

Dossier zur Nutzenbewertung gemäß § 35a SGB V

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

Bristol-Myers Squibb GmbH & Co.KGaA und
AstraZeneca GmbH

Modul 4 C

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

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	5
Abbildungsverzeichnis	13
Abkürzungsverzeichnis	15
4 Modul 4 – allgemeine Informationen	18
4.1 Zusammenfassung der Inhalte von Modul 4.....	19
4.2 Methodik.....	32
4.2.1 Fragestellung.....	32
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	34
4.2.3 Informationsbeschaffung.....	37
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	38
4.2.3.2 Bibliografische Literaturrecherche.....	38
4.2.3.3 Suche in Studienregistern.....	40
4.2.3.4 Selektion relevanter Studien.....	41
4.2.4 Bewertung der Aussagekraft der Nachweise.....	41
4.2.5 Informationssynthese und -analyse.....	43
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	43
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	44
4.2.5.3 Meta-Analysen.....	48
4.2.5.4 Sensitivitätsanalysen.....	51
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	51
4.2.5.6 Indirekte Vergleiche.....	53
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	55
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	56
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	56
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	56
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	59
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	61
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	61
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	62
4.3.1.2.1 Studiendesign und Studienpopulationen.....	62
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	66
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	67
4.3.1.3.1 Ausmaß der Blutzuckerkontrolle (HbA1c) – RCT.....	68
4.3.1.3.1.1 Hypoglykämien – RCT.....	75
4.3.1.3.1.2 Gewichtsveränderung – RCT.....	79
4.3.1.3.1.3 Therapiezufriedenheit (DTSQ) – RCT.....	85
4.3.1.3.1.4 Blutdrucksenkung – RCT.....	89
4.3.1.3.1.5 Unerwünschte Ereignisse – RCT.....	96

4.3.1.3.1.6	Kardiale Ereignisse – RCT	102
4.3.1.3.2	Subgruppenanalysen – RCT.....	104
4.3.1.3.2.1	Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach dem Geschlecht– RCT.....	107
4.3.1.3.2.2	Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach dem Alter – RCT	111
4.3.1.3.2.3	Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach asiatischer bzw. nicht-asiatischer Abstammung - RCT.....	115
4.3.1.3.2.4	Hypoglykämien stratifiziert nach dem Geschlecht – RCT	118
4.3.1.3.2.5	Hypoglykämien stratifiziert nach dem Alter– RCT	121
4.3.1.3.2.6	Hypoglykämien stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT	123
4.3.1.3.2.7	Gewichtsveränderung stratifiziert nach dem Geschlecht– RCT....	125
4.3.1.3.2.8	Gewichtsveränderung stratifiziert nach dem Alter– RCT	129
4.3.1.3.2.9	Gewichtsveränderung stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT	132
4.3.1.3.2.10	Therapiezufriedenheit stratifiziert nach dem Geschlecht – RCT ..	135
4.3.1.3.2.11	Therapiezufriedenheit stratifiziert nach dem Alter – RCT	138
4.3.1.3.2.12	Therapiezufriedenheit stratifiziert nach asiatischer und nicht nicht-asiatischer Abstammung – RCT.....	141
4.3.1.3.2.13	Blutdrucksenkung stratifiziert nach dem Geschlecht – RCT	144
4.3.1.3.2.14	Blutdrucksenkung stratifiziert nach dem Alter – RCT	149
4.3.1.3.2.15	Blutdrucksenkung stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT	154
4.3.1.3.2.16	Unerwünschte Ereignisse stratifiziert nach dem Geschlecht – RCT	159
4.3.1.3.2.17	Unerwünschte Ereignisse stratifiziert nach dem Alter– RCT	163
4.3.1.3.2.18	Unerwünschte Ereignisse stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT	165
4.3.1.3.2.19	Kardiale Ereignisse stratifiziert nach dem Geschlecht – RCT	168
4.3.1.3.2.20	Kardiale Ereignisse stratifiziert nach dem Alter – RCT	169
4.3.1.3.2.21	Kardiale Ereignisse stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT	170
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	171
4.3.2	Weitere Unterlagen.....	175
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	175
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	176
4.3.2.1.1.1	Studien des pharmazeutischen Unternehmers	176
4.3.2.1.1.2	Studien aus der bibliografischen Literaturrecherche	179
4.3.2.1.1.3	Studien aus der Suche in Studienregistern	186
4.3.2.1.1.4	Resultierender Studienpool: RCT für indirekte Vergleiche	187
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	189
4.3.2.1.2.1	Studiendesign und Studienpopulationen	189
4.3.2.1.2.2	Verzerrungspotenzial auf Studienebene	201
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	202

