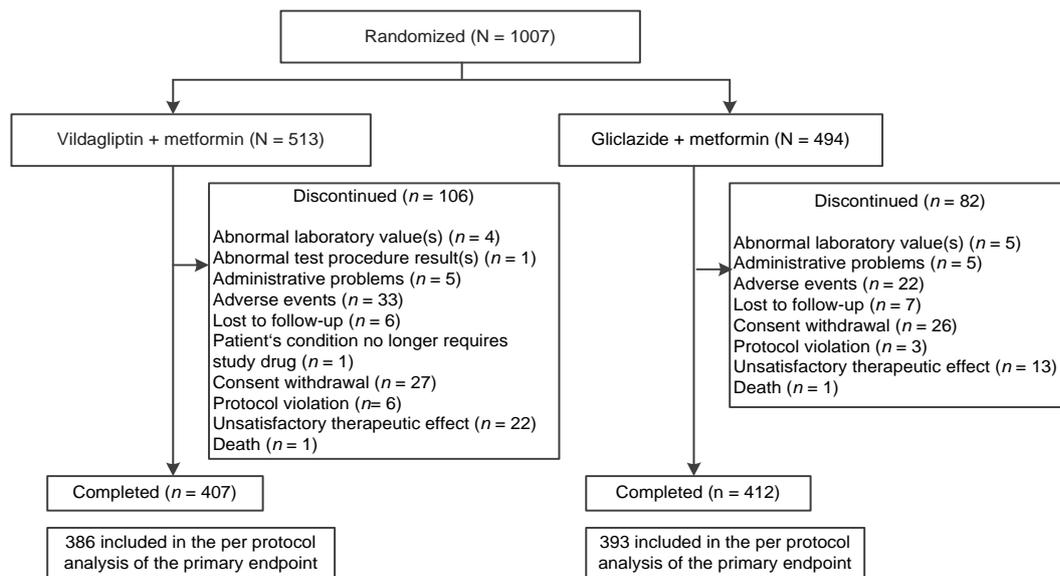


Abbildung 38: Flussdiagramm Studiendesign und -methodik für die Studie von Filozof, 2010

Tabelle 4-180 (Anhang): Studiendesign und -methodik für die Studie von Derosa 2011b

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Studienziel war es, die Wirksamkeit von Exenatide im Vergleich mit Glimperid im Hinblick auf glykämische Kontrolle und Parameter, die durch die Insulinresistenz verursacht werden, bei Typ-2-Diabetikern, die mit Metformin behandelt werden, zu untersuchen.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Parallelzuteilte (im Verhältnis 1:1), aktivkontrollierte Studie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	Typ-2-Diabetiker, die mit Metformin unzureichend kontrolliert sind
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Nicht berichtet
4b	Studienorganisation und Ort der Studiendurchführung	Italien; exakte Details sind nicht berichtet
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Exenatide 5 µg zweimal täglich und nach einem Monat eine zu 10 µg erhöhte Dosis zweimal täglich (n=57); Glimperid 1 mg dreimal täglich und nach einem Monat eine erhöhte Dosis von 2 mg dreimal täglich (n=52), beides zusammen mit Metformin (1000–2000 mg/Tag)
6	Zielkriterien	BMI, HbA1c, Nüchternblutglukose, postprandialer Blutzucker, Nüchternplasmainsulin, HOMA Index und Sammelplasmaanalysen Adiponektin und Tumornekrosefaktor- α als Ausgangswert und nach 12 Monaten
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Nicht berichtet
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	Die randomisierte Fallzahl wurde nicht für die einzelnen Studienarme berichtet. Jedoch betrug die Anzahl derer, die die Studie vollständig beendeten n=57 für den Exenatide-behandelten Studienarm und n=54 für den Glimperid-behandelten Studienarm.
7a	Wie wurden die Fallzahlen bestimmt?	Nicht berichtet

Item ^a	Charakteristikum	Studieninformation
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht berichtet
8	Randomisierung, Erzeugung der Behandlungsfolge	Nicht berichtet
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	Nicht berichtet
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Nicht berichtet
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	Nicht berichtet
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Nicht berichtet
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 39)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 39)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Nicht berichtet
14	Aufnahme / Rekrutierung	Nicht berichtet
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Nicht berichtet

Item ^a	Charakteristikum	Studieninformation
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht berichtet
a: nach CONSORT 2010.		

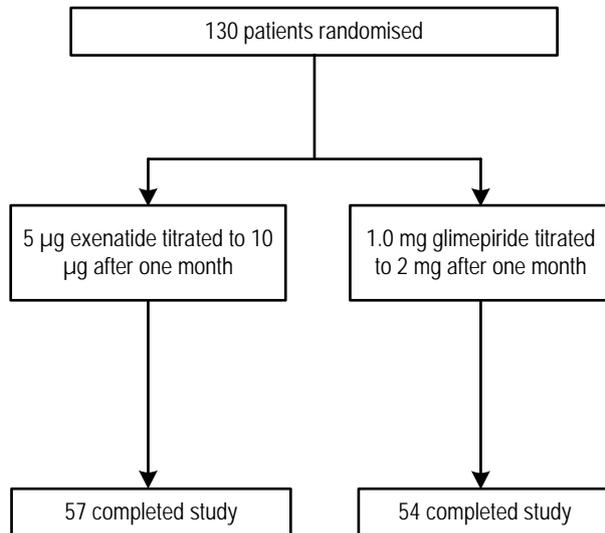


Abbildung 39: Flussdiagramm Studiendesign und -methodik für die Studie von Derosa 2011b

Anmerkung: Der Abstract berichtete keine Anzahl für die Registrierung und für die Randomisierung oder Gründe für den Abbruch der Studie

Tabelle 4-181 (Anhang): Studiendesign und -methodik für die Studie von Derosa 2010

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Ziel war es, die Wirkungen einer einjährigen Behandlung mit Exenatide im Vergleich zu Glibenclamid, zusätzlich zu Metformin verabreicht, bei Patienten mit Typ-2-Diabetes auf Körpergewicht, Blutzuckerkontrolle und Beta-Zellfunktion, sowie auf Insulin-Resistenz und Entzündungsparameter wie Resistin, Retinol bindendes Protein-4 und hochempfindliches C-reaktives Protein zu bewerten.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie (RCT)
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Es handelte sich um eine multizentrische, randomisierte, einfachblinde Studie, mit einem Zuteilungsverhältnis von 1:1.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Es wurden keine Änderungen vorgenommen.
4	Probanden / Patienten	Typ-2-Diabetiker mit unzureichender Blutzuckerkontrolle und Unverträglichkeit gegenüber Metformin in der Höchstdosis. Passende Patienten, identifiziert durch Überprüfen von Fallaufzeichnungen und/oder computerunterstützten klinischen Registern, wurden von den Prüfern persönlich oder telefonisch kontaktiert.
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Weißer Männer und Frauen, über 18 Jahre alt, mit Typ-2-Diabetes, unzureichender Blutzuckerkontrolle (HbA1c-Wert >8.0%) und Übergewicht (BMI ≥ 25 und < 30 kg/m²), unter Metformin-Behandlung mit einer Durchschnittsdosierung von 1500+/-500 mg/Tag. Patienten wiesen eine Unverträglichkeit gegenüber Metformin in Höchstdosierung (3000 mg/Tag) auf, mit Auftreten von Magen-Darm-Beschwerden, wie Diarrhoe und Meteorismus, wenn Metformin auf die Maximalkonzentration titriert wurde.</p> <p>Ausschlusskriterien:</p> <p>Diabetische Ketoazidose oder instabile oder schnell voranschreitende diabetische Retinopathie, Nephropathie oder Neuropathie, Beeinträchtigung der Leberfunktion (der Plasma-Aminotransferase- und/oder der γ-Glutamyltransferase-Spiegel liegt über der für das jeweilige Alter und Geschlecht normalen Obergrenze.), Beeinträchtigung der</p>

Item ^a	Charakteristikum	Studieninformation
		Nierenfunktion (Serum-Kreatinin-Spiegel liegt über der für das jeweilige Alter und Geschlecht normalen Obergrenze), oder schwere Anämie. Patienten mit ernstesten kardiovaskulären Erkrankungen (z. B. <i>New York Heart Association</i> Herzinsuffizienz der Stadien I–IV oder anamnestischer Herzinfarkt oder Schlaganfall) oder mit zerebrovaskuläre Ereignissen während der 6 Monate vor Studienbeginn wurden ebenfalls ausgeschlossen. Schwangere oder stillende Frauen oder Frauen, die im gebärfähigen Alter waren und keine adäquaten empfängnisverhütenden Maßnahmen vornahmen, wurden ebenfalls ausgeschlossen.
4b	Studienorganisation und Ort der Studiendurchführung	8 Zentren in Italien (Abteilung für Innere Medizin und Therapeutika, Universität Pavia, Pavia, Italien; „G. Descovich“ Astheosklerose Studienzentrum, Abteilung für Innere Medizin, Alters- und Nierenkrankheiten, Universität Bologna, Bologna, Italien; Geschäftseinheit Diabetes Care, S. Carlo Krankenhaus, Milano, Italien; Pesenti Fenaroli Krankenhaus, Alzano Lombardo, Bergamo, Italien; Unitá Metabolica, Regionales Krankenhaus, Varese, Italien. Abteilung für Medizin, städtisches Krankenhaus, Cittiglio, Varese, Italien; RSA Don Leone Porta, Milano, Italien; und Fondazione Ospedale della Carita`, Casalbuttano, Cremona, Italien.)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Die Patienten wurden angewiesen, Exenatide 5 µg zweimal am Tag (n=63) oder Glibenclamid 2.5 mg dreimal am Tag (n=65) und nach einem Monat Exenatide titriert auf 10 µg zweimal am Tag oder Glibenclamid titriert auf 5 mg dreimal am Tag für 12 Monate, beides zusätzlich zu Metformin 1000-2000 mg/Tag, zu nehmen.
6	Zielkriterien	Bewertung der Wirkungen einer 1-jährigen Behandlung mit Exenatide verglichen mit Glibenclamid bei Typ-2-Diabetes-Patienten auf Körpergewicht, Blutzuckerkontrolle und Beta-Zellfunktion
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Änderungen des Körpergewichtes, der Werte der Blutzuckerkontrolle und der Beta-Zellfunktion zwischen Ausgangswert und Messung in Woche 52 Sekundäre Zielkriterien: Änderungen der Insulin-Resistenz und von Entzündungsparametern, wie Resistin, Retinol bindendes Protein-4 und hochempfindliches C-reaktives Protein zwischen Ausgangswert und Messung in Woche 52

Item ^a	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	Randomisiert wurden: Exenatide (5 µg, zweimal am Tag, nach einem Monat titriert auf 10 µg, zweimal am Tag): n=63; Glibenclamid (2.5 mg dreimal am Tag, nach einem Monat titriert auf 5 mg, dreimal am Tag): n= 65.
7a	Wie wurden die Fallzahlen bestimmt?	Nicht berichtet
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Zur Randomisierung wurden Umschläge gezogen, die Randomisierungscodes enthielten, die von einem Statistiker vorbereitet worden waren.
8a	Methode zur Generierung der zufälligen Zuteilung	Zur Randomisierung wurden Umschläge gezogen, die Randomisierungscodes enthielten, die von einem Statistiker vorbereitet worden waren.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Zur Randomisierung wurden Umschläge gezogen, die Randomisierungscodes enthielten, die von einem Statistiker vorbereitet worden waren.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Zum Großteil nicht berichtet. Ein Statistiker bereitete Randomisierungscodes vor.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar. Die Autoren bezeichneten diese Studie als einfachblind aber berichteten nicht ausdrücklich, wer oder auf welche Weise verblindet wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Durchgängige Variablen wurden mit einer Varianzanalyse verglichen. Auswirkungen der Intervention wurden an hinzukommende, potentielle Störfaktoren angepasst mit Hilfe einer Kovarianzanalyse. Eine Varianzanalyse wurde ebenfalls verwendet, um die Signifikanz innerhalb und zwischen den Gruppen zu bemessen. Die statistische Signifikanz der unabhängigen

Item ^a	Charakteristikum	Studieninformation
		Behandlungswirkungen auf die anderen Variablen wurde mit einer Kovarianzanalyse bestimmt. Mit Hilfe eines t-Tests bei einer Stichprobe wurden die Werte verglichen, die vor und nach der Verabreichung der Behandlung erhoben wurden; t-Tests bei zwei Stichproben wurden für die Vergleiche zwischen den Gruppen verwendet. Die Bonferroni-Korrektur für mehrfache Vergleiche wurde ebenfalls durchgeführt. Die statistische Datenanalyse wurde mit der <i>Statistical Package for Social Sciences software</i> Version 11.0 (SPSS Inc., Chicago, IL) durchgeführt. Die Daten wurden als Mittelwert±Standardabweichungen dargestellt. Für alle statistischen Analysen wurde $p < 0.05$ als statistisch signifikant festgelegt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 40)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 40) Jeder Patient, der mindestens eine Dosis der Studienmedikation erhielt, wurde einer Beobachtung bezüglich der Verträglichkeit unterzogen, um das Auftreten akuter unerwünschter Ereignisse auszuschließen. Danach wurde eine ITT-Analyse bei Patienten durchgeführt, die eine oder mehrere Dosen der Studienmedikation erhielten, keine akuten unerwünschten Reaktionen zeigten und eine darauffolgende Wirksamkeitsmessung hatten.
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 40)
14	Aufnahme / Rekrutierung	Daten bezüglich der Patientenrekrutierung wurden nicht berichtet.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studiendaten wurden nicht spezifiziert. Alle Sicherheits- und Wirksamkeitsparameter wurden gemessen zu Beginn der Studie und nach 3, 6, 9 und 12 Monaten. Um die Verträglichkeitsmessungen zu bewerten, wurden alle unerwünschten Ereignisse berichtet.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
^a :nach CONSORT 2010.		

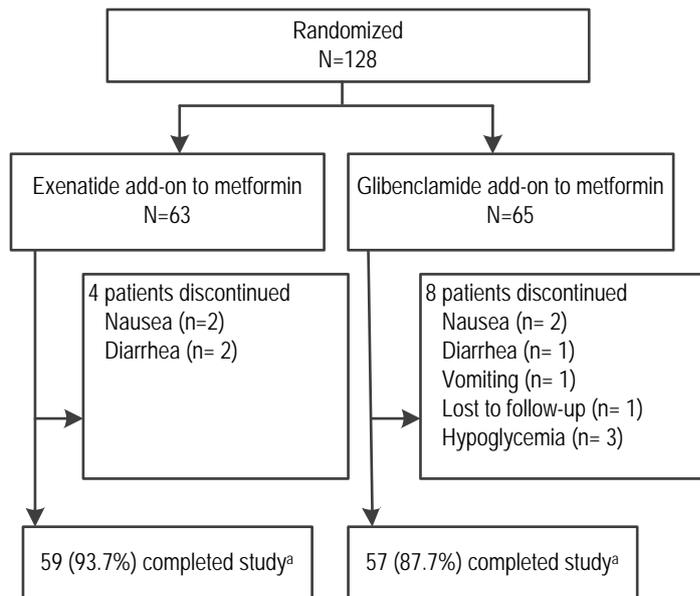


Abbildung 40: Flussdiagramm Studiendesign und -methodik für die Studie von Derosa, 2010

^aDie Studienprüfer berichteten, dass eine veränderte ITT-Analyse durchgeführt wurde, aber die verwendeten Angaben in die Analyse und die Methode, um fehlende klinische Messungen aufzufüllen ist unklar. Die Studienprüfer gaben keine Patientenanzahl für die Rekrutierung und das Screening an.

Tabelle 4-182 (Anhang): Studiendesign und -methodik für die Studie von Matthews, 2010

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Randomisierte, doppelblinde, aktivkontrollierte Studie bei Patienten mit Typ-2-Diabetes, die durch Metformin allein nicht ausreichend kontrolliert waren (HbA1c 6,5–8,5%)
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte kontrollierte Studie; dies war eine 2-jährige, multizentrische, randomisierte, doppelblinde, double-dummy, aktivkontrollierte Studie; Patienten waren im Verhältnis 1:1 randomisiert.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Um die Beständigkeit der Behandlung zu bewerten, war ursprünglich eine 5-jährige Studie geplant, die als primären Endpunkt das Therapieversagen der glykämischen Kontrolle (definiert als HbA1c >8%) hatte. Wegen der unerwartet hohen Studienabbruchrate und weil weniger Patienten als erwartet den Endpunkt HbA1c >8% erreichten, wurde die Studie später modifiziert. Die <i>Power</i> wurde mit Hilfe von der verblindeten Stichprobenpopulation neu berechnet und wurde für unzureichend befunden. Das Studienziel wurde deshalb zu einer 2-jährigen Nicht-Unterlegenheitsstudie und der primäre Endpunkt zu 'Änderung des HbA1c-Wertes im Vergleich zwischen Ausgangswert und dem Wert in Woche 104' geändert. Die Studie endete, als der letzte randomisierte Patient in der Studie zwei Jahre gewesen war.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien (Ferrannini, 2009): Männer und Frauen (nicht fruchtbar oder mit einer medizinisch zugelassenen Verhütungskontrolle) mit Typ-2-Diabetes und HbA1c-Wert von 6,5-8,5%, die Metformin für ≥ 3 Monate erhalten hatten und die eine stabile Dosis ≥ 1500 mg täglich für mindestens ≥ 4 Wochen vor der ersten Visite hatten, im Alter von 18-73 Jahren waren und einen BMI von 22-45 kg/m² hatten, waren zur Teilnahme berechtigt.</p> <p>Ausschlusskriterien (Ferrannini, 2009): Patienten mit einer Vorgeschichte von Typ-1-Diabetes oder einem anderen Diabetes wurden ausgeschlossen, wie auch Patienten, die akute metabolische diabetische Komplikationen in den letzten 6 Monaten gehabt hatten, ferner Patienten, die akute</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Infektionen, die die Blutglukosekontrolle beeinflussen könnten, in den 4 Wochen vor der ersten Visite gehabt hatten, auch Patienten mit folgenden Erkrankungen: schwere Herzerkrankungen (Vorgeschichte von <i>torsades de pointes</i> oder ventrikuläre Tachykardie; perkutane koronare Intervention in den letzten 3 Monaten; Herzinfarkt, koronare Bypassoperation, instabile Angina oder Schlaganfall in den letzten 6 Monaten; kongestive Herzinsuffizienz, die pharmakologische Behandlung erfordert; Grad II oder III atrioventrikulärer Block oder verlängerte QTc) oder klinisch signifikante Leber- oder Nierenerkrankung. Eine oder mehrere Abweichungen von den folgenden Laborwerten bei der Eingangsuntersuchung schlossen die Teilnahme ebenfalls aus: Alanin-Aminotransferase oder Aspartat-Aminotransferase mehr als das Dreifache der Obergrenze der Normalwerte (<i>upper limit of normal, ULN</i>), direktes Bilirubin mehr als das 1,3-Fache ULN, Serumkreatinin ≥ 132 mmol/L bei Männern oder ≥ 123 mmol/L bei Frauen, klinisch relevanter Wert des schilddrüsenstimulierenden Hormons außerhalb des Normalbereichs bei der Eingangsuntersuchung; oder Nüchterntriglyceride $>7,9$ mmol/L.</p>
4b	Studienorganisation und Ort der Studiendurchführung	International mit 402 Zentren
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Die geeigneten Patienten wurde 1:1 am Tag 0 zu Vildagliptin (50 mg zweimal täglich) oder Glimepirid (Anfangsdosis 2 mg/Tag) zusätzlich zu Metformin (Dosis blieb unverändert) randomisiert. Glimepirid konnte maximal auf 6 mg/Tag in Wochen 4, 8 oder während einer späteren Visite, falls der FPG-Wert 6,2 mmol/L überschritt, auftitriert werden, oder in Fällen von wiederholter Hypoglykämie, ausstitriert werden.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäres Zielkriterium: Änderung des HbA1c-Wertes zwischen dem Ausgangswert und dem Endpunkt (Woche 104). Die Methode des Übertragens der letzten Beobachtung (<i>last observation carried forward</i>) wurde für die Patienten verwendet, die die Studie vorzeitig abbrachen. Für Patienten, die Notfallmedikation erhielten, wurde als Endpunkt für die Woche 104 die Messung von der letzten Messung vor dem Notfall definiert.</p> <p>Sekundäre Zielkriterien: Sekundäre Wirksamkeitsbewertungen</p>

Item ^a	Charakteristikum	Studieninformation
		<p>beinhalteten HbA1c-Ansprechraten (HbA1c-Ausgangswert $\geq 7\%$ und $< 7\%$ in Woche 104 oder HbA1c-Ausgangswert $> 6.5\%$ und $\leq 6.5\%$ in Woche 104), prädefinierte Subgruppenanalysen nach Alter (≥ 65 oder < 65 Jahre) bei der Eingangsuntersuchung, und <i>coefficient of failure</i> (CoF; Steigung der Regressionsgerade von HbA1c vs. Zeit, nach dem Prinzip der kleinsten Quadrate) (Wallace, 2002), wie auch FPG, Körpergewicht, Nüchternlipide, β-Zellfunktion und Insulinresistenzparameter.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Um die Beständigkeit der Behandlung zu bewerten, war ursprünglich eine 5-jährige Studie geplant, die als primären Endpunkt das Therapieversagen der glykämischen Kontrolle (definiert als HbA1c $> 8\%$) hatte. Wegen der unerwartet hohen Studienabbruchrate und weil weniger Patienten als erwartet den Endpunkt HbA1c $> 8\%$ erreichten, wurde die Studie später modifiziert. Die <i>Power</i> wurde mit Hilfe von der verblindeten Stichprobenpopulation neu berechnet und wurde für unzureichend befunden. Das Studienziel wurde deshalb zu einer 2-jährigen Nicht-Unterlegenheitsstudie und der primäre Endpunkt zu 'Änderung des HbA1c-Wertes im Vergleich zwischen Ausgangswert und dem Wert in Woche 104' geändert.</p>
7	Fallzahl	<p>Randomisierungszahlen (52 Wochen, Ferrannini, 2009): Vildagliptin 50 mg zweimal täglich: n=1396; Glimperid bis zu 6 mg/Tag: n=1393</p>
7a	Wie wurden die Fallzahlen bestimmt?	<p>Mit 3120 randomisierten Patienten (d. h. 1560 Patienten pro Behandlungsarm) und einer Abbruchrate von 20%, hatte die Studie 96% <i>Power</i>, um die Nicht-Unterlegenheit von Vildagliptin im Vergleich zu Glimperid zu zeigen (einseitiger α-Level von 0,0125, vorausgesetzt eine Nicht-Unterlegenheitsmarge von 0,3% HbA1c und eine Standardabweichung von 1,25%).</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Eine geplante Zwischenauswertung nach 52 Wochen, die Vildagliptin (50 mg zweimal täglich) mit Glimperid (bis zu 6 mg/Tag) bei Typ-2-Diabetes-Patienten verglich, die mit Metformin-Monotherapie (Ferrannini, 2009) nicht ausreichend kontrolliert waren, wurde abgeschlossen. Ca. 6000 Patienten wurden für die Randomisierung von 3000 Patienten gescreent; die Zwischenauswertung wurde durchgeführt, als ca. 2800 Patienten die 52-wöchige Behandlung beendet hatten (oder beendet haben würden).</p> <p>Die primäre Wirksamkeitsbewertung war die Änderung des HbA1c-Ausgangswertes. Die sekundären Wirksamkeitsbewertungen beinhalteten HbA1c-Responderraten und</p>

Item ^a	Charakteristikum	Studieninformation
		HbA1c-Reduktion kategorisiert nach HbA1c-Ausgangswert, Altersgruppen, FPG und Körpergewicht.
8	Randomisierung, Erzeugung der Behandlungsfolge	Nicht berichtet
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar, die Autoren beschreiben die Studie als doppelblind, aber beschreiben nicht explizit wer verblindet war und wie die Verblindung durchgeführt wurde; dennoch, die Autoren berichteten dass eine unabhängige kardiovaskuläre und zerebrovaskuläre (CCV) Beurteilungskomitee alle Vorkommnisse von CCV-Ereignissen verblindet überprüften (Ferrannini, 2009).
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Verwendung einer Kontrollgruppe wird in der Zwischenauswertung nach 52 Wochen erwähnt.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Änderungen von Ausgangswerten in primären und sekundären Endpunkten wurden mittels ANCOVA-Modell analysiert, mit Behandlung und gepoolten Studienzentren als Klassifikationsvariablen, und HbA1-Ausgangswert als Kovariate. Der Test für die Nicht-Unterlegenheit von Vildagliptin im Vergleich zu Glimperid wurde mit einer Konfidenzintervall-Annäherung durchgeführt. Die Nicht-Unterlegenheit für HbA1c war gezeigt, wenn die Obergrenze des 97,5% Konfidenzintervalls (CI) für den Behandlungsunterschied in der berichtigten Mittelwertsveränderungen vom Ausgangswert zum Endpunkt in dem ANCOVA-Modell weniger als 0,3% HbA1c war, ein Unterschied, der nicht als klinisch signifikant angesehen wird. Eine Zwischenauswertung wurde in Woche 52 durchgeführt und die Nicht-Unterlegenheits-Hypothese wurde mit einseitigem 0,0125-Niveau getestet. Um die einseitige Typ-1-Fehlerrate auf einem Niveau von 0,025 zu

Item ^a	Charakteristikum	Studieninformation
		halten, wurde der Nicht-Unterlegenheitstest mit einem nominalen einseitigen α -Level von 0,0125 mittels Bonferroni-Korrektur durchgeführt. Die primäre Analyse basierte auf der PP-Population, und die Hauptwirksamkeitsanalysen wurden auch mit der ITT-Population durchgeführt. Der Anteil der Responder im jeweiligen Behandlungsarm wurde zusammengefasst und mittels <i>chi-squared</i> -Test verglichen. Die Mittelwertdifferenz der Behandlungen in der Nachhaltigkeit wurde mit zweiseitigen t-Tests in der Gesamtgruppe und in den Subgruppen verglichen. Multiple Schrittweise Regressionanalyse wurde verwendet, um unabhängige Einflussfaktoren für die Nachhaltigkeit zu suchen. Alle Parameter mit einem univariaten $p < 0,10$, wurden in ein multivariates Verfahren eingeschlossen.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Die Subgruppenanalysen wurden getrennt nach BMI, Alter, Dauer der Krankheit und HbA1c-Ausgangswert durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 41)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 41)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Siehe <i>Flow-Chart</i> (Abbildung 41)</p> <p><i>Intent-to-treat</i> (ITT)-Population: Alle Patienten, die mindestens eine Dosis des Studienmedikaments erhalten hatten, und die mindestens eine Wirksamkeitsmessung (primäre oder sekundäre) nach der Eingangsuntersuchung hatten. Notfallmedikation wurde nicht in die Bewertung eingeschlossen.</p> <p><i>Per-protocol</i> (PP)-Population umfasste Patienten zumindest aus einer der folgenden Kategorien (52 Wochen; Ferrannini, 2009): (i) beendeten mindestens 48 Wochen von der Behandlung ohne Notfallmedikation und ohne schwerwiegende Protokollverletzung; (ii) begannen Notfallmedikation wegen fehlender Wirksamkeit nach 24 Wochen der Behandlung (wie <i>per protocol</i>) ohne schwerwiegende Protokollverletzung; (iii) Abbruch der Studie wegen fehlender Wirksamkeit (wie <i>per protocol</i>) und ohne schwerwiegende Protokollverletzung.</p>

Item ^a	Charakteristikum	Studieninformation
		<i>Per protocol</i> (PP)-Population (104 Wochen): ITT Patienten, die keine schwerwiegenden Protokollverletzungen und eine valide Abschlußmessung von HbA1c hatten, und nicht auf Notfalltherapie waren und (i) ≥ 96 Wochen der Behandlung beendeten oder (ii) mindestens 24 Wochen der Behandlung ohne Notfallmedikation beendeten und eine Notfallmedikation nach 24 Wochen begannen oder (iii) wegen des unzufriedenstellenden therapeutischen Ansprechens die Studie abbrachen.
14	Aufnahme / Rekrutierung	Die Rekrutierungsdaten wurden nicht berichtet
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Patienten besuchten eine Screeningsvisite (Woche -4, 1. Visite), bei der Einschluss- und Ausschlusskriterien festgesetzt wurden. Geeignete Patienten wurden 1:1 am Tag 0 (Ausgangswert) randomisiert. Weitere Visiten wurden für die Wochen 4, 8, 12, 16, 20, 24, 32, 40, 46 und 52 geplant. Eine Zwischenauswertung nach 52 Wochen (Ferrannini, 2009) wurde für die 104-wöchige Studie abgeschlossen (Matthews, 2010). Für die 52-wöchige Studie umfasste die randomisierte Population die ersten 2800 Patienten, die die 52 Wochen in der Studie mit vollständig bereinigten und gesperrten Wirksamkeits- und Sicherheitsdaten beendeten (oder beendet haben würden).
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

Abkürzungen: ANCOVA, Analysis of covariance- Kovarianzanalyse; BMI, *Body-Mass-Index*; CCV, Kardiovaskulär und zerebrovaskulär; CI, Konfidenzintervall; FPG, Nüchternplasmaglukose; HbA1c, Glykiertes Hämoglobin; ITT, *Intent-to-treat*; PP, *Per protocol*; T2DM, Typ-2-Diabetes mellitus; ULN, Obergrenze des Normalwertes.

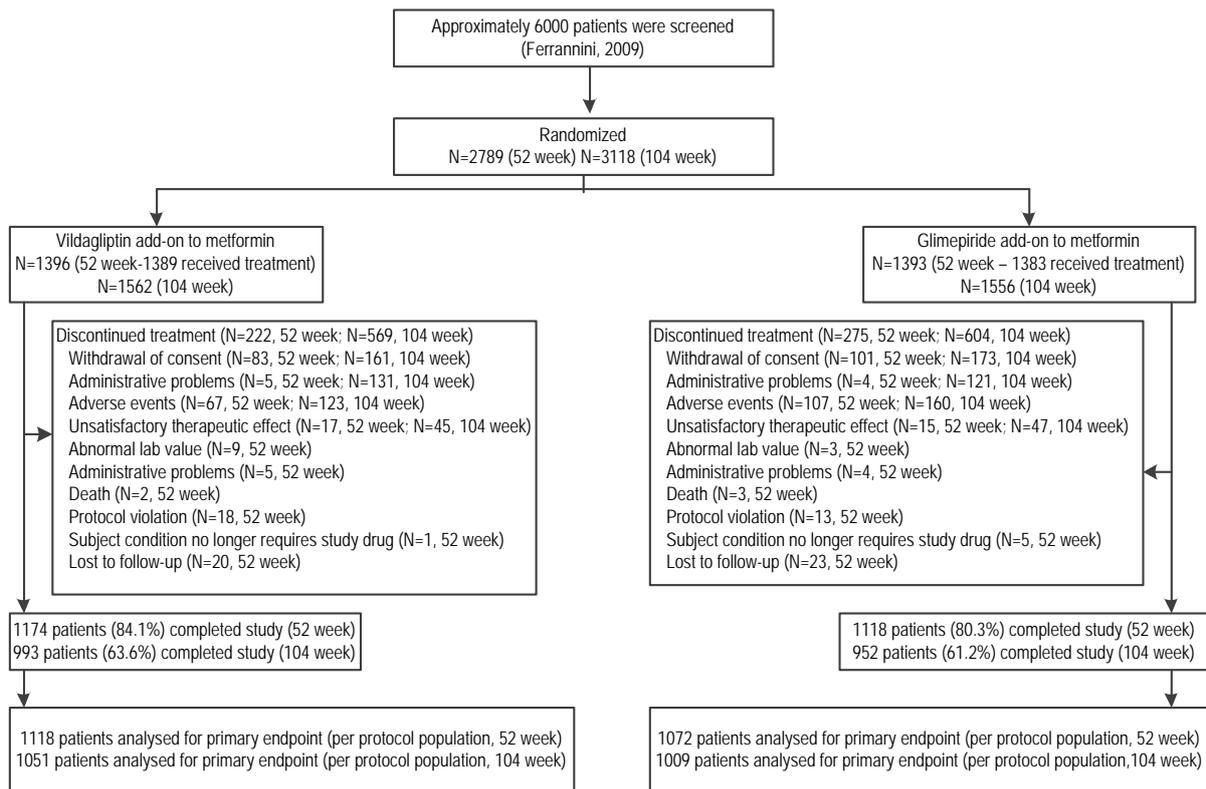


Abbildung 41: Flussdiagramm Studiendesign und -methodik für die Studie von Matthews, 2010 und Ferrannini, 2009

Tabelle 4-183 (Anhang): Studiendesign und -methodik für die Studie von Yang 2011

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	In dieser Studie sollte die Wirksamkeit und Sicherheit von 5 mg Saxagliptin einmal täglich <i>add on</i> zu Metformin versus Placebo <i>add on</i> zu Metformin bei asiatischen Patienten mit Typ-2-Diabetes mellitus (T2DM), deren Blutzucker inadäquat mit Metformin-Monotherapie kontrolliert ist, untersucht werden.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Es handelt sich um eine parallele, Placebo-kontrollierte, randomisierte Studie. Zuteilungsverhältnis: 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet.
4	Probanden / Patienten	Typ-2-Diabetes Patienten, deren Blutzucker inadäquat durch Metformin-Monotherapie kontrolliert ist.
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> - Alter ≥ 18 Jahre - Typ-2-Diabetes Patient - Glykiertes Hämoglobin (HbA1c) 7–10%, trotz Metformin-Monotherapie (≥ 1500 mg/Tag für ≥ 8 Wochen) - C-Peptid-Konzentration ≥ 0.33 nmol/L <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> - Typ-1-Diabetes - Krankheitsgeschichte einer diabetischen Ketoazidose, hyperosmolare nicht-ketotische Koma oder Symptome einer schlecht kontrollierten Diabeteserkrankung - <i>New York Heart Association</i> (NYHA)-Klassifizierung III oder IV - kongestive Herzinsuffizienz und/oder eine linke ventrikuläre Ejektionsfraktion von ≤ 40 - eine signifikante kardiovaskuläre Erkrankung während der 6 Monate der Rekrutierungs-Phase - aktuelle Lebererkrankung und/oder signifikante abnormale Leberfunktion - Krankheitsgeschichte einer Hämoglobinopathie - instabile oder schnell fortschreitende Nierenerkrankung - autoimmune Hauterkrankung - Patient hatte eine gastrointestinale Chirurgie, welche die Resorption des

Item ^a	Charakteristikum	Studieninformation
		<p>Medikaments beeinflussen könnte</p> <ul style="list-style-type: none"> - Alkohol- oder Drogenmissbrauch in den letzten 12 Monaten - Immunsupprimierte Patienten (Humanes Immundefizienz-Virus [HIV] oder Transplantationspatienten) - schwangere oder stillende Frauen - Patienten mit einer signifikanten Auffälligkeit, die während der körperlichen Untersuchung, des Elektrokardiogramms oder der Laboruntersuchung festgestellt worden ist, welche nach der Einschätzung des untersuchenden Arztes die Sicherheit des Patienten oder die erfolgreiche Beendigung der Therapie gefährden würde. - nicht spezifizierte Grenzen bestimmter Laborparameter (Serumkreatinin, Aspartat Aminotransferase, Alanin Aminotransferase, Kreatinkinase) - Insulin Therapie innerhalb des letzten Jahres vor Rekrutierung (Ausnahme: während eine Krankenhausaufenthaltes oder bei Schwangerschaftsdiabetes) - Behandlung mit DPP-4 Inhibitoren - Behandlung mit einem antidiabetischen Medikament mit der Ausnahme von ≥ 1500 mg/Tag Metformin in den letzten 8 Wochen vor Rekrutierung (12 Wochen bei vorhergehender Behandlung mit Thiazolidindionen) - Verwendung von systemischen Glucocorticoiden oder Cytochrom P450 3A4 Induzierern
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie in 40 Zentren (21 Zentren in China, 7 in Indien und 12 in Südkorea)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	5 mg Saxagliptin täglich plus Metformin (n=283) oder Placebo plus Metformin (n=287). Die Gabe von Metformin fand für die Dauer der Studie unverblindet statt (1500, 2000, 2500, oder 3000 mg/Tag [2500 mg/Tag = maximale Dosis in China]).
6	Zielkriterien	Veränderung des HbA1c vom Ausgangswert bis Woche 24
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Endpunkt: Veränderung des HbA1c vom Ausgangswert bis Woche 24</p> <p>Sekundäre key-Endpunkte:</p> <ul style="list-style-type: none"> - Nüchtern-Blutzucker (FPG)-Veränderung vom Ausgangswert bis Woche 24 - postprandialer Blutzucker (PPG) während eines Toleranztests auf ein gemischtes Essen (425 kcal, 85 g Glukose) gemessen an der Fläche unter der Kurve von 0 bis 180 min