4.3.2.1.3.1	Ausmaß der Blutzuckerkontrolle (HbA1c) – indirekte Vergleiche aus RCT.....	204
4.3.2.1.3.1.1	Hypoglykämien – indirekte Vergleiche aus RCT.....	212
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.....	233
4.3.2.1.3.1.5	Unerwünschte Ereignisse – indirekte Vergleiche aus RCT....	239
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	254
4.3.2.2	Nicht randomisierte vergleichende Studien.....	254
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	255
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	255
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	256
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	256
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	257
4.3.2.3	Weitere Untersuchungen.....	257
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	257
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	257
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	258
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	258
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	258
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	258
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	262
4.4.1	Beurteilung der Aussagekraft der Nachweise	262
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	263
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	272
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	273
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	274
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	274
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	274
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	275
4.5.4	Verwendung von Surrogatendpunkten	275
4.6	Liste der eingeschlossenen Studien.....	278
4.7	Referenzliste.....	279
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		285
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		300
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund		305
Anhang 4-D : Liste der abgebrochenen Studien		311
Anhang 4-E : Liste der laufenden Studien.....		312
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....		315

Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 364

Tabellenverzeichnis

	Seite
Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen	22
Tabelle 4-2: Zusammenfassende Darstellung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit.....	30
Tabelle 4-3: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen	37
Tabelle 4-4: Methodik zur gemeinsame Betrachtung der Endpunkte Blutzuckerkontrolle (HbA1c) und Hypoglykämien.....	46
Tabelle 4-5: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	57
Tabelle 4-6: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	59
Tabelle 4-7: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	61
Tabelle 4-8: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	62
Tabelle 4-9: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	64
Tabelle 4-10: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	65
Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	65
Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	65
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	66
Tabelle 4-14: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	67
Tabelle 4-15: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	67
Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	68
Tabelle 4-17: Operationalisierung von Ausmaß der Blutzuckerkontrolle (HbA1c)	69
Tabelle 4-18: Bewertung des Verzerrungspotenzials für „Ausmaß der Blutzuckerkontrolle (HbA1c)“ in RCT mit dem zu bewertenden Arzneimittel	70
Tabelle 4-19: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	71
Tabelle 4-20: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	72

Tabelle 4-21: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	73
Tabelle 4-22: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	74
Tabelle 4-23: Operationalisierung von Hypoglykämien	75
Tabelle 4-24: Bewertung des Verzerrungspotenzials für Hypoglykämien in RCT mit dem zu bewertenden Arzneimittel	76
Tabelle 4-25: Ergebnisse für Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation).....	77
Tabelle 4-26: Ergebnisse für Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	78
Tabelle 4-27: Operationalisierung von Gewichtsveränderung.....	80
Tabelle 4-28: Bewertung des Verzerrungspotenzials für Gewichtsveränderung in RCT mit dem zu bewertenden Arzneimittel.....	80
Tabelle 4-29: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation).....	81
Tabelle 4-30: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	82
Tabelle 4-31: Ergebnisse für Gewichtsveränderung ($\geq 3\%$ und $\geq 5\%$ Gewichtsreduktion) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	83
Tabelle 4-32: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	84
Tabelle 4-33: Operationalisierung von Therapiezufriedenheit (DTSQ)	85
Tabelle 4-34: Bewertung des Verzerrungspotenzials für Therapiezufriedenheit (DTSQ) in RCT mit dem zu bewertenden Arzneimittel	86
Tabelle 4-35: Ergebnisse für Therapiezufriedenheit (DTSQ) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)	87
Tabelle 4-36: Ergebnisse für Therapiezufriedenheit aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	88
Tabelle 4-37: Operationalisierung von Blutdrucksenkung	90
Tabelle 4-38: Bewertung des Verzerrungspotenzials für Blutdrucksenkung in RCT mit dem zu bewertenden Arzneimittel.....	90
Tabelle 4-39: Ergebnisse für Blutdrucksenkung aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation).....	92
Tabelle 4-40: Ergebnisse für Blutdrucksenkung aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	94
Tabelle 4-41: Operationalisierung von Unerwünschte Ereignissen.....	97
Tabelle 4-42: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel	97