Item ^a	Charakteristikum	Studieninformation
		<p>(AUC_{0–180}) in einer Subgruppe an Patienten</p> <ul style="list-style-type: none"> - Anteil an Patienten, die ein therapeutisch glykämisches Ansprechen zeigen (definiert mit HbA_{1c} <7%) <p>Zusätzliche sekundäre Wirksamkeitsendpunkte, welche als maßgeblich für den Aufschluss erachtet worden sind:</p> <ul style="list-style-type: none"> - 2 h PPG während eines gemischten Essens bei einem Anteil der Patienten - Nüchtern und postprandialer AUC₀₋₁₈₀ Insulin - C-Peptid - Glukagon - Proinsulin - die β-Zellfunktion wurde mit dem <i>Homeostasis Modell Assessment</i> (HOMA-2β), basierend auf der Messung des C-Peptids und des FPG (Wallace, 2004) kalkuliert - <i>Body Mass Index</i> (BMI) - Taillenumfang - Körpergewicht - prozentuale Veränderung des Nüchternlipids (Gesamtcholesterol, Low Density Lipoprotein Cholesterol, High Density Lipoprotein Cholesterol und Triglycerid) vom Ausgangswert <p>Sicherheit und Verträglichkeit:</p> <ul style="list-style-type: none"> - unerwünschte Ereignisse - häufige unerwünschte Ereignisse - schwerwiegende unerwünschte Ereignisse - therapieabhangige unerwünschte Ereignisse - Therapieabbruch aufgrund von unerwünschten Ereignissen
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet.
7	Fallzahl	Randomisierte Patienten: Saxagliptin+Metformin: n=283 Placebo+Metformin: n=287
7a	Wie wurden die Fallzahlen bestimmt?	Mit der Gesamtanzahl an randomisierten und behandelten Patienten von 530 (n=265 pro Behandlungsgruppe), ergab sich eine <i>Power</i> von 98% um einen Unterschied von 0.5% zwischen den 2 randomisierten Behandlungsgruppen zu detektieren (absolute Veränderung vom Ausgangswert bis Woche 24 des HbA _{1c} bei einem Level von 5%, geschätzte Standardabweichung vom Ausgangswert bis Woche 24 des HbA _{1c} von 1,2%).
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen	Nicht anwendbar

Item ^a	Charakteristikum	Studieninformation
	Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	Computererzeugte Generierung
8a	Methode zur Generierung der zufälligen Zuteilung	Computererzeugte Generierung
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgte zum Ausgangszeitpunkt, wurde zwischen den Behandlungsgruppen ausbalanciert und stratifiziert nach Land auf Grundlage einer Computergenerierten Sequenz.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Metformin wurde <i>open-label</i> verabreicht. Die Gabe von Saxagliptin verlief verblindet, mittels einer doppelblinden, single-dummy Methode.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Der Versuchsleiter rekrutierte die Patienten der Studie.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar. Der Autor berichtet von einer doppelblinden, single-dummy Methode zur Verblindung der Saxagliptin-Gabe, während Metformin <i>open-label</i> verabreicht wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die primäre Wirksamkeitsanalyse wurde mit dem Full-Analysis-Set (≥ 1 Dosis der Studienmedikamentation, sowie Ausgangswert- und Folgewert-Wirksamkeitsdaten) mit <i>der last observations carried forward</i> (LOCF)-Methode, durchgeführt. Die Behandlungsgruppen wurden mittels der Kovarianz Analyse (ANCOVA) miteinander verglichen, bei der die Behandlung und das Land als feste Effekte und der HbA1c-Ausgangswert als Kovariante verwendet wurden. Um die generellen Typ-1- und Typ-2-Fehler innerhalb der Studie zu kontrollieren, wurde ein hierarchischer Test für die primären Endpunkte und die drei sekundären <i>key</i> Endpunkte verwendet. Die Veränderung des FPG vom Ausgangswert wurde mittels <i>Full-Analysis-Set</i> ausgewertet. Die Behandlungsgruppen wurden mittels ANCOVA miteinander verglichen. Die Veränderung des AUC0-180 vom Ausgangswert, bei dem Anteil der Patienten die ein gemischtes Essen zu sich nahmen,

Item ^a	Charakteristikum	Studieninformation
		<p>wurde mittels ANCOVA verglichen.</p> <p>Der Anteil an Patienten mit einem glykämischen Ansprechen (HbA1c <7%) vom Ausgangswert wurde mittels dem Full-Analysis-Set ausgewertet. Die Behandlungsgruppen wurden mittels Fishers exakten Tests miteinander verglichen.</p> <p>Alle Vergleiche waren zweiseitig bei einem Signifikanzniveau von 5%.</p> <p>Die explorativen Wirksamkeitsendpunkte wurden ähnlich analysiert, mit der Ausnahme dass kein formaler Signifikanztest durchgeführt wurde.</p> <p>Sicherheitsdaten wurden deskriptiv aufgelistet.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 42)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 42)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 42)
14	Aufnahme / Rekrutierung	Nicht berichtet
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	<p>Die Studie bestand aus 3 Phasen: Rekrutierungs-Phase, <i>Lead-In</i>-Phase, Behandlungs-Phase (nach Randomisierung).</p> <p>Es fanden 9 Visiten während der Behandlungs-Phase statt (Woche 0 [Ausgangswert], 2, 4, 6, 8, 12, 16, 20 und 24.</p>
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete regulär
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

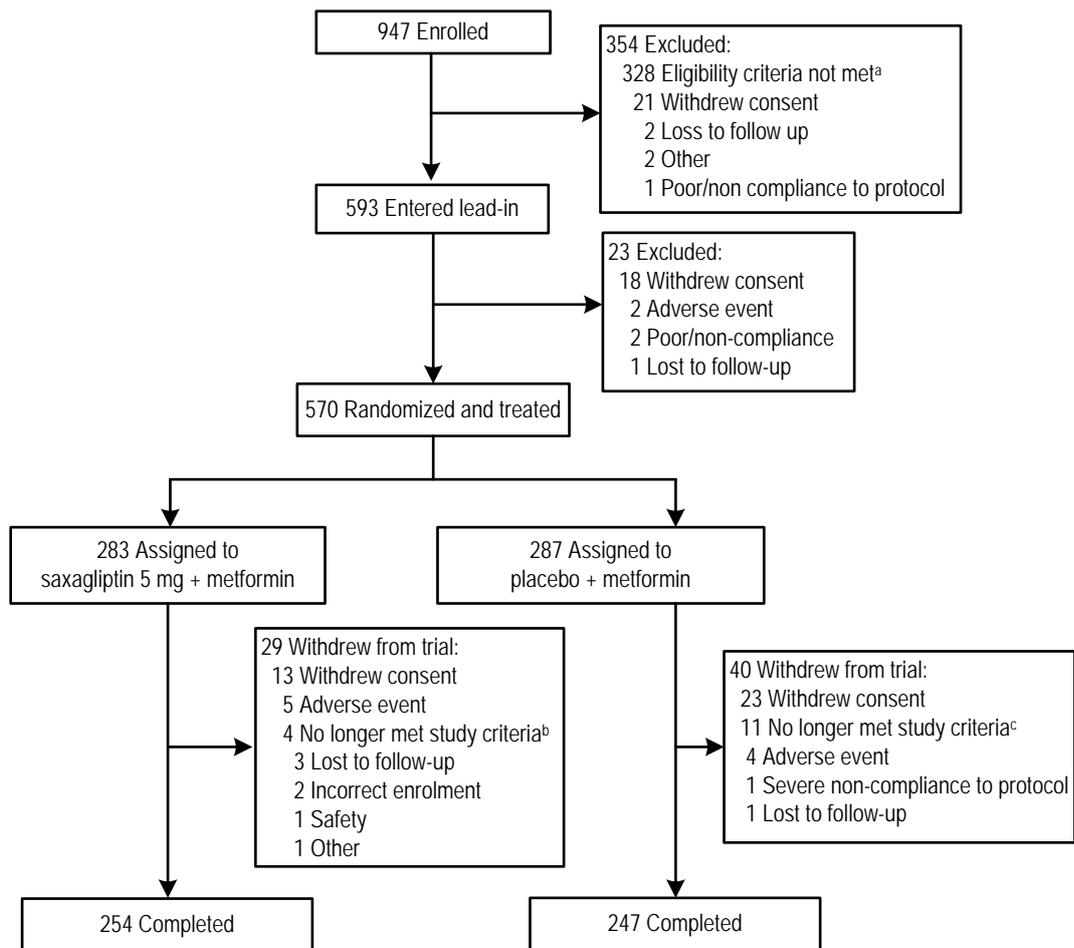


Abbildung 42: Flussdiagramm Studiendesign und -methodik für die Studie von Yang 2011

Tabelle 4-184 (Anhang): Studiendesign und -methodik für die Studie von Derosa 2011a

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Ziel der Studie war es, die Wirkung von Pioglitazon und Glibenclamid auf das Lipidprofil und auf Entzündungsparameter während einer oralen Fettbelastung zu untersuchen
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Parallele (1:1), aktivkontrollierte, randomisierte Studie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	Typ-2 Diabetiker, die mit Metformin unzureichend kontrolliert sind
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien: Typ-2-Diabetiker im Alter von mindestens 18 Jahren mit unzureichender glykämischer Kontrolle, gekennzeichnet durch einen Anteil von glykiertem Hämoglobin (HbA1c) von >7,0% unter einer Therapie mit Diät, körperlicher Aktivität und Metformin (mittlere Dosierung 1700 ±850 mg/Tag)</p> <p>Ausschlusskriterien: Frühere Ketoazidosen oder instabile oder schnell fortschreitende diabetische Retinopathie, Nephropathie (definiert als Beginn einer Albuminausscheidung >300 mg/24 h oder eine Albuminausscheidungsrate >200 µg/min während eines Zeitraums von 6 Monaten) oder Neuropathie; Beeinträchtigte Leberfunktion (definiert als Aminotransferase- und/oder Gamma-Glutamyltransferase-Plasmaspiegel oberhalb des dem Alter und Geschlecht entsprechenden Normbereichs); Beeinträchtigte Nierenfunktion (definiert als Serumkreatininspiegel oberhalb des dem Alter und Geschlecht entsprechenden Normbereichs); Schwere Anämie (definiert als Hämoglobinspiegel unter 8 g/dl); Patienten mit schwerwiegenden kardiovaskulären Erkrankungen, Herzversagen oder früherem Herzversagen (Stadium I-IV gemäß der <i>New York Heart Association</i>) sowie Patienten mit früherem Herzinfarkt, Schlaganfall oder zerebrovaskulären Erkrankungen (ischämischer Schlaganfall, Hirnblutung oder</p>

Item ^a	Charakteristikum	Studieninformation
		transitorische ischämische Attacke) innerhalb von 6 Monaten vor Aufnahme in die Studie. Frauen nach der Menopause mit früherer Osteoporose wurden aufgrund des unter Pioglitazon beobachteten höheren Risikos für Frakturen von Unterarm, Hand oder Handgelenk sowie Fuß, Knöchel, Schienbein oder Wadenbein ausgeschlossen. Ebenfalls ausgeschlossen wurden schwangere oder stillende Frauen sowie gebärfähige Frauen ohne adäquate Empfängnisverhütung.
4b	Studienorganisation und Ort der Studiendurchführung	Italien, zwei Zentren
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Pioglitazon, titriert auf bis zu 45 mg/Tag bzw. Glibenclamid, titriert auf bis zu 15 mg/Tag und zusätzlich Metformin, dessen Dosis alle 3 Monate hochtitriert wurde (unabhängig von der glykämischen Kontrolle bis zum Erreichen der maximalen verträglichen Dosis) über eine Dauer von 12 Monaten.
6	Zielkriterien	Metabolische und Entzündungsparameter nach einer oralen Fettbelastung.
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Der primäre Endpunkt dieser Studie ist die Untersuchung der Wirkung von Pioglitazon bzw. Glibenclamid zusätzlich zur Behandlung mit Metformin auf metabolische und Entzündungsparameter nach einer oralen Fettbelastung Der sekundäre Endpunkt dieser Studie ist die Untersuchung der Wirkung von Pioglitazon bzw. Glibenclamid zusätzlich zur Behandlung mit Metformin auf metabolische und Entzündungsparameter nach einer Behandlungsdauer von 12 Monaten.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	Pioglitazon: 102 Patienten Glibenclamid: 99 Patienten
7a	Wie wurden die Fallzahlen bestimmt?	Ein Unterschied von mindestens 10% gegenüber dem Ausgangswert und ein Alpha-Fehler von 0,05 wurden als klinisch signifikant betrachtet, die tatsächliche Patientenzahl war adäquat, um eine <i>Power</i> von 0,80 für alle gemessenen Variablen zu erreichen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Computergenerierte Randomisierungs-codes, die von einem Statistiker erstellt wurden
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung wurde durchgeführt, indem ein Umschlag gezogen wurde, der die von einem Statistiker erstellten

Item ^a	Charakteristikum	Studieninformation
		Randomisierungscodes enthielt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Patienten wurden entsprechend der prädefinierten Randomisierungscodes einer Studienmedikation zugeteilt.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Randomisierung wurde durchgeführt, indem ein Umschlag gezogen wurde, der die von einem Statistiker erstellten Randomisierungscodes enthielt. Nur die Person, die für die Durchführung der statistischen Analyse verantwortlich war, erhielt eine Kopie des Codes. Der Code wurde erst offengelegt, nachdem die Datenbank geschlossen war, konnte jedoch in Notfällen für einzelne Patienten offengelegt werden.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Patienten und Prüfarzte waren verblindet. Pioglitazon und Glibenclamid wurden als identisch aussehende, undurchsichtige, weiße Kapseln in codierten Flaschen verabreicht, um die Verblindung der Studie zu gewährleisten.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Pioglitazon und Glibenclamid wurden als identisch aussehende, undurchsichtige, weiße Kapseln verabreicht.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Bei Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten und bei denen danach eine Untersuchung zur Wirksamkeit durchgeführt wurde, wurde eine <i>Intention-to-treat</i> -Analyse durchgeführt. Patienten wurden in die Untersuchung der Verträglichkeit aufgenommen, wenn sie mindestens eine Dosis der Studienmedikation erhalten hatten und danach auf Verträglichkeit hin untersucht wurden. Kontinuierlichen Variablen wurden getestet, indem eine Zwei-faktorielle Varianzanalyse (ANOVA) mit wiederholten Messungen verwendet wurde, um die Gesamtunterschiede der postprandialen Reaktion zu beurteilen. Der Zuwachs der Fläche unter der Kurve wurde berechnet als die gegenüber dem Ausgangswert erhöhte Reaktion abzüglich jeglichen Abfalls unter den Ausgangswert, basiert auf der Trapezoidregel. Die Behandlungseffekte wurden mittels Kovarianzanalyse (ANCOVA) hinsichtlich

Item ^a	Charakteristikum	Studieninformation
		<p>zusätzlicher möglicher Störvariablen (Geschlecht, Raucherstatus) adjustiert.</p> <p>Eine ANOVA wurde auch verwendet, um die Signifikanz innerhalb der Gruppen und zwischen den Gruppen zu beurteilen.</p> <p>Die statistische Signifikanz der unabhängigen Behandlungseffekte auf die anderen Variablen wurde mittels ANCOVA bestimmt, indem jeweils der Ausgangswert des Parameters als Kovariate verwendet wurde.</p> <p>Es wurden auch paarweise Tests verwendet:</p> <p>Ein Ein-Stichproben-t-Test, um die Werte vor und nach der Behandlung zu vergleichen, sowie ein Zwei-Stichproben-t-Tests, um die Differenz (Endwert minus Ausgangswert) für einen gegebenen Parameter zu vergleichen.</p> <p>Unterschiede über die Zeit und die Assoziierung zwischen Blutglukosespiegel und anderen Variablen wurden mittels schrittweiser mehrfacher linearer Regressionsanalyse evaluiert.</p> <p>Die Statistische Analyse der Daten wurde mit Hilfe der <i>Statistical Package for Social Sciences Software</i>, Version 14.0 durchgeführt (SPSS Inc., Chicago, Illinois, USA).</p> <p>Für alle statistischen Analysen wurde $p < 0,05$ als statistisch signifikant angesehen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 43)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 43)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 43)
14	Aufnahme / Rekrutierung	Geeignete Patienten, die bei der Durchsicht von Patientenunterlagen und/oder elektronischer klinischer Register identifiziert wurden, wurden von den Prüfarzten persönlich oder telefonisch kontaktiert. Alle Patienten gaben eine schriftliche Einwilligungserklärung für die Teilnahme ab.
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Patienten wurden nach Randomisierung 12 Monate beobachtet und nach 0 (Ausgangswert), 3, 6, 9 und 12 Monaten untersucht.