Tabelle 4-43: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation).....	98
Tabelle 4-44: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	99
Tabelle 4-45: Ergebnisse für Unerwünschte Ereignisse – Anzeichen einer Genital- bzw. Harnwegsinfektion aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation) ...	100
Tabelle 4-46: Ergebnisse für Unerwünschte Ereignisse – Anzeichen einer Genital- bzw. Harnwegsinfektion aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	100
Tabelle 4-47: Operationalisierung von Kardiale Ereignisse	102
Tabelle 4-48: Bewertung des Verzerrungspotenzials für Kardiale Ereignisse in RCT mit dem zu bewertenden Arzneimittel.....	103
Tabelle 4-49: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel.....	103
Tabelle 4-50: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)	104
Tabelle 4-51: Interaktionstest der präspezifizierten Subgruppenanalysen für den Endpunkt Ausmaß der Blutzuckerkontrolle	105
Tabelle 4-52: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel	107
Tabelle 4-53: Ergebnisse für Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach dem Geschlecht	109
Tabelle 4-54: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel	111
Tabelle 4-55: Ergebnisse für Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach dem Alter	113
Tabelle 4-56: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel	115
Tabelle 4-57: Ergebnisse für Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach asiatischer und nicht-asiatischer Abstammung.....	117
Tabelle 4-58: Ergebnisse für Hypoglykämien stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel.....	118
Tabelle 4-59: Ergebnisse für Hypoglykämien stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel	121
Tabelle 4-60: Ergebnisse für Hypoglykämien stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel	123
Tabelle 4-61: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach dem Geschlecht	125

Tabelle 4-62: Ergebnisse für Gewichtsverlust ($\geq 5\%$) aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach dem Geschlecht	127
Tabelle 4-63: Ergebnisse für Gewichtsveränderung stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel.....	129
Tabelle 4-64: Ergebnisse für Gewichtsverlust ($\geq 5\%$) aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach dem Alter	131
Tabelle 4-65: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach asiatischer und nicht-asiatischer Abstammung.....	132
Tabelle 4-66: Ergebnisse für Gewichtsverlust ($\geq 5\%$) aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach asiatischer und nicht-asiatischer Herkunft	134
Tabelle 4-67: Ergebnisse für Therapiezufriedenheit stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel	135
Tabelle 4-68: Ergebnisse für Therapiezufriedenheit stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel.....	138
Tabelle 4-69: Ergebnisse für Therapiezufriedenheit stratifiziert nach asiatischer und nicht-asiatischer Abstammung - RCT mit dem zu bewertenden Arzneimittel.....	141
Tabelle 4-70: Ergebnisse für Blutdrucksenkung stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel.....	144
Tabelle 4-71: Ergebnisse für Blutdrucksenkung stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel	149
Tabelle 4-72: Ergebnisse für Blutdrucksenkung stratifiziert nach asiatischer und nicht-asiatischer Abstammung aus RCT mit dem zu bewertenden Arzneimittel.....	154
Tabelle 4-73: Ergebnisse für Unerwünschte Ereignisse stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel	159
Tabelle 4-74: Ergebnisse für Anzahl der Patienten mit Anzeichen einer Genital- oder Harnwegsinfektion stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel.....	161
Tabelle 4-75: Ergebnisse für Unerwünschte Ereignisse stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel	163
Tabelle 4-76: Ergebnisse für Ergebnisse für Anzahl der Patienten mit Anzeichen einer Genital- oder Harnwegsinfektion stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel	164
Tabelle 4-77: Ergebnisse für Unerwünschte Ereignisse stratifiziert nach nach asiatischer und nicht-asiatischer Abstammung - RCT mit dem zu bewertenden Arzneimittel	165
Tabelle 4-78: Ergebnisse für Ergebnisse für Anzahl der Patienten mit Anzeichen einer Genital- oder Harnwegsinfektion stratifiziert nach asiatischer und nicht-asiatischer Abstammung - RCT mit dem zu bewertenden Arzneimittel.....	167
Tabelle 4-79: Ergebnisse für Kardiale Ereignisse stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel	168
Tabelle 4-80: Ergebnisse für Kardiale Ereignisse stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel	169

Tabelle 4-81: Ergebnisse für Kardiale Ereignisse stratifiziert nach asiatischer und nicht-asiatischer Abstammung aus RCT mit dem zu bewertenden Arzneimittel.....	170
Tabelle 4-82: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	177
Tabelle 4-83: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	179
Tabelle 4-84: Relevante Studien aus der Suche in Studienregistern – RCT mit der zweckmäßigen Vergleichstherapie.....	186
Tabelle 4-85: Relevante Studien aus der Suche in Studienregistern – RCT mit der alternativen Vergleichstherapie.....	187
Tabelle 4-86: Studienpool – RCT für indirekte Vergleiche.....	188
Tabelle 4-87: Charakterisierung der eingeschlossenen Studien – Studien für indirekte Vergleiche	190
Tabelle 4-88: Charakterisierung der Interventionen – Studien für indirekte Vergleiche.....	194
Tabelle 4-89: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche	197
Tabelle 4-90: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche	198
Tabelle 4-91: Verzerrungspotenzial auf Studienebene – Studien für indirekte Vergleiche...	202
Tabelle 4-92: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	203
Tabelle 4-93: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	204
Tabelle 4-94: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden.....	205
Tabelle 4-95: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden	205
Tabelle 4-96: Operationalisierung von Ausmaß der Blutzuckerkontrolle (HbA1c).....	206
Tabelle 4-97: Bewertung des Verzerrungspotenzials für Ausmaß der Blutzuckerkontrolle (HbA1c) in RCT für indirekte Vergleiche	207
Tabelle 4-98: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT für indirekte Vergleiche	208
Tabelle 4-99: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) - indirekte Vergleiche	210
Tabelle 4-100: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) - indirekte Vergleiche	212
Tabelle 4-101: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden.....	213

Tabelle 4-102: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden	213
Tabelle 4-103: Operationalisierung von Hypoglykämien.....	214
Tabelle 4-104: Bewertung des Verzerrungspotenzials für Hypoglykämien in RCT für indirekte Vergleiche	215
Tabelle 4-105: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche.....	217
Tabelle 4-106: Ergebnisse für Hypoglykämien aus indirekte Vergleiche	220
Tabelle 4-107: Ergebnisse für Gesamtrate der Patienten mit mindestens einer Hypoglykämie - indirekte Vergleiche	222
Tabelle 4-108: Ergebnisse für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel – indirekter Vergleich	223
Tabelle 4-109: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden.....	224
Tabelle 4-110: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden	225
Tabelle 4-111: Operationalisierung von Gewichtsveränderung.....	226
Tabelle 4-112: Bewertung des Verzerrungspotenzials für Gewichtsveränderung in RCT für indirekte Vergleiche	227
Tabelle 4-113: Ergebnisse für Gewichtsveränderung aus RCT für indirekte Vergleiche.....	228
Tabelle 4-114: Ergebnisse für Gewichtsveränderung aus indirekte Vergleiche	230
Tabelle 4-115: Ergebnisse für Gewichtsveränderung - indirekte Vergleiche	233
Tabelle 4-116: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden.....	234
Tabelle 4-117: Operationalisierung von Blutdrucksenkung	235
Tabelle 4-118: Bewertung des Verzerrungspotenzials für Blutdrucksenkung in RCT für indirekte Vergleiche	235
Tabelle 4-119: Ergebnisse für Blutdrucksenkung aus RCT für indirekte Vergleiche	236
Tabelle 4-120: Ergebnisse für Blutdrucksenkung - indirekte Vergleiche.....	238
Tabelle 4-121: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden.....	239
Tabelle 4-122 Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden	240
Tabelle 4-123: Operationalisierung von Unerwünschte Ereignisse.....	241

Tabelle 4-124: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT für indirekte Vergleiche.....	243
Tabelle 4-125: Ergebnisse für Unerwünschte Ereignisse aus RCT für indirekte Vergleiche	245
Tabelle 4-126: Ergebnisse für Gesamtrate unerwünschter Ereignisse – indirekte Vergleiche	249
Tabelle 4-127: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse – indirekte Vergleiche	250
Tabelle 4-128: Ergebnisse der Studienabbrüche aufgrund unerwünschter Ereignisse - indirekte Vergleiche	250
Tabelle 4-129: Ergebnisse für Gesamtrate unerwünschter Ereignisse - indirekte Vergleiche	253
Tabelle 4-130: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse - indirekte Vergleiche	253
Tabelle 4-131: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse - indirekte Vergleiche	254
Tabelle 4-132: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	255
Tabelle 4-133: Operationalisierung von <Endpunkt xxx>.....	256
Tabelle 4-134: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	256
Tabelle 4-135: Operationalisierung von <Endpunkt xxx>.....	258
Tabelle 4-136: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	273
Tabelle 4-137: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	274
Tabelle 4-138 (Anhang): DIMDI: Suchstrategie MEDLINE 1950 bis heute: am 23. Oktober 2012.....	286
Tabelle 4-139 (Anhang): DIMDI: Suchstrategie EMBASE 1947 bis heute: am 23. Oktober 2012.....	288
Tabelle 4-140 (Anhang): DIMDI: Suchstrategie Cochrane 1948 bis heute: am 23. Oktober 2012.....	289
Tabelle 4-141 (Anhang): DIMDI: Suchstrategie MEDLINE 1950 bis heute: am 23. Oktober 2012.....	290
Tabelle 4-142 (Anhang): DIMDI: Suchstrategie EMBASE 1947 bis heute: am 23. Oktober 2012.....	292
Tabelle 4-143 (Anhang): DIMDI: Suchstrategie Cochrane 1948 bis heute: am 23. Oktober 2012.....	294
Tabelle 4-144 (Anhang): DIMDI: Suchstrategie MEDLINE 1946 bis heute: am 17. Oktober 2012.....	296