Item ^a	Charakteristikum	Studieninformation
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend – die Studie endete wie geplant.
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

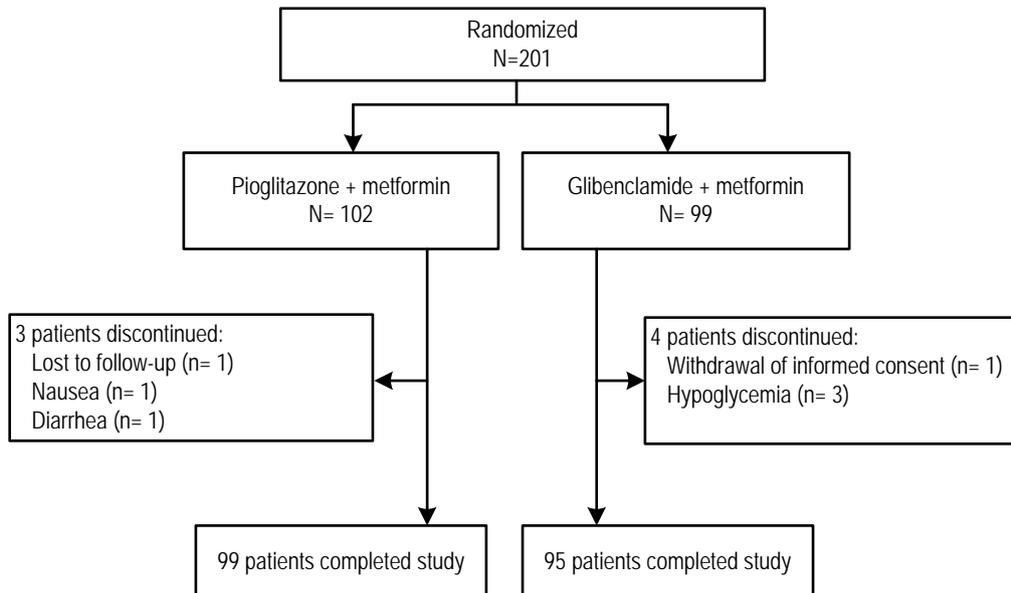


Abbildung 43: Flussdiagramm Studiendesign und -methodik für die Studie von Derosa 2011a

Tabelle 4-185 (Anhang): Studiendesign und -methodik für die Studie von Pan 2012

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Ziel der Studie war es, die Wirksamkeit und Sicherheit von Vildagliptin <i>add on</i> zu Metformin, bei chinesischen Patienten mit Typ-2-Diabetes mellitus (T2DM), die durch Metformin allein unzureichend kontrolliert waren, zu untersuchen.
-	Methoden	
3	Studiendesign	Randomisierte, kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Bei der Studie handelt es sich um eine randomisierte, doppelblinde, parallelgruppengeführte (im Verhältnis 1:1) ^b , placebokontrollierte Sicherheits- und Wirksamkeitsstudie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	Mit Metformin allein nicht ausreichend kontrollierte Typ-2-Diabetiker
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <p>Erwachsene Patienten im Alter von 18-78 Jahren mit Typ-2-Diabetes mit HbA1c-Wert von 7–10%, obwohl sie mindestens für 8 Wochen mit Metformin und mindestens für 4 Wochen mit einer stabilen Dosis Metformin (≥ 1500 mg/Tag) vor der Eingangsuntersuchung behandelt wurden.</p> <p>Zusätzliche Einschlusskriterien:</p> <p>Patienten (Männer, nicht fruchtbare Frauen oder Frauen, die medizinisch anerkannte Verhütungsmethoden anwandten) mit einem BMI von 20-40 kg/m² und einem Nüchternblutzucker (FPG)-Wert <270 mg/dL (15 mmol/L)</p> <p>Ausschlusskriterien:</p> <p>Patienten mit Typ-1-Diabetes, mit durch Bauchspeicheldrüsenverletzungen verursachtem Diabetes oder sekundären Formen von Diabetes und Patienten mit akuten metabolischen diabetischen Komplikationen (Ketoazidose oder hyperosmolarer Zustand) innerhalb der vergangenen 6 Monate; Patienten mit Myokardinfarkt, instabiler Angina oder Bypassoperation an Herzkranzgefäßen, in den vergangenen 6 Monaten.</p> <p>Patienten mit (kongestiver) Herzinsuffizienz, Lebererkrankungen, wie Zirrhose oder</p>

Item ^a	Charakteristikum	Studieninformation
		chronischer aktiver Hepatitis oder mit einem der folgenden anomalen Laborbefunde bei der Eintrittsuntersuchung: Alanin-Aminotransferase (ALT) oder Aspartat-Aminotransferase (AST) mehr als das Zweifache der Obergrenze der Normalwerte (ULN), Gesamt-Bilirubin mehr als das Zweifache der Obergrenze der Normalwerte, Serumkreatinin [Männer: ≥ 1.5 mg/dL (132 $\mu\text{mol/L}$); Frauen: ≥ 1.4 mg/dL (123 $\mu\text{mol/L}$)] oder Thyroidea-stimulierendes-Hormon außerhalb des normalen Rahmens, Nüchtern-Triglyceride >500 mg/dL (5.64 mmol/L)
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie an 20 Orten in China
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Vildagliptin 50 mg zweimal täglich mit Metformin (n=146) verglichen mit Placebo und Metformin (n=144) über 24 Wochen. Metformin blieb durchgehend auf einer stabilen Dosis von ≥ 1500 mg täglich.
6	Zielkriterien	Veränderung des HbA1c-Ausgangswertes in 24 Wochen
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Ziel: Veränderung des HbA1c-Ausgangswertes bis Woche 24 Sekundäres Ziel: Veränderung des FPG-Ausgangswertes bis Woche 24 Außerdem wurden unerwünschte Ereignisse untersucht und aufgezeichnet.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	Randomisierungszahl: Vildagliptin 50 mg zweimal täglich + Metformin: n=146; Placebo+Metformin: n=144
7a	Wie wurden die Fallzahlen bestimmt?	Es war eine Stichprobengröße von 369 randomisierten Patienten erforderlich, um mit einer statistischen <i>Power</i> von mindestens 90% einen statistisch signifikanten Unterschied im HbA1c-Wert (0.5%) nachzuweisen und um die Überlegenheit von Vildagliptin über Placebo durch Senkung des HbA1c-Werts vom Ausgangswert nach 24 Behandlungswochen, ein zweiseitiges Signifikanz-Niveau von 0.05 angenommen, zu testen. Unter Berücksichtigung einer

Item ^a	Charakteristikum	Studieninformation
		12%-igen Abbruchrate, wurden 420 Patienten randomisiert.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht zutreffend
8	Randomisierung, Erzeugung der Behandlungsfolge	Nicht berichtet
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar. Die Autoren bezeichneten diese Studie als "doppelblind" aber berichteten nicht ausdrücklich, wer verblindet wurde und wie die Verblindung vorgenommen wurde.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Für die Wirksamkeitsanalyse wurde die <i>Full-Analysis-Set</i>-Population verwendet und die Robustheit der Ergebnisse wurde im <i>Per-Protokoll-Set</i> untersucht.</p> <p>Die Sicherheitspopulation bestand aus allen Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten.</p> <p>Für die primäre Analyse wurde, für Patienten, die frühzeitig abbrachen, die <i>Last-Observation-Carried-Forward</i>-Methode (LOCF) angewandt; es ist unklar ob diese auch für die sekundären Wirksamkeitsvariablen angewandt wurde.</p> <p>Die primären und sekundären Wirksamkeitsvariablen wurden anhand der Kovarianzanalyse analysiert mit Behandlung und Region (<i>pooled-centre</i>) als Klassifikationsfaktoren und der Ausgangswertvariablen als Kovariable</p> <p>Die Analysen wurden anhand zweiseitigen</p>

Item ^a	Charakteristikum	Studieninformation
		Tests und einem statistischen Signifikanzniveau von 0,05 durchgeführt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Präspezifizierte Subgruppenanalysen über die Veränderung des HbA1c vom Ausgangswert bis zum Endpunkt wurden durchgeführt in Übereinstimmung zum Ausgangswert HbA1c ($\leq 8\%$ und $> 8\%$; $\leq 9\%$ und $> 9\%$), Alter (< 65 Jahre und ≥ 65 Jahre), Geschlecht Ausgangswert BMI ($< 30 \text{ kg/m}^2$ und $\geq 30 \text{ kg/m}^2$)</p> <p>Während der Metformin-Therapie und während der Diabetes-Erkrankung wurden Post-hoc-Subgruppenanalysen durchgeführt, um die Placeboresponse zu untersuchen. Der Prozentsatz der Responder hinsichtlich der jeweiligen Responsekriterien wurde zwischen Behandlungsgruppen, unter Verwendung von <i>Chi-Squared</i>-Tests (Responsekriterien: Endpunkt HbA1c $< 7\%$, oder Endpunkt HbA1c $\leq 6.5\%$, oder HbA1c Veränderung von Ausgangswert am Endpunkt $\geq 1\%$, oder HbA1c Veränderung von Ausgangswert am Endpunkt $\geq 0.7\%$ oder $\geq 0.5\%$)</p>
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 44)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 44)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 44)
14	Aufnahme / Rekrutierung	Nicht berichtet
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Studie bestand aus zwei Perioden: Screenings- und randomisierte Behandlungsperiode. Es gab 7 Visiten während der randomisierten Behandlungsperiode, in Woche 0 (Ausgangswert), 4, 8, 12, 16, 20 und 24.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend – Studie endete wie geplant
a: nach CONSORT 2010.		

Anmerkungen: ^a, Nach CONSORT 2010 checklist (Moher, 2010); ^b, Randomisierung erfolgte zu zwei Dosen Vildagliptin (50 mg zweimal täglich und 50 mg täglich) und Placebo d. h. in einem Verhältnis 1:1:1, aber nur Vildagliptin 50 mg (zweimal täglich)-Arm und Placebo-Arm wurden für die Nutzenbewertung einbezogen

Abkürzungen: ALT, Alanin-Aminotransferase; AST, Aspartat-Aminotransferase; BMI, *Body-Mass-Index*; FPG,

,Nüchternplasmablutzucker; HbA1c, Glykiertes Hämoglobin; LOCF, Last observation carried forward; T2DM, Typ-2-Diabetes mellitus; ULN, Obergrenze des Normalwertes

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

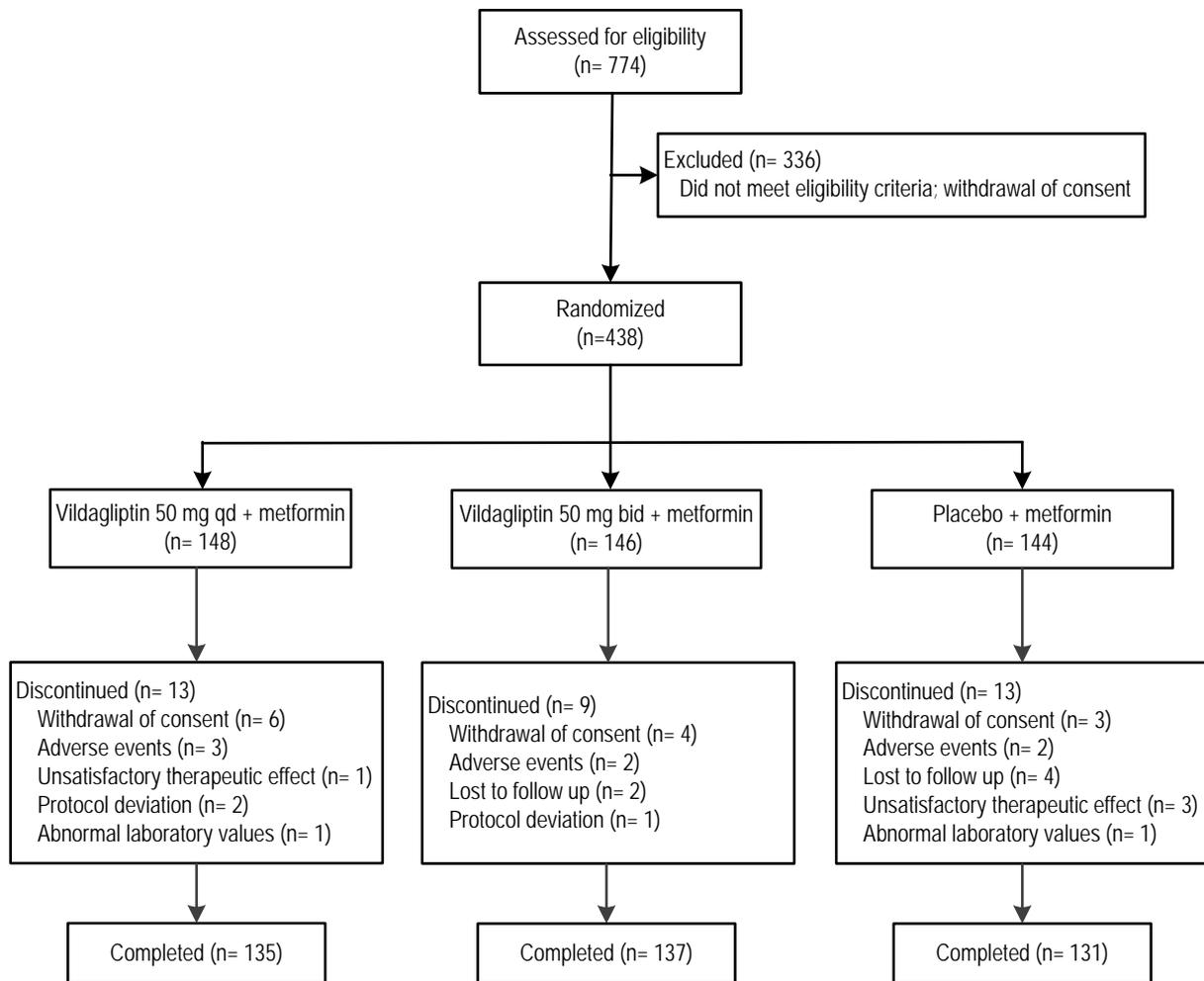


Abbildung 44: Flussdiagramm Studiendesign und -methodik für die Studie von Pan, 2012

Anmerkungen: Randomisierung erfolgte zu zwei Dosen Vildagliptin (50 mg zweimal täglich und 50 mg täglich) und Placebo d. h. in einem Verhältnis 1:1:1, aber nur Vildagliptin 50 mg (zweimal täglich)-Arm und Placebo-Arm wurden für die Nutzenbewertung einbezogen.

Tabelle 4-186 (Anhang): Studiendesign und -methodik für die Studie von DeFronzo 2012

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Die Studie untersuchte die Wirksamkeit und Verträglichkeit von Alogliptin mit Pioglitazon bei Metformin-behandelten Patienten mit Typ-2-Diabetes mellitus (T2DM).
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Bei der Studie handelt es sich um eine 26-wöchige, multizentrische, randomisierte, doppelblinde, placebo- kontrollierte, parallele Studie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Änderungen der Methodik nach dem Studienbeginn
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Eingeschlossen wurden Typ-2-Diabetiker, die ihren Blutzucker unter einer Metformin-Monotherapie (stabile Metformin-Dosis ≥ 1500 mg/Tag für ≥ 2 Monate) nicht mehr ausreichend kontrollieren können.</p> <p>Einschlusskriterien:</p> <ul style="list-style-type: none"> • Einverständniserklärung • Erwachsene Patienten im Alter von 18-80 Jahren • Diagnose: Typ-2-Diabetes, • BMI von 23 bis 45 kg/m², • C-Peptid-Konzentration von $\geq 0,26$ nmol/L, • Systolischer/diastolischer Blutdruck von maximal 160/100 mm Hg, • Hämoglobin-Wert von mindestens 12 g/dl für Männer und 10 g/dl für Frauen, • Alanin-Aminotransferase nicht mehr als das 2,5-fache der Obergrenze des Normalwertes, • TSH nicht größer als die Obergrenze des Normalwertes • Serumkreatinin ≤ 133 μmol für Männer und ≤ 124 μmol für Frauen <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Orale oder systematisch injizierte Glukokortikoide oder Medikamente, zur Gewichtsreduktion innerhalb der 3 Monate der Randomisierung, • Urinalbumin/Kreatininverhältnis

Item ^a	Charakteristikum	Studieninformation
		<p>>113 mg/mmol</p> <ul style="list-style-type: none"> • Eine Laserbehandlung für proliferative Retinopathie innerhalb der letzten 6 Monate, • Behandelte diabetische Gastroparese (Lähmung der Magenperistaltik), • Einen Herzfehler der Klasse III oder IV laut New York Heart Association, eine Herzoperation oder einen Herzinfarkt innerhalb der letzten 6 Monate.
4b	Studienorganisation und Ort der Studiendurchführung	Internationale Studie mit 327 Zentren in 20 Ländern
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>12 Behandlungsgruppen mit folgenden Behandlungen zusätzlich zu einer stabilen Metformin-Dosis:</p> <p>Placebo</p> <p>Alogliptin in der Dosierung von 12,5 mg oder 25 mg</p> <p>Pioglitazon in der Dosierung von 15 mg, 30 mg oder 45 mg</p> <p>Die Kombination von Alogliptin (12,5 mg) mit 15 mg, 30 mg oder 45 mg Pioglitazon</p> <p>Alogliptin (25 mg) in Kombination mit 15 mg, 30 mg oder 45 mg Pioglitazon.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Das primäre Zielkriterium war:</p> <p>Veränderung des HbA1c-Wertes in Woche 26 im Vergleich zu Studienbeginn.</p> <p>Die sekundären Zielkriterien waren:</p> <p>Änderung der Nüchternplasmaglukose (FPG)-Konzentration in Woche 26,</p> <p>Änderung von FPG und des HbA1c bei jeder Untersuchung,</p> <p>Auftreten einer hyperglykämischen Rescue, Fälle, in denen der HbA1c-Wert in Woche 26 nicht größer als 7,0% war,</p> <p>Veränderungen der β-Zell-Funktion im Vergleich zum Ausgangswert,</p> <p>Insulinresistenz und Änderung des Körpergewichts vom Ausgangswert zu Woche 26</p> <p>Unerwünschte Ereignisse, Laborbewertungen und Vitalwerte wurden erhoben und berichtet</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderung der Zielkriterien nach Studienbeginn
7	Fallzahl	<p>Randomisierungszahl: 1554 Patienten</p> <p>Placebo n=129^b,</p> <p>Alogliptin 12,5 mg n = 128</p> <p>Alogliptin 25 mg n=129</p> <p>Pioglitazon 15 mg n=130</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Pioglitazon 30 mg n=129^b, Pioglitazon 45 mg n= 129^b, Alogliptin (12,5 mg)+Pioglitazon 15 mg n=130 Alogliptin (12,5 mg)+Pioglitazon 30 mg n=130 Alogliptin (12,5 mg)+Pioglitazon 45 mg n= 130 Alogliptin (25 mg)+Pioglitazon 15 mg n=130 Alogliptin (25 mg)+Pioglitazon 30 mg n=130 Alogliptin (25 mg)+Pioglitazon 45 mg n=130 Jeweils in Kombination zu Metformin</p>
7a	Wie wurden die Fallzahlen bestimmt?	<p>Eine Gesamtzahl von 1440 Patienten (120 pro Untersuchungsgruppe) sollte randomisiert werden für eine 95% statistische <i>Power</i>, um Unterschiede in der HbA1c-Response zwischen der Gruppe Alogliptin 25 mg+Pioglitazon oder Alogliptin 12,5 mg+Pioglitazon und der Behandlungsgruppe mit Pioglitazon allein von 0,3% zu erkennen.</p> <p>Die Gesamtzahl wurde als ausreichend bestimmt, um einen Unterschied von etwa 0,46% bei einem zweiseitigen Signifikanzniveau von 0,05 für alle paarweisen Vergleiche zu erkennen.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Hyperglykämische Nofallmedikation bei folgenden Kriterien:</p> <p>Nach 1 Woche aber bevor Woche 4: FPG \geq 16,7mmol/ L Nach 4 Wochen aber bevor Woche 8: FPG \geq 15,3mmol/L Nach 8 Wochen aber bevor Woche 12: FPG \geq 13,9mmol/L Nach 12 Wochen bis zum Studienende: HbA1c \geq 8,5% und weniger als 0,5% Differenz zum Ausgangswert hatte</p>
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Keine Angaben
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Gruppen wurden nach HbA1c-Wert (<9,0% versus \geq 9,0%) in Woche 1 und nach geographischen Regionen stratifiziert..
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Metformin wurde als offene Gabe verabreicht; anderen Medikamente wurde blind anhand einer doppel-blindet double-dummy Technik verabreicht
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte	Keine Angaben

Item ^a	Charakteristikum	Studieninformation
	die Probanden/Patienten den Gruppen zu?	
11	Verblindung	Keine Angaben
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar; es wird von doppelblind und „double-dummy“ berichtet
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Es wird von double dummy berichtet, aber keine weitere Angaben dazu in der Publikation
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die Veränderungen von den Ausgangswerten bezüglich der Wirksamkeit wurden mit Hilfe einer Kovarianzanalyse (ANCOVA) durchgeführt, die Behandlungsgruppe und die geographische Region als Klasseneffekte und der HbA1c-Ausgangswert und die Ausgangs-Metformin-Dosis als Kovariablen.</p> <p>Die primäre Analyse verglich die Mittelwerts (der kleinsten Quadrate)-Veränderung des HbA1c bis Woche 26 von Patienten, die 25 Alogliptin+Pioglitazon erhielten, mit Patienten, die Pioglitazon allein verabreicht bekamen. Bei einem Signifikanzniveau von 0,05 wurde ein Kontrast gebraucht, der aus dem ANCOVA-Model abgeleitet war (er verglich den Durchschnittswert der Gruppen, die mit Pioglitazon in Kombination behandelt wurden mit dem der Gruppen, die mit Pioglitazon alleine behandelt wurden).</p> <p>War der Test signifikant, wurde die Response der Patienten, die Alogliptin+Pioglitazon erhielten, verglichen mit denjenigen, die nur Pioglitazon erhielten; mit einem Signifikanzniveau von 0,05, mit einem ähnlichen Kontrast. Mit dieser schrittweisen Strategie wurde keine Anpassung des Signifikanzniveaus für Mehrfachvergleichstests vorgenommen.</p> <p>Als unterstützende Analysen wurde jede Behandlungsgruppe mit der Alogliptin- oder Pioglitazongruppe verglichen. Jede Gruppe wurde außerdem mit der Placebogruppe verglichen</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Subgruppenanalysen bezüglich der Wirksamkeit :</p> <ul style="list-style-type: none"> • Rasse • Geschlecht • Ethnische Zugehörigkeit • BMI-Gruppe • Alter

Item ^a	Charakteristikum	Studieninformation
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 45)
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Siehe <i>Flow-Chart</i> (Abbildung 45)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Es wurden alle Patienten in die Analyse eingeschlossen.
14	Aufnahme / Rekrutierung	Mai 2006 – März 2008
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	<i>Pre-Screening</i> Phase (nur für Patienten mit einer Mefomin-Dosis von 1000mg/Tag) 12 Wochen Titrationsphase (optional) <i>Screening</i> -Phase <i>Run-in</i> / Stabilisierungsphase Behandlungsphase 2 Wochen <i>Follow-up</i> -Phase
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

Abkürzungen: ANCOVA, Analysis of covariance; HbA1c, Glykiertes Hämoglobin; T2DM, Type-2- Diabetes mellitus; FPG, Nüchternplasmaglukose; ^b, nur Pioglitazon 30 mg and 45 mg sind zugelassen

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

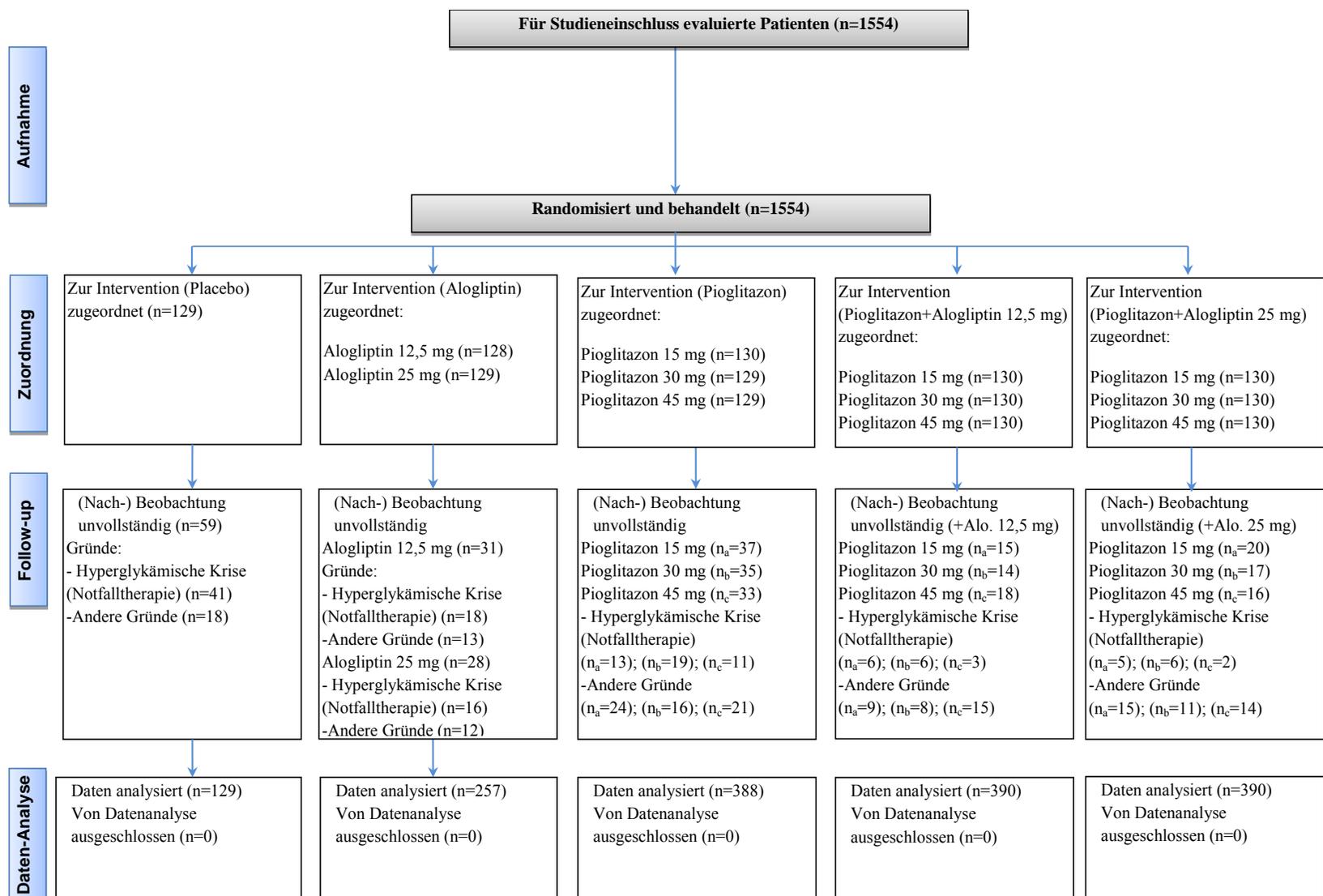


Abbildung 45: Flussdiagramm Studiendesign und -methodik für die Studie von DeFronzo 2012

Tabelle 4-187 (Anhang): Studiendesign und -methodik für die Studie von Gallwitz 2012

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Studienziel war die Überprüfung der Wirksamkeit und Verträglichkeit von Linagliptin verglichen mit Glimperid bei Patienten mit Diabetes Typ 2, die mit Metformin nicht ausreichend kontrolliert waren
-	Methoden	
3	Studiendesign	Randomisierte, kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Es handelt sich um eine randomisierte, doppelblinde, parallele, aktivkontrollierte Studie (non-inferiority) im Zuteilungsverhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	Patienten mit Diabetes Typ 2, die mit Metformin nicht ausreichend kontrolliert waren
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> • Patienten mit Diabetes Typ 2 im Alter von 18-80 Jahren, die eine stabile Dosis Metformin von 1500 mg/Tag oder mehr (oder aber eine geringere, maximal verträgliche Dosis) erhielten. • Patienten, die unter Metformin-Monotherapie einen HbA1c von 6,5–10% aufwiesen oder in Kombination mit einem oralen Antidiabetikum einen HbA1c von 6–9% hatten. Bei Beginn des Placebo <i>run-in</i> Phase mussten alle Patienten einen HbA1c-Wert von 6,5-10% aufweisen. • Patienten, die einen maximalen BMI von 40 kg/m² hatten. <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Patienten, die einen myokardialen Infarkt, Schlaganfall oder eine transitorische, ischämischen Attacke innerhalb der letzten 6 Monate hatten; • Patienten mit beeinträchtigter Leberfunktion bei Screening und • Behandlung mit Rosiglitazon, Pioglitazon, GLP-1-Analoga oder Antagonisten, Insulin oder Medikamente gegen Fettleibigkeit in den letzten 3 Monaten.

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische, internationale Studie durchgeführt an 209 Zentren in 16 Ländern (Bulgarien, Dänemark, Deutschland, Frankreich, Großbritannien, Hongkong, Irland, Indien, Italien, Niederlande, Norwegen, Polen, Südafrika, Schweden, Ungarn, USA)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Linagliptin 5 mg täglich (n=776) oder Glimepirid 1 mg täglich verabreicht zur stabilen Dosis Metformin. Nach der anfänglichen Dosis von 1 mg, wurde die Dosis Glimepirid schrittweise alle vier Wochen erhöht bis zur maximalen Dosis von 4 mg. Die Erhöhung der Dosis wurde von den Prüfarzten verabreicht, wenn die FPG-Werte der Patienten größer als 6,1 mmol/L waren. Die Dosis konnte zu jeder Zeit verringert werden, um Hypoglykämien vorzubeugen.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäres Zielkriterium: Veränderungen der HbA1c-Ausgangswerte an Woche 104</p> <p>Sekundäre Zielkriterien:</p> <ul style="list-style-type: none"> • Auftreten von hypoglykämischen Vorfällen und Veränderung des Körpergewichts bis Woche 104, • Veränderungen der HbA1c-Ausgangswerte bis Woche 52, • Verringerung des HbA1c (von 0,5% oder mehr in Woche 104), • HbA1c-Werte von weniger als 7% oder weniger als 6,5% in Woche 104 • Veränderungen des Nüchternplasmaglukosewerts (FPG) bis Woche 52 und Woche 104 • Veränderungen der 2-Stunden postprandialen Glukose bis Woche 104 während einem Mahlzeiten-Toleranz-Test mit einer Teilpopulation der Patienten, die ein standardisiertes Frühstück erhielten. • Veränderungen des Cholesterin, LDL Cholesterin, HDL Cholesterin und der Triglyceride. • Veränderungen des Plasma Proinsulin/Insulin-Verhältnisses sowie HOMA-IR bis Woche 104 und der pankreatischen β-Zell-Funktion. • Unerwünschte Ereignisse

Item ^a	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	Anzahl der randomisierten Patienten: n=1552 Behandelt mit Linagliptin: n= 776 Behandelt mit Glimperid: n=775
7a	Wie wurden die Fallzahlen bestimmt?	SAS 9.2 wurde für alle Analysen benutzt. Bei einer Annahme einer Standardabweichung des HbA1c-Ausgangswertes von 1,3% wird eine Gesamtzahl von 707 Patienten pro Behandlungsgruppe benötigt für eine 90%-ige Power, um Nichtunterlegenheit (mit einem KI von 97,5%) für die Behandlungsdifferenz in der erwarteten Mittelwertsänderung von weniger als 0,35% HbA1c mit einem Level von $\alpha=0,0125$ (einseitig) zu zeigen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurde eine Zwischenanalyse in Woche 52 vorgenommen. Ein <i>Data Safety Monitoring Board</i> wertete die Daten aus mit dem Ergebnis, die Studie bis Woche 104 weiterzuführen. Um das gesamte Signifikanzniveau von 0,05 (zweiseitig) trotz multipler Tests des primären Endpunkts zu gewährleisten, wurde das α Level unterteilt in 0,025 (zweiseitig) für die Zwischen- und die Vollanalyse. Die Nicht-Unterlegenheit des HbA1c wurde mit einem KI von 97,5% getestet (mehr als der Standard von 95% KI). Notfall-Behandlung (Pioglitazon) konnte während der Studie begonnen werden, wenn der Teilnehmer zu irgendeinem Zeitpunkt einen FPG-Wert von mehr als 13,3 mmol/L oder einen HbA1c-Wert von mehr als 8,5% von Woche 28-104 aufwies. Wenn Patienten anschließend die vorgegeben Kriterien der glykämischen Kontrolle nicht erreichten, wurden sie von der Studienteilnahme ausgeschlossen.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Computergenerierte Randomisierung
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratifiziert wurde nach HbA1c (<8,5% versus $\geq 8,5\%$) und Vorbehandlung (Metformin Monotherapie versus Metformin+anderes orales Antidiabetika).
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax /	Metformin wurde offen verabreicht, wohingegen die Gabe von Linagliptin und Glimperid verblindet war (doppel-blind und doppel-dummy)

Item ^a	Charakteristikum	Studieninformation
	Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Zuteilung zur Studienmedikation wurde durch das <i>interactive web response system</i> (IWRS) durchgeführt, mit Randomisierungs-codes, die durch den Sponsor generiert wurden.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Prüfarzte nahmen die Patienten auf. Diese wählten Patienten ihres Zentrums aus, die sie für geeignet hielten.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Studienteilnehmer und Prüfarzte waren für die Dauer der Studie verblindet. Nur dedizierte Personen hatten Zugriff auf die Randomisierungs-codes für die Zuteilung und konnten diese im Notfall zur Verfügung stellen. Während der doppelblinden Behandlung erhielten alle Studienteilnehmer außerhalb der USA eine Tablette und eine Kapsel täglich: entweder 5 mg Linagliptin+Placebo oder Glimepirid+Placebo. Teilnehmer der USA (Hier waren 3 mg-Tabletten nicht erhältlich) erhielten eine Tablette und 2 Kapseln täglich: entweder eine 5 mg Tablette Lingagliptin+2 Kapseln Placebo oder eine 1 mg, 2 mg oder 4 mg Kapsel Glimepirid+2 Placebo-Kapseln oder 1 mg Glimepirid+2 mg Glimepirid+Placebo.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Placebos waren mit der anderen Behandlung äußerlich identisch.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Primäranalysen wurden mittels ANCOVA-Modell durchgeführt, mit der Behandlung und der vorausgehenden Einnahme von oralen Antidiabetika als feste Effekte und des HbA1c-Ausgangswertes als Kovariate. Die Analyse wurde mit dem <i>Full Analysis Set</i> vorgenommen und mittels der LOCF-Methode die fehlenden Daten imputiert. (FAS schloss alle Patienten ein, die mindestens eine Behandlungsdosis bekamen und die eine behandelte und eine unbehandelte Messung erfuhren) Es gab eine Zwischenanalyse in Woche 52 und eine Hauptanalyse in Woche 104. Um das Signifikanzniveau α von 0,05 (zweiseitig) trotz multipler Tests des primären Endpunkts zu gewährleisten, wurde das α Level unterteilt in 0,025 (zweiseitig) für die Zwischen- und die Hauptanalyse. Die

Item ^a	Charakteristikum	Studieninformation
		<p>Nicht-Unterlegenheit des HbA1c wurde mit einem KI von 97,5% getestet (mehr als der Standard von 95% KI). Zusätzlich wurde α 0,0125 (einseitig) festgelegt, um die Nicht-Unterlegenheit zu testen.</p> <p>Die Analysen, die für die Veränderung des HbA1c beschrieben wurden, wurden ebenfalls für die Veränderung der FPG verwendet. Veränderungen der 2-h PPG, des Plasma Proinsulin:Insulin-Verhältnisses und der HOMA-IR wurden mit ANCOVA festgestellt. Veränderungen der PPG wurden mittels MTT festgestellt (LOCF).</p> <p>Veränderungen des Plasma Proinsulin:Insulin-Verhältnisses und der HOMA-IR schlossen das <i>Full Analysis Set</i> mit ein.</p> <p>Sekundäranalysen für das Auftreten von Hypoglykämien wurden mit dem <i>Cochran-Mantel-Haenszel</i> Test im <i>Treated Set</i> (alle Patienten, die mindestens eine Dosis erhielten) festgestellt. Die Analyse für die Veränderung des Körpergewichts wurde mittels ANCOVA Model durchgeführt (LOCF).</p> <p>Sicherheitsanalysen wurden mit den bereits erwähnten Methoden durchgeführt. Das relative Risiko (RR) kardiovaskulärer Ereignisse nach Behandlung wurde mit einem χ^2 Test verglichen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Sensitivitätsanalysen wiederholten die Primäranalyse, welche Patienten einschloss, die dem <i>Per-Protocol Set (PPS)</i> angehörten und abschlossen (Patienten des <i>Full Analysis Set</i>, die im Wesentlichen nicht gegen das Protokoll verstießen, 684 Tage an der Behandlung teilnahmen und eine HbA1c-Messung an Woche 104 hatten).</p> <p>Um die Patienten mit dem größten Behandlungseffekt zu identifizieren, wurde eine Post-hoc <i>Responder</i> Analyse für die Veränderung des HbA1c-Wertes bis Woche 104 an der oben beschriebenen Patientenpopulation durchgeführt. Diese Patienten wiesen zudem bestimmte glykämische Werte auf: (Woche 4-12: FPG \leq13,3 mmol/L; Woche 12-16: FPG \leq12,2 mmol/L; Woche 16-28: FPG \leq1,11 mmol/L; Woche 28- 52: HbA1c \leq8%; Woche 52-104: HbA1c \leq7,5%).</p>
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe <i>Flow-Chart</i> (Abbildung 46)

Item ^a	Charakteristikum	Studieninformation
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) 1552 b) 1551 c) 1519
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flow-Chart</i> (Abbildung 46)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Studie begann am 12. Februar 2008 und endete am 21. Dezember 2010. Der Studie beinhaltete einen 2-wöchigen <i>Run-in</i> (mit einer vorausgehenden 6-wöchigen <i>Washout</i> -Phase für mit oralen Antidiabetika Behandelte), eine doppelblinde, 104-wöchige Behandlungsphase und eine 1-wöchige Nachbeobachtung (<i>follow-up</i>).
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend
a: nach CONSORT 2010.		

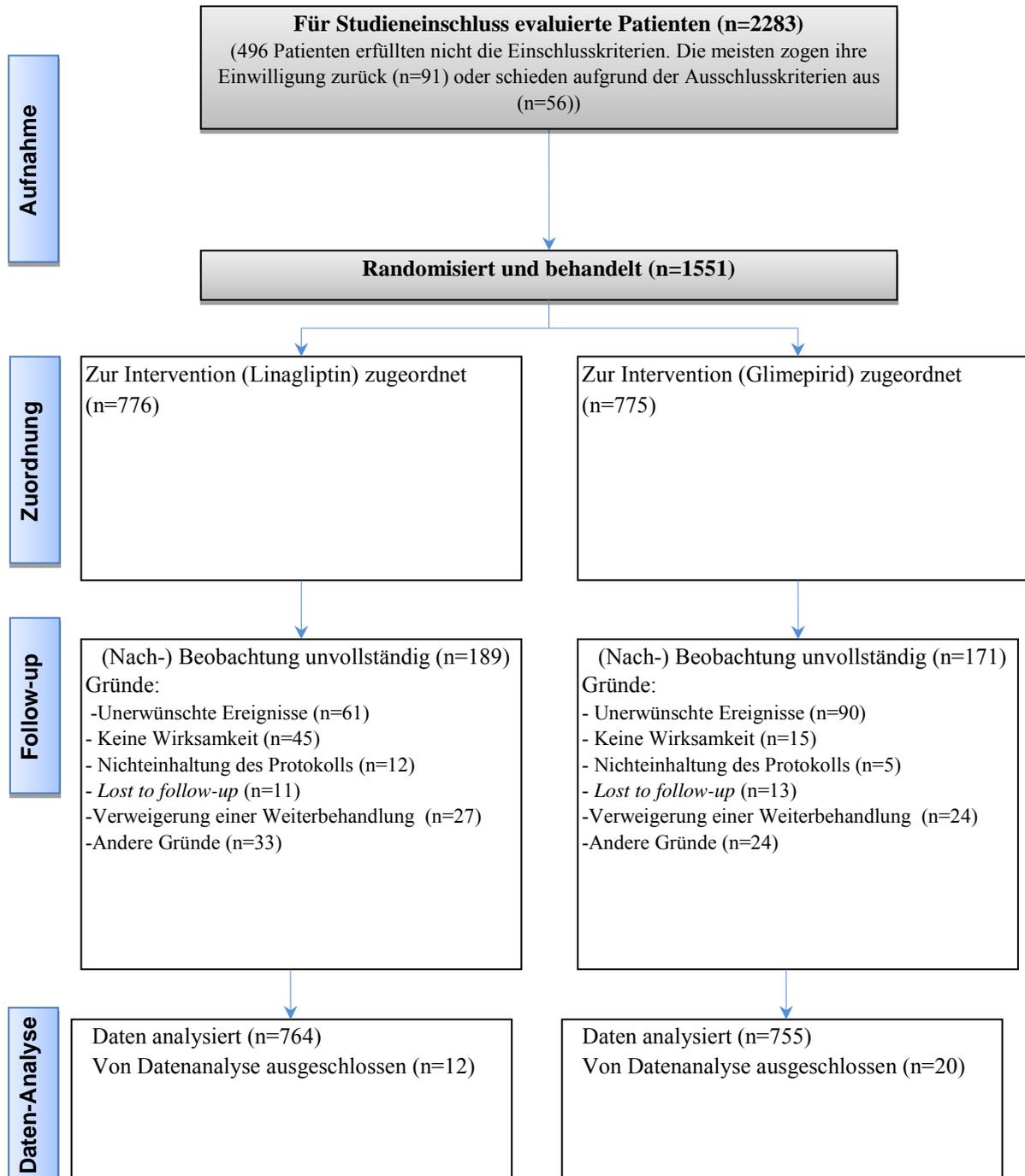


Abbildung 46: Flussdiagramm Studiendesign und -methodik für Studie Gallwitz 2012

Tabelle 4-188 (Anhang): Studiendesign und -methodik für die Studie von Yang 2012

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Das Ziel der vorliegenden Studie war die Evaluierung der Wirksamkeit, Sicherheit und Verträglichkeit von Sitagliptin zusätzlich zur bestehenden Metformin-Therapie in chinesischen Typ-2-Diabetes-Patienten, die unter Metformin-Monotherapie keine ausreichende Blutzuckerkontrolle erzielten.
-	Methoden	
3	Studiendesign	Randomisierte kontrollierte Studie
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Zuordnung parallel (Verhältnis 1:1), placebo-kontrolliert, randomisiert
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	Typ 2 Diabetiker, deren Blutzucker mit Metformin nicht ausreichend kontrolliert war
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> • Chinesische Typ-2-Diabetiker mit unzureichender Blutzuckerkontrolle unter Metformin • Alter 18 – 78 Jahre • Metformin Monotherapie mit stabiler Dosis vom 1000 oder 1700 mg/Tag (500 oder 850 mg zweimal täglich) entweder zum Zeitpunkt des Einschlusses oder nach Metformin-Dosisstabilisierung, OAD washout und Einlaufphase • Metformin-basierte duale Kombinationstherapie (d.h. Metformin in Kombination mit einem anderen OAD, außer einem PPARγ-Agonist) • HbA_{1c} \geq 7,5% und \leq 11,0% bei Metformin-Monotherapie, bzw. \geq 7,0% und \leq 9,0% bei Metformin-basierter dualer Kombinationstherapie <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Type-1-Diabetes, vorbelastet mit diabetischer Ketoazidose • Vorliegen einer Leber- oder Gallenblasenerkrankung • kongestive Herzinsuffizienz • instabile koronare Herzkrankheit • erhöhte Werte der Leberenzyme (d.h. mehr als zweifach erhöht im Vergleich zum höchsten Normalwert)

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Schwangerschaft, Stillen • jegliche Kontraindikation zur Verabreichung von Metformin • andere OADs waren während der Studie nicht erlaubt • Ein Studienprotokoll mit zusätzlichen Details zu Ein- und Ausschluss ist als Anhang zur Hauptpublikation öffentlich erhältlich.
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie in 17 Zentren in China
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Sitagliptin 100 mg einmal täglich plus Metformin (n=197) Placebo plus Metformin (n=189) Metformin wurde <i>open label</i> (1000, 1700 mg/Tag) während der Studie zur Verfügung gestellt.
6	Zielkriterien	Veränderung des HbA1c zum Zeitpunkt 24 Wochen im Vergleich zu baseline
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäres Zielkriterium:</p> <ul style="list-style-type: none"> • Veränderung des HbA1c von baseline zu Woche 24 <p>Sekundäre Zielkriterien:</p> <ul style="list-style-type: none"> • Veränderung des HbA1c im Vergleich zu Baseline bei individueller Metformin-Dosierung, • Veränderung des Blutzuckerwerts 2h postprandial (2-h PMG) und nüchtern (FPG) zum Zeitpunkt 24 Wochen im Vergleich zu Baseline für die gesamte Kohorte <p>weitere Schwerpunkte:</p> <p>Anteile der Patienten mit HbA1c unterhalb der Zielkriterien < 7,0% und < 6,5%, Veränderungen von Nüchtern-Insulin, HOMA-β (<i>homeostasis model assessment of β-cell function</i>), HOMA-IR (<i>homeostasis model assessment of insulin resistance</i>), QUICKI (<i>quantitative insulin sensitivity check index</i>) und prozentuale Veränderung von Lipiden (Gesamt-Cholesterin, LDL-Cholesterin, Triglyceride) zum Zeitpunkt 24 Wochen verglichen mit Baseline vorher bestimmte Endpunkte zur Sicherheit: Hypoglykämie, Körpergewicht, ausgewählte gastrointestinale unerwünschte Nebenwirkungen (abdominale Schmerzen, Übelkeit, Erbrechen, Diarrhoe)</p> <p>Labordaten zur Sicherheit wurden während der Studie erhoben und umfassten</p>

Item ^a	Charakteristikum	Studieninformation
		blutchemische Messungen, komplettes Blutbild und Urinstatus.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	randomisiert: Sitagliptin + Metformin: n=197; Placebo + Metformin: n=198
7a	Wie wurden die Fallzahlen bestimmt?	Im Rahmen der Studie war vorgesehen jeweils 200 Patienten (Verhältnis 1:1) in die Sitagliptin und Placebogruppe einzuschließen um 99,0% power zu erreichen (2-seitig, $\alpha=0,05$) um die primäre Hypothese zu beweisen. Diese behauptet, dass Sitagliptin effektiver ist als Placebo bei Patienten mit Metformintherapie, gemessen an einer Veränderung des HbA1c zum Zeitpunkt 24 Wochen im Vergleich zu <i>Baseline</i> .
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Bei Patienten, die die zunehmend strengeren Blutzuckerzielwerte nicht erreichten, war die Initiierung einer <i>Rescue</i> -Therapie mit Glipizide <i>open-label</i> (Dosierung nach Ermessen des Prüfarztes) erforderlich, wenn der Nüchtern-Blutzuckerwert (FPG) folgende spezifische Kriterien überstieg: >15,0 mmol/l nach Tag 1 bis Woche 6, >13,3 mmol/l nach Woche 6 bis Woche 12 und >11,1 mmol/l nach Woche 12
8	Randomisierung, Erzeugung der Behandlungsfolge	computer-basierte Generierung
8a	Methode zur Generierung der zufälligen Zuteilung	computer-basierte Generierung
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung erfolgte bei <i>Baseline</i> stratifiziert nach Metformin-Dosis (1000 oder 1700 mg/Tag) gemäß einer computer-generierten Reihenfolge
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Metformin wurde <i>open label</i> zur Verfügung gestellt wohingegen Sitagliptin doppelblind bereitgestellt wurde
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Unklar: Autoren beschreiben die Studie als doppelblind (mit Metformin open-label), die Methode der Verblindung wurde jedoch nicht berichtet. Angestellte eines zentralen Labors, die die Laboruntersuchungen durchführten, waren

Item ^a	Charakteristikum	Studieninformation
		verblindet gegenüber der Zugehörigkeit zur Behandlungsgruppe
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die Wirksamkeit wurde mittels <i>full-analysis set</i> analysiert (separat definiert für jeden Endpunkt als alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben und bei denen Daten zum entsprechenden Endpunkt an baseline und mindestens einmal nach baseline erhoben wurden bevor die <i>Rescue</i>-Medikation eingesetzt wurde) und fehlende Daten gemäß der LOCF-Methode berücksichtigt. Die Veränderung des HbA1c und die prozentuale Veränderung der Lipide verglichen mit baseline wurden mit Hilfe eines ANCOVA (analysis of covariance)-Modells analysiert, wobei für Behandlung, Studienzentrum, Metformin-Dosis und baseline-Wert des entsprechenden Endpunkts kontrolliert wurde.</p> <p>Die analytische Vorgehensweise für weitere Parameter der Wirksamkeit und Sicherheit umfasste unter anderem ein logistisches Regressionsmodell um den Anteil der Patienten, die die HbA1c-Ziele erreichten zu analysieren, und ein nicht-parametrisches ANCOVA-Modell zur Analyse der prozentualen Veränderung der Triglyceride. Zur Analyse der Gewichtsveränderungen wurde die bereits in der Wirksamkeitsanalyse verwendete ANCOVA angewandt. Für andere unerwünschte Nebenwirkungen und prädefinierte Grenzwerte für Abweichungen der Laborwerte wurde die Methode nach Miettinen und Nurminen (1985) angewandt. Eine Adjustierung für multiple Vergleiche wurde nach dem Hochberg-Verfahren (1988) bzw. mit einem geordneten Testverfahren durchgeführt, um die Typ 1 Fehlerrate bei den Schlüssel-Endpunkten zur Wirksamkeit zu kontrollieren.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Vordefinierte Subgruppenanalysen der HbA1c-Veränderung verglichen mit baseline wurden nach Geschlecht, Alter (< 65 oder ≥ 65 Jahre), Metformin-Dosislevel, baseline BMI (> Median oder ≤ Median), <i>Baseline</i> HbA1c (> Median oder ≤ Median), vorangegangener OAD-Therapie und bekannter Dauer des T2DM durchgeführt.

Item ^a	Charakteristikum	Studieninformation
-		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	siehe Abbildung 47
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	siehe Abbildung 47
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	siehe Abbildung 47
14	Aufnahme / Rekrutierung	Die Studie wurde zwischen 13.01.2009 und 09.08.2010 durchgeführt
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Studie bestand aus fünf Abschnitten: screening, Metformin Dosiserhöhung, Metformin Dosisstabilisierung, einfach-blinde placebo und doppelblinde Behandlungsperiode. Es erfolgten fünf Untersuchungen während der randomisierten Behandlungsperiode: Tag 1 (baseline), Woche 6, 12, 18, und 24.
14b	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend - Studie endete wie geplant
a: nach CONSORT 2010.		

Abkürzungen: ANCOVA, Analysis of covariance; FPG, Fasting plasma glucose; HbA1c, Glykosyliertes Hämoglobin; HOMA- β , homeostasis model assessment of insulin resistance of β -cell function; HOMA-IR, homeostasis model assessment of insulin resistance; LOCF, Last observation carried forward; OAD, orales Antidiabetikum; PMG, Post-meal glucose; PPAR, peroxisome proliferator-activated receptor; QUICKI, quantitative insulin sensitivity check index; T2DM, Typ 2 Diabetes mellitus

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

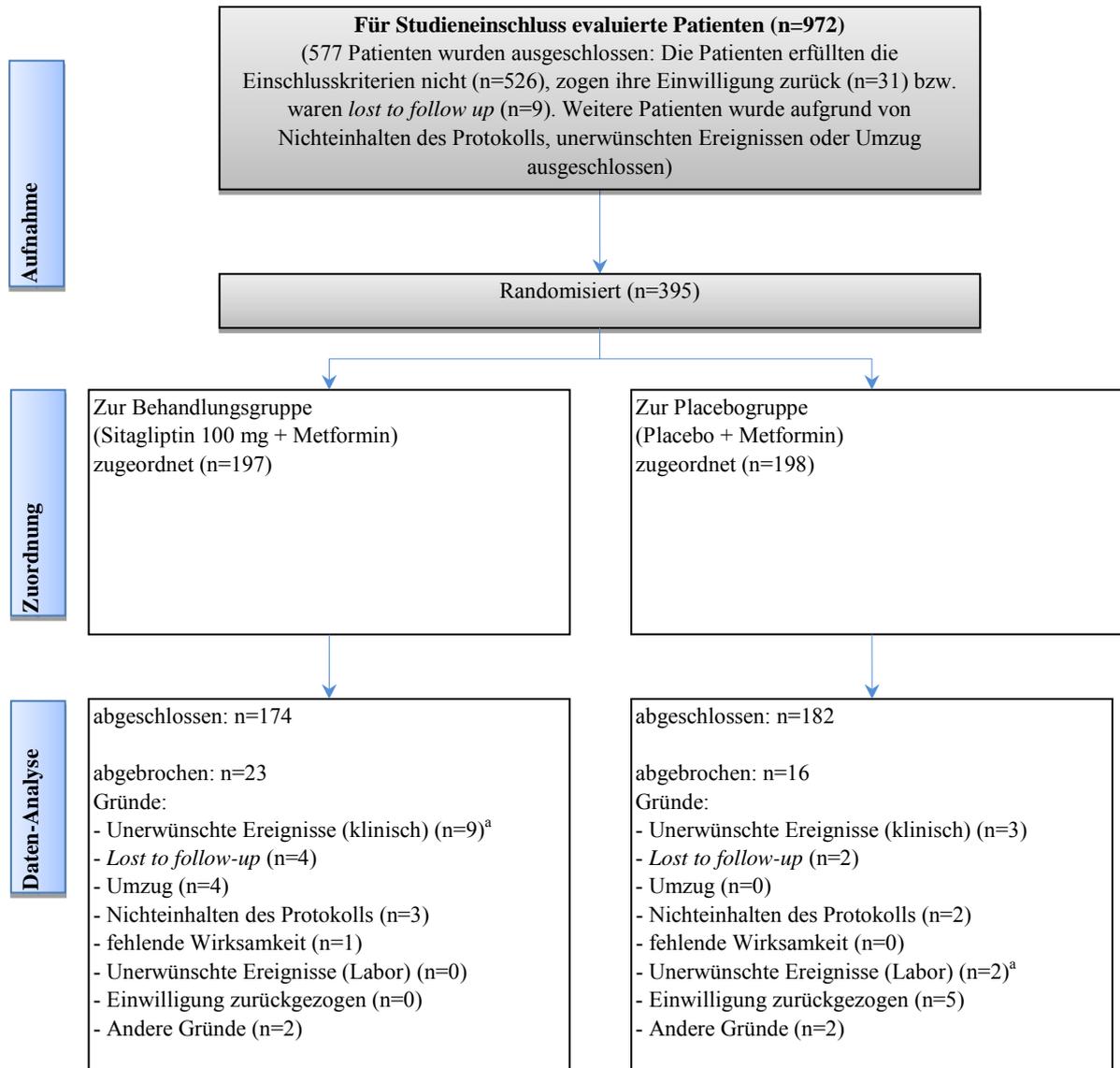


Abbildung 47: Flussdiagramm Studiendesign und -methodik für Studie Yang 2012

^aeinschließlich einem Patient, der die Studie aufgrund eines unerwünschten Ereignisses, das bereits vor Randomisierung begann, abbrach

Anhang 4-G: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.

Tabelle 4-189 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie D1690C00004

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: D1690C00004

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studie D1690C00004 (AstraZeneca, 2010b)	CSR04short
Studie D1690C00004 Verlängerungsphase I (AstraZeneca, 2011b)	CSR04long
Nauck et. al, 2011 (Nauck et al., 2011)	PubNauck2011

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (CSR04short)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die verblindete Behandlung erfolgte anhand einer double-dummy Technik;

Prüf- und Vergleichsmedikation waren identisch. (CSR04short)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die verblindete Behandlung erfolgte anhand einer double-dummy Technik;

Prüf- und Vergleichsmedikation waren identisch. (CSR04short)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis

kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer

Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbrüche aufgrund unzureichender glykämischer Kontrolle,

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in

begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Hypoglykämien

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Gewichtsveränderung**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Patienten mit einer Gewichtsabnahme von mehr als 5%**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Therapiezufriedenheit (DTSQ)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Blutdrucksenkung

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Kardiale Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Erhebung und Beurteilung des Endpunktes erfolgte verblindet. (CSR04short)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-190 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie MB102014 / Bailey, 2010

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Studie MB102014 / Bailey, 2010

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Bailey et al, 2010 (Bailey et al., 2010)	PubBailey2010
Studie MB102014 (Bristol-Myers Squibb Company, 2009)	MB102014

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.
→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin. Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung* *inadäquater* *statistischer* *Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung Gewichts (kg) vom Ausgangswert bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine*Beachte die Hinweise zu Punkt A.4!*

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-191 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie D1690C00012

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: D1690C00012

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Bolinder et al, 2012 (Bolinder et al., 2012)	PubBolinder2012
Studie D1690C00012 (AstraZeneca, 2010a)	D1690C00012
Studie D1690C00012 Verlängerungsphase (AstraZeneca, 2011a)	D1690C00012 Verlängerungsphase

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin. Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- intransparenter Patientenfluss*
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-192 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Scott, 2008

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Scott, 2008

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Scott et al, 2008 (Scott et al., 2008)	PubScott2008

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubScott2008)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Angaben zur Erzeugung der Randomisierungssequenz gemacht. (PubScott2008)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Angaben zur Verdeckung der Gruppenzuteilung gemacht. (PubScott2008)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren berichten von einer doppelblinden Studie. Es wurde jedoch nicht darauf eingegangen, wer und wie verblindet wurde. (PubScott2008)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren berichten von einer doppelblinden Studie. Es wurde jedoch nicht darauf eingegangen, wer und wie verblindet wurde. (PubScott2008)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 18)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 18**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 18

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in

begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-193 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Charbonnel, 2006

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Charbonnel, 2006

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Charbonnel et al, 2006 (Charbonnel et al., 2006)	PubCharbonnel2006

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubCharbonnel2006)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde keine Information zur Erzeugung der Randomisierungssequenz zur Verfügung gestellt. (PubCharbonnel2006)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde keine Information zur Verdeckung der Gruppenzuteilung zur Verfügung gestellt. (PubCharbonnel2006)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschrieben die Studie als doppelblind, es wurde jedoch nicht berichtet wie und wer verblindet wurde. (PubCharbonnel2006)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschrieben die Studie als doppelblind, es wurde jedoch nicht berichtet wie und wer verblindet wurde. (PubCharbonnel2006)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.

- Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubCharbonnel2006)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubCharbonnel2006)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubCharbonnel2006)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubCharbonnel2006)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubCharbonnel2006)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubCharbonnel2006)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-194 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von DeFronzo, 2009

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: DeFronzo, 2009

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
DeFronzo et al, 2009 (DeFronzo et al., 2009)	PubDeFronzo2009

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubDeFronzo2009)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2009).

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde nicht genügend Information zur Verfügung gestellt. Die Durchführung des ITT-Prinzips in Nicht-Unterlegenheits-Studien ist nicht immer anwendbar oder der beste Weg, um die Daten zu analysieren. Im Studienbericht werden beide Ergebnisse dargestellt („Per Protokoll“ und „alle behandelten Patienten“ bzw. modifiziertes ITT). (PubDeFronzo2009)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- unplausible Angaben*
- Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-195 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Bosi, 2007

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Bosi, 2007

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Bosi et al, 2007 (Bosi et al., 2007)	PubBosi2007

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubBosi2007)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Angaben zur Erzeugung der Randomisierungssequenz gemacht. (PubBosi2007)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Angaben zur Verdeckung der Gruppenzuteilung gemacht. (PubBosi2007)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschrieben die Studie als doppelblind. Es wurde jedoch nicht darüber berichtet, wer und wie verblindet worden ist. (PubBosi2007)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschrieben die Studie als doppelblind. Es wurde jedoch nicht darüber berichtet, wer und wie verblindet worden ist. (PubBosi2007)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Die Definition der primären ITT-Population ist unklar. Der Artikel gab an, dass die primären Wirksamkeitsanalysen mit Daten von Patienten gemacht worden sind, welche eine verlässliche Untersuchung der HbA1c-Wertes von $\geq 7,4\%$, mindestens ein Studienmedikament erhalten, einen verlässlichen Ausgangswert und mindestens einen Folgewert aus einer HbA1c-Messung hatten. Diese Population wurde als primäre ITT-Population bezeichnet und wurde präspezifiziert als Hauptwirksamkeit-Population. Es wurde nicht klar, was ein „verlässlicher“ HbA1c bedeutet. (PubBosi2007)

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 24)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBosi2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die primäre Wirksamkeitsanalyse wurde mit Daten von Patienten durchgeführt, die einen verlässlichen HbA1c-Wert von $\geq 7,4\%$, mindestens ein Studienmedikament erhalten, einen verlässlichen Ausgangswert und mindestens einen verlässlichen Folgewert für HbA1c hatten. Diese Population wurde als primäre ITT-Population definiert und wurde präspezifiziert als Hauptwirksamkeits-Population. Es wurde nicht darüber berichtet, wie „verlässlicher“ HbA1c definiert wurde. Deshalb könnte dies eine Quelle für eine Verzerrung darstellen. (PubBosi2007)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Die Definition der primären ITT-population war unklar. Aus diesem Grund konnte keine Abschätzung gemacht werden. (PubBosi2007)

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBosi2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die primäre Wirksamkeitsanalyse wurde mit Daten von Patienten durchgeführt, die einen verlässlichen HbA1c-Wert von $\geq 7,4\%$, mindestens ein Studienmedikament erhalten, einen verlässlichen Ausgangswert und mindestens einen verlässlichen Folgewert für HbA1c hatten. Diese Population wurde als primäre ITT-Population definiert und wurde präspezifiziert als Hauptwirksamkeits-Population. Es wurde nicht darüber berichtet, wie „verlässlicher“ HbA1c definiert wurde. Deshalb könnte dies eine Quelle für eine Verzerrung darstellen. (PubBosi2007)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Die Definition der primären ITT-population war unklar. Aus diesem Grund konnte keine Abschätzung gemacht werden. (PubBosi2007)

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBosi2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund der hohen Verzerrung auf Studienebene (PubBosi2007)

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu

definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund der hohen Verzerrung auf Studienebene (PubBosi2007)

Endpunkt: Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBosi2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine randomisierte Studie.(PubBosi2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund der hohen Verzerrung auf Studienebene (PubBosi2007)

Tabelle 4-196 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Taskinen, 2011

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Taskinen, 2011

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Taskinen et al, 2011 (Taskinen et al., 2011)	PubTaskinen2011

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubTaskinen2011)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Diese Information wurde nicht zur Verfügung gestellt.
(PubTaskinen2011)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Diese Information wurde nicht zur Verfügung gestellt.
(PubTaskinen2011)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren bezeichnen ihre Studie als doppelblind, berichten aber nicht wer und wie verblindet wurde. (PubTaskinen2011)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren bezeichnen ihre Studie als doppelblind, berichten aber nicht wer und wie verblindet wurde. (PubTaskinen2011)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Unklar: Patienten, die eine Notfallmedikament erhalten haben, blieben in der Studie. Es wurde jedoch nicht berichtet, ob die LOCF-Methode für diese Patienten genutzt wurde. Falls nicht könnten die Ergebnisse verzerrt sein. (PubTaskinen2011)

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubTaskinen2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Unklar: Patienten, die eine Notfallmedikament erhalten haben, blieben in der Studie. Es wurde jedoch nicht berichtet, ob die LOCF-Methode für diese Patienten genutzt wurde. Falls nicht könnten die Ergebnisse verzerrt sein. (PubTaskinen2011)

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund fehlender Informationen (PubTaskinen2011)

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubTaskinen2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Gewichtsveränderung (vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubTaskinen2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund fehlender Informationen (PubTaskinen2011)

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubTaskinen2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubTaskinen2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubTaskinen2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung

angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-197 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Bergenstal, 2010a

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Bergenstal, 2010a

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Bergenstal et al, 2010a (Bergenstal et al., 2010a)	PubBergenstal2010a

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubBergenstal2010a)

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Angaben zur Erzeugung der

Randomisierungssequenz angegeben. (PubBergenstal2010a)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Angaben zur Verdeckung der Gruppenzuteilung angegeben. (PubBergenstal2010a)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine verblindete Studie. (PubBergenstal2010a)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine verblindete Studie. (PubBergenstal2010a)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden nicht genügend Angaben im Abstract angegeben, um eine Abschätzung abzugeben. (PubBergenstal2010a)

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es wurden nicht genügend Angaben im Abstract angegeben, um eine Abschätzung abzugeben. (PubBergental2010a)

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Im Abstract sind nicht genügend Informationen angegeben um eine Verzerrung auszuschließen. (PubBergental2010a)

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu

anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBergenstal2010a)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden nicht genügend Angaben im Abstract angegeben, um eine Abschätzung über die adäquate Umsetzung des ITT-Prinzips abzugeben. (PubBergental2010a)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Im Abstract waren nicht genügend Angaben, um eine Abschätzung abzugeben. (PubBergental2010a)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Im Abstract waren nicht genügend Angaben, um eine Abschätzung abzugeben. (PubBergenstal2010a)

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Im Abstract wurden nicht genügend Informationen zur Verfügung gestellt, um eine Verzerrung ausschließen zu können. (PubBergenstal2010a)

PanTabelle 4-198 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Bolli, 2008

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Bolli, 2008

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Bolli et al, 2008 (Bolli et al., 2008)	PubBolli2008

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubBolli2008)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es handelt sich um eine verblindete Studie. (PubBolli2008)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBolli2008)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse des Endpunkts wurde mit der „per Protokoll“-Population durchgeführt (benötigt wurde hierfür mindestens eine Dosis des Studienmedikaments und mindestens ein Folgewert nach Messung des Ausgangswerts der primären und sekundären Wirksamkeitsvariablen nicht länger als 7 Tage nach Erhalt des letzten Studienmedikaments). Die LOCF-Methode wurde für die Patienten angewendet, die frühzeitig die Studie aufgrund von unzufrieden stellendem Ansprechen auf das Medikament abbrechen, oder falls die Patienten nur 22 Wochen behandelt wurden. (PubBolli2008)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubBolli2008)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse des Endpunkts wurde mit der „per Protokoll“-Population durchgeführt (benötigt wurde hierfür mindestens eine Dosis des Studienmedikaments und mindestens ein Folgewert nach Messung des Ausgangswerts der primären und sekundären Wirksamkeitsvariablen nicht länger als 7 Tage nach Erhalt des letzten Studienmedikaments). Die LOCF-Methode wurde für die Patienten angewendet, die frühzeitig die Studie aufgrund von unzufrieden stellendem Ansprechen auf das Medikament abbrechen, oder falls die Patienten nur 22 Wochen behandelt wurden. (PubBolli2008)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBolli2008)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBolli2008)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war. In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubBolli2008)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBolli2008)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-199 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Bergenstal, 2010b

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Bergenstal, 2010b

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Bergenstal et al, 2010b (Bergenstal et al., 2010b)	PubBergenstal2010b

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubBergenstal2010b)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine verblindete Studie. (PubBergenstal2010b)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine verblindete Studie. (PubBergenstal2010b)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 26)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie (PubBergenstal2010b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 26**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie (PubBergenstal2010b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 26

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie (PubBergenstal2010b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 26**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubBergenstal2010b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine verblindete Studie. (PubBergental2010b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubBergenstal2010b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine verblindete Studie. (PubBergenstal2010b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-200 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von DeFronzo, 2005

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: DeFronzo, 2005

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
DeFronzo et al, 2005 (DeFronzo et al., 2005)	PubDeFronzo2005

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubDeFronzo2005)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Informationen über Erzeugung der Randomisierungssequenz zur Verfügung gestellt. (PubDeFronzo2005)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Informationen zur Verdeckung der Gruppenzuteilung zur Verfügung gestellt. (PubDeFronzo2005)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren bezeichnen ihre Studie als dreifach verblindet, aber beschreiben nicht wer und wie verblindet wurde. (PubDeFronzo2005)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren bezeichnen ihre Studie als dreifach verblindet, aber beschreiben nicht wer und wie verblindet wurde. (PubDeFronzo2005)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 30)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 30**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 30

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in

begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubDeFronzo2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-201 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Nauck, 2009

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Nauck, 2009

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Nauck et al, 2009 (Nauck et al., 2009)	PubNauck2009

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubNauck2009)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde keine Information zur Verdeckung der Gruppenzuteilung zur Verfügung gestellt. (PubNauck2009)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschreiben die Studie als doppelblind und die Verabreichung der Medikamente als double-dummy. Es wurde jedoch nicht berichtet, wer und wie verblindet worden ist. (PubNauck2009)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschreiben die Studie als doppelblind und die Verabreichung

der Medikamente als double-dummy. Es wurde jedoch nicht berichtet, wer und wie verblindet worden ist. (PubNauck2009)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 26)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 26**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es gibt hierzu keine ausreichenden Information. (PubNauck2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 26

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-202 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Pratley, 2010

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Pratley, 2010

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Pratley et al, 2010 (Pratley et al., 2010)	PubPratley2010

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubPratley2010)

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine offene Studie. (PubPratley2010)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine offene Studie. (PubPratley2010)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 26)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. (PubPratley2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 26**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. (PubPratley2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 26

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. (PubPratley2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 26**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. (PubPratley2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. (PubPratley2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Es besteht ein Verzerrungspotenzial bei der Berichterstattung von Nebenwirkungen (z.B. Übelkeit), da es sich um eine offene Studie handelt und Glucagon-like Peptid-1 Analoge (z.B. Liraglutide) bekannt für das Erzeugen von Übelkeit sind. (PubPratley2010)

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. (PubPratley2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu

anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine offene Studie. (PubPratley2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Es besteht ein Verzerrungspotenzial bei der Berichterstattung von Nebenwirkungen (z.B. Übelkeit), da es sich um eine offene Studie handelt und Glucagon-like Peptid-1 Analoge (z.B. Liraglutide) bekannt für das Erzeugen von Übelkeit sind. (PubPratley2010)

Tabelle 4-203 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Kaku, 2009

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Kaku, 2009

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Kaku et al, 2009 (Kaku, 2009)	PubKaku2009

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubKaku2009)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde keine Information zur Erzeugung der Randomisierung zur Verfügung gestellt. (PubKaku2009)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde keine Information zur Verdeckung der Gruppenzuteilung zur Verfügung gestellt. (PubKaku2009)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschreiben die Studie als doppelblind, es wurde jedoch nicht darüber berichtet, wer und wie verblindet wurde. (PubKaku2009)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschreiben die Studie als doppelblind, es wurde jedoch nicht darüber berichtet, wer und wie verblindet wurde. (Kaku, 2009)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Patienten-Subgruppen, bei denen die Behandlung angeschlagen hatte, wurden selektiv im Text des Ergebnisteils hervorgehoben. (PubKaku2009)

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*

- *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Patienten in der Behandlungsgruppe waren signifikant kürzer an Diabetes erkrankt (4,5 Jahre im Vergleich zu 5,6 Jahre, $p=0,04$). Die Studienpopulation war jünger als normal (Durchschnittsalter etwa 52 Jahre) (PubKaku2009)

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Das Verzerrungspotenzial liegt auf mittlere Stufe. Wie bereits beschrieben, waren die Patienten der Behandlungsgruppe gesünder. Deshalb könnten sie eine unterschiedliche Reaktion auf die Behandlung zeigen. (PubKaku2009)

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 28)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubKaku2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Bei der Behandlungsgruppe handelte es sich um Patienten mit einer geringeren Diabetesdauer. (PubKaku2009)

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Unklar. Die Ergebnisse sind möglicherweise aufgrund des Alters und der Diabetesdauer der Patienten verzerrt. (PubKaku2009)

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubKaku2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund der hohen Verzerrung auf Studienebene (PubKaku2009)

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubKaku2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu

definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund der hohen Verzerrung auf Studienebene (PubKaku2009)

Tabelle 4-204 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Charpentier, 2001

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Charpentier, 2001

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Charpentier et al, 2001 (Charpentier et al., 2001)	PubCharpentier2001

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie (PubCharpentier2001)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren bezeichnen die Behandlungsmethode als doppelblind und double-dummy. Es wurde jedoch nicht berichtet, wer und wie verblindet wurde. (PubCharpentier2001)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren bezeichnen die Behandlungsmethode als doppelblind und

double-dummy. Es wurde jedoch nicht berichtet, wer und wie verblindet wurde. (PubCharpentier2001)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Adäquate methodische Information, die relevant für die Abschätzung des Zusatznutzens anhand des jeweiligen Endpunkts waren (z.B. Verblindung, Kalkulation der Studiengröße) wurde nicht berichtet. (PubCharpentier2001)

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).

- Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
- Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Adäquate methodische Information, die relevant für die Abschätzung des Zusatznutzens anhand des jeweiligen Endpunkts waren (z.B. Verblindung, Kalkulation der Studiengröße) wurde nicht berichtet. Deshalb war es schwer ein Verzerrungspotenzial auszuschließen. (PubCharpentier2001)

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Veränderung des systolischen Blutdrucks vom Ausgangswert bis Woche 20

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren bezeichnen die Behandlungsmethode als doppelblind und double-dummy. Es wurde jedoch nicht berichtet, wer und wie verblindet wurde. (PubCharpentier2001)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Adäquate methodische Information, die relevant für die Abschätzung des Zusatznutzens anhand zu diesem Endpunkt waren (z.B. Verblindung, Kalkulation der Studiengröße) wurden nicht berichtet. (PubCharpentier2001)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Adäquate methodische Information, die relevant für die Abschätzung des Zusatznutzens anhand des Endpunkts waren (z.B. Verblindung, Kalkulation der Studiengröße) wurde nicht berichtet. Deshalb war es schwer ein Verzerrungspotenzial auszuschließen. (PubCharpentier2001)

Tabelle 4-205 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Papathanassiou, 2009

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Papathanassiou, 2009

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Papathanassiou et al, 2009 (Papathanassiou et al., 2009)	PubPapathanassiou2009

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie.
(PubPapathanassiou2009)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Randomisierungssequenz der Patienten beruhte auf der Reihenfolge der Patientenaufnahme in den Ambulanzen.
(PubPapathanassiou2009)

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Methode zur Verdeckung der Gruppenzuteilung wurde nicht berichtet. Des Weiteren handelte es sich um eine offene Studie. (PubPapathanassiou2009)

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**

- ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

- ja:** Die Patienten waren verblindet.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine offene Studie. (PubPapathanassiou2009)

Behandler bzw. Weiterbehandler

- ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine offene Studie. (PubPapathanassiou2009)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*

- Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Adäquate methodische Information, die relevant für die Beurteilung des Zusatznutzens des Endpunkts (z.B. systolischer Blutdruck), wurde nicht berichtet. (PubPapathanassiou2009)

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).

- *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
- *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Adäquate methodische Information, die relevant für die Beurteilung des Zusatznutzens des Endpunkts (z.B. systolischer Blutdruck), wurde nicht berichtet. (PubPapathanassiou2009)

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Adäquate methodische Information, die relevant für die Beurteilung des Zusatznutzens des Endpunkts (z.B. systolischer Blutdruck), wurde nicht berichtet. Deshalb war es schwer ein Verzerrungspotenzial auszuschließen. (PubPapathanassiou2009)

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelte sich um eine offene Studie. (PubPapathanassiou2009)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Adäquate methodische Information zu diesem Endpunkt wurde nicht zur Verfügung gestellt. (PubPapanthassiou2009)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Adäquate methodische Information zu diesem Endpunkt wurde nicht zur Verfügung gestellt. (PubPapathanassiou2009)

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Adäquate methodische Information zu diesem Endpunkt wurde nicht zur Verfügung gestellt. Deshalb war es schwer ein Verzerrungspotenzial auszuschließen. (PubPapathanassiou2009)

Tabelle 4-206 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Matthews, 2005

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Matthews, 2005

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Matthews et al, 2005 (Matthews et al., 2005)	PubMatthews2005

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubMatthews2005)

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubMatthews2005)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubMatthews2005)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 52**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“-)Analyse werden alle randomisierten Patienten

entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

*Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.
In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.*

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante

Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- unplausible Angaben*
- Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2005)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und

Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-207 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Nauck, 2007

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Nauck, 2007

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Nauck et al, 2007 (Nauck et al., 2007)	PubNauck2007

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubNauck2007)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Diese Information wurde nicht zur Verfügung gestellt.
(PubNauck2007)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Diese Information wurde nicht zur Verfügung gestellt.
(PubNauck2007)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubNauck2007)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine verblindete Studie. (PubNauck2007)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis zur Woche 52)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde nicht genügend Information zur Verfügung gestellt. Die Durchführung des ITT-Prinzips in Nicht-Unterlegenheits-Studien ist nicht immer anwendbar oder der beste Weg, um die Daten zu analysieren. Im Studienbericht werden beide Ergebnisse dargestellt („Per Protokoll“ und „alle behandelten Patienten“ bzw. modifiziertes ITT). (PubNauck2007)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Information wurde nicht zur Verfügung gestellt. (PubNauck2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- unplausible Angaben*
- Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und

Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

*Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.
In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.*

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante

Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Diese Information wurde nicht zur Verfügung gestellt. (PubNauck2007)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung

angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-208 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Göke, 2010

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Göke, 2010

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Göke et al, 2010 (Göke et al., 2010)	PubGöke2010

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubGöke2010)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubGöke2010)

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubGöke2010)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- intransparenter Patientenfluss*
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 52)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche52**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubGöke2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubGöke2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 52**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubGöke2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit einer Nebenwirkung**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubGöke2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubGöke2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch,

insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubGöke2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-209 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Filozof, 2010

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Filozof, 2010

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Filozof et al, 2010 (Filozof und Gautier, 2010)	PubFilozof2010

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubFilozof2010)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Information zur Erzeugung der Randomisierungssequenz fehlt.
(PubFilozof2010)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Information zur Verdeckung der Gruppenzuteilung wurde nicht zur Verfügung gestellt. (PubFilozof2010)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es handelt sich um ein doppelblindes Studiendesign. (PubFilozof2010)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es handelt sich um ein doppelblindes Studiendesign. (PubFilozof2010)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- intransparenter Patientenfluss*
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 52)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubFilozof2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse dieses Endpunkts beruht auf der „per Protokoll“-Population. Die ITT-Population wurde für die Sensitivitätsanalyse (Ergebnisse werden als vergleichbar angegeben, aber nicht berichtet) verwendet. (PubFilozof2010)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubFilozof2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse dieses Endpunktes basiert auf der „per Protokoll“-Population.
(PubFilozof2010)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um ein doppelblindes Studiendesign. (PubFilozof2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubFilozof2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um ein doppelblindes Studiendesign. (PubFilozof2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-210 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Salvadeo, 2010

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Salvadeo, 2010

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Salvadeo et al, 2010 (Abstract) (Salvadeo et al., 2010)	PubSalvadeo2010

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubSalvadeo2010)

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht.
(PubSalvadeo2010)

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht.
(PubSalvadeo2010)

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**

- ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

- ja:** Die Patienten waren verblindet.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

Behandler bzw. Weiterbehandler

- ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.
- unklar:** Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*

- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*

- *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
- *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
- *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend das Verzerrungspotenzial der Ergebnisse zu beurteilen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden

relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 52)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch,

insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend das Verzerrungspotenzial der Ergebnisse zu

beurteilen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht.
(PubSalvadeo2010)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend um eine Abschätzung zu treffen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Da nur der Abstract der Studie publiziert wurde, war die erhältliche Information nicht ausreichend das Verzerrungspotenzial der Ergebnisse zu beurteilen. Die Autoren wurden kontaktiert, jedoch antworteten sie nicht. (PubSalvadeo2010)

Tabelle 4-211 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Derosa, 2010

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Derosa, 2010

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Derosa et al, 2010 (Derosa et al., 2010)	PubDerosa2010

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubDerosa2010)

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten erhielten eine verblindete Behandlung. (PubDerosa2010)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine einfachblinde Studie. (PubDerosa2010)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 52)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelte sich um eine einfachblinde Studie. (PubDerosa2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Im Studienbericht wurde beschrieben, dass eine modifizierte ITT-Analyse durchgeführt wurde. Allerdings lässt die Ergebnisdarstellung in den Tabellen darauf schließen, dass nur Daten von Patienten verwendet wurden, die noch in der Studie waren. (PubDerosa2010)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 52)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelte sich um eine einfachblinde Studie. (PubDerosa2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Im Studienbericht wurde beschrieben, dass eine modifizierte ITT-Analyse durchgeführt wurde. Allerdings lässt die Ergebnisdarstellung in den Tabellen darauf schließen, dass nur Daten von Patienten verwendet wurden, die noch in der Studie waren. (PubDerosa2010)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelte sich um eine einfachblinde Studie. (PubDerosa2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei

einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-212 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Matthews, 2010

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Matthews, 2010

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Matthews et al, 2010 (Matthews et al., 2010)	PubMatthews2010
Ferranini, 2009 (Ferrannini et al., 2009)	PubFerrannini2009

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubMatthews2010)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Im Studienbericht wird beschrieben, dass die Patienten randomisiert wurden. Es wurde jedoch keine Information zur Erzeugung der Randomisierung berichtet. (PubMatthews2010)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Methode zur Verdeckung der Gruppenzuteilung wurde nicht berichtet. (PubMatthews2010)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubMatthews2010)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubMatthews2010)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 52)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die „per Protokoll“-Population wurde zur Berechnung dieses Endpunktes genutzt. Die ITT-Population wurde für die Sensitivitätsanalyse verwendet. (PubMatthews2010)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts (kg) vom Ausgangswert bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die LOCF-Methode wurde für die Analyse des HbA1c genutzt. Es ist jedoch nicht klar, ob diese Methode auch für die anderen Wirksamkeitsparameter genutzt worden ist. Die „per Protokoll“-Population wurden für die Analyse dieses Endpunktes verwendet. (PubMatthews2010)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 52

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu

definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es handelt sich um eine doppelblinde Studie. (PubMatthews2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubMatthews2010)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu

definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-213 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Derosa, 2011a

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Derosa, 2011a

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Derosa et al., 2011a (Derosa et al., 2011a)	PubDerosa2011a

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubDerosa2011)

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubDerosa2011)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie. (PubDerosa2011)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Veränderung des HbA1c nach 52 Wochen im Vergleich zum Ausgangswert**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubDerosa2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in

begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Autoren geben an, das ITT-Prinzip angewendet zu haben, sie berichten aber nicht, wie mit fehlenden Daten umgegangen wurde. (PubDerosa2011)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des Körpergewichts nach 52 Wochen im Vergleich zum Ausgangswert

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubDerosa2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Autoren geben an, das ITT-Prinzip angewendet zu haben, sie berichten aber nicht, wie mit fehlenden Daten umgegangen wurde. (PubDerosa2011)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbrüche aufgrund unerwünschter Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es handelt sich um eine doppelblinde Studie. (PubDerosa2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Autoren geben an, das ITT-Prinzip angewendet zu haben, sie berichten aber nicht, wie mit fehlenden Daten umgegangen wurde. (PubDerosa2011)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es ist unklar, ob sämtliche Hypoglykämien erfasst wurden, da nur diejenigen Hypoglykämien berichtet wurden, die zum Studienabbruch führten. Im Methodenteil wird die Beurteilung der Hypoglykämien nicht beschrieben. (PubDerosa2011)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-214 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Derosa 2011b

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Derosa, 2011b

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Derosa et al 2011b (Derosa et al., 2011b)	PubDerosa2011b
Salvadeo et al 2010 (Abstract) (Salvadeo et al., 2010)	PubSalvadeo2010

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubDerosa2011b)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine single-blind Studie. (PubDerosa2011b)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren beschrieben die Studie als „single blind“ ohne zu erläutern wer verblindet war. (PubDerosa2011b)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte

erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- intransparenter Patientenfluss*
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Änderung abweichend vom Ausgangswert in Prozent HbA1c nach 52 Wochen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die

Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Die Autoren beschrieben die Studie als „single blind“ ohne zu erläutern wer verblindet war. (PubDerosa2011b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Autoren berichten dass ein ITT-Ansatz Verwendung fand. Es geht jedoch nicht hervor wie fehlende Daten gehandhabt wurden. (PubDerosa2011b)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Änderung abweichend vom Ausgangswert des Gesamtkörpergewichts (kg) nach 52 Wochen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren beschrieben die Studie als „single blind“ ohne zu erläutern wer verblindet war. (PubDerosa2011b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und

Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Autoren berichten dass ein ITT-Ansatz Verwendung fand. Es geht jedoch nicht hervor wie fehlende Daten gehandhabt wurden. (PubDerosa2011b)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

5.

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbrüche aufgrund unerwünschter Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Die Autoren beschrieben die Studie als „single blind“ ohne zu erläutern wer verblindet war. (PubDerosa2011b)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Autoren berichten dass ein ITT-Ansatz Verwendung fand. Es geht jedoch nicht hervor wie fehlende Daten gehandhabt wurden. (PubDerosa2011b)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-215 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Yang et al., 2011

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Yang, 2011

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Yang et al., 2011 (Yang et al., 2011)	PubYang2011

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubYang2011)

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es handelt sich um eine randomisierte Studie. (PubYang2011)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Angaben zur Verdeckung der Gruppenzuteilung fehlen. (PubYang2011)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- intransparenter Patientenfluss*
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*

- *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung Gewichts (kg) vom Ausgangswert bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- unplausible Angaben*
- Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung des systolischen Blutdrucks (mmHg) vom Ausgangswert bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs,

Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Diese Information wurde für diesen Endpunkt nicht berichtet.
(PubYang2011)

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und

Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Autoren bezeichnen ihre Studie als doppelblind und single-dummy, ohne darauf einzugehen wer verblindet worden ist. (PubYang2011)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr

unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-216 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Pan et al., 2012

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Pan, 2012

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Pan et al 2012 (Pan et al., 2012)	PubPan2012

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubPan2012)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind

adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es finden sich keine Angaben zur Verblindung (PubPan2012)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es finden sich keine Angaben zur Verblindung. (PubPan2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung Gewichts (kg) vom Ausgangswert bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubPan2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine ausreichende Darstellung mit dem Umgang von fehlenden Werten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubPan2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei

einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubPan2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubPan2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubPan2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-217 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von DeFronzo., 2012

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: DeFronzo 2012

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
DeFronzo et al 2012 (DeFronzo et al., 2012)	PubDeFronzo2012

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubDeFronzo2012)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine Angaben in der Publikation (PubDeFronzo2012).

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine Informationen in der Publikation. (PubDeFronzo2012)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind

adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es finden sich keine Angaben zur Methodik der Verblindung.
(PubDeFronzo2012)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es finden sich keine Angaben zur Methodik der Verblindung. (PubDeFronzo2012)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte

erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 26)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubDeFronzo2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung Gewichts (kg) vom Ausgangswert bis Woche 26**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubDeFronzo2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 26

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubDeFronzo2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei

einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubDeFronzo2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubDeFronzo2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubDeFronzo2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-218 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Gallwitz., 2012

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Gallwitz 2012

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Gallwitz et al 2012 (inklusive des supplement) (Gallwitz et al., 2012)	PubGallwitz2012

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunkübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubGallwitz2012)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind

adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden

muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- intransparenter Patientenfluss*
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%) vom Ausgangswert bis Woche 52)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung Gewichts (kg) vom Ausgangswert bis Woche 52**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 104

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei

einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse (104 Wochen)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubGallwitz2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (104 Wochen)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubGallwitz2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse (104 Wochen)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung (DeFronzo, 2012).

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-219 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie von Yang., 2012

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Studie: Yang 2012

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Yang et al 2012 (inklusive 2 supplements) (Yang et al., 2012)	PubYang2012

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunkübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie. (PubYang2012)

1.**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine Angaben in der Publikation. (PubYang2012)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine Informationen in der Publikation. (PubYang2012)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind

adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es finden sich keine Angaben zur Methodik der Verblindung.
(PubYang2012)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es finden sich keine Angaben zur Methodik der Verblindung. (PubYang2012)

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte

erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- intransparenter Patientenfluss*
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Ausmaß der Blutzuckerkontrolle (HbA1c (%)) vom Ausgangswert bis Woche 24)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es finden sich keine Angaben zur Verblindung. (PubYang2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung Gewichts (kg) vom Ausgangswert bis Woche 24**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Es finden sich keine Angaben zur Verblindung. (PubYang2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil an Patienten mit Hypoglykämien bis Woche 24

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubYang2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei

einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubYang2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubYang2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Es finden sich keine Angaben zur Verblindung. (PubYang2012)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