Tabelle 4-145 (Anhang): DIMDI: Suchstrategie EMBASE 1947 bis heute: am 17. Oktober 2012.....	297
Tabelle 4-146 (Anhang): DIMDI: Suchstrategie Cochrane 1948 bis heute: am 17. Oktober 2012.....	298
Tabelle 4-147 (Anhang): Im Volltext ausgeschlossene nicht relevante Publikationen.....	305
Tabelle 4-148 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel.....	311
Tabelle 4-149 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	313
Tabelle 4-150 (Anhang): Studiendesign und -methodik für Studie D1690C00005.....	316
Tabelle 4-151 (Anhang): Studiendesign und -methodik für Studie DeFronzo et al. 1995	326
Tabelle 4-152 (Anhang): Studiendesign und –methodik für Studie Goldstein et al. 2003	334
Tabelle 4-153 (Anhang): Studiendesign und -methodik für Studie Hermansen et al. 2007 ..	340
Tabelle 4-154 (Anhang): Studiendesign und -methodik für Studie Chacra et al. 2009	345
Tabelle 4-155 (Anhang): Studiendesign und -methodik für Studie Garber et al. 2008	350
Tabelle 4-156 (Anhang): Studiendesign und -methodik für Studie Pratley et al. 2009	355
Tabelle 4-157 (Anhang): Studiendesign und -methodik für Studie Lewin et al. 2012	360
Tabelle 4-158 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D1690C00005	365
Tabelle 4-159 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Publikation von DeFronzo 1995.....	390
Tabelle 4-160 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Goldstein 2003.....	408
Tabelle 4-161 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Hermansen K et al. 2007	426
Tabelle 4-162 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Chacra AR et al. 2009.....	445
Tabelle 4-163 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Garber AJ et al. 2008.....	464
Tabelle 4-164 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Pratley et al. 2009	483
Tabelle 4-165 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Lewin et al. 2012	502

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	60
Abbildung 2: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	74
Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	79
Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	84
Abbildung 5: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	89
Abbildung 6: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	96
Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	101
Abbildung 8: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	104
Abbildung 9: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie Metformin in Kombination mit Glimепирид	181
Abbildung 10: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie Metformin in Kombination mit Glibenclamid oder Glipizid	183
Abbildung 11: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der alternativen Vergleichstherapie DPP-4-Inhibitor in Kombination mit Sulfonylharnstoff	185
Abbildung 12: Meta-Analyse der eingeschlossenen Studie zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Ausmaß der Blutzuckerkontrolle“	211
Abbildung 13: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate der Patienten mit mindestens einer Hypoglykämie“	221
Abbildung 14: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gewichtsveränderung“	231
Abbildung 15: Sensitivitätsanalyse: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor (nur Studien mit Angabe von Streuungsmaßen) vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gewichtsveränderung“	232

Abbildung 16: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate unerwünschter Ereignisse“	251
Abbildung 17: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate schwerwiegender unerwünschter Ereignisse“	252
Abbildung 18: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Studienabbrüche aufgrund unerwünschter Ereignisse“	252
Abbildung 19: Flussdiagramm Studiendesign und -methodik für Studie D1690C00005	325
Abbildung 20: Flussdiagramm Studiendesign und -methodik für Studie DeFronzo 1995 (Protokoll 1)	332
Abbildung 21: Flussdiagramm Studiendesign und –methodik für DeFronzo 1995 (Protokoll 2)	333
Abbildung 22: Flussdiagramm Studiendesign und-methodik für Goldstein 2003.....	339
Abbildung 23: Flussdiagramm Studiendesign und-methodik für Hermansen 2007	344
Abbildung 24: Flussdiagramm Studiendesign und-methodik für Chacra 2009	349
Abbildung 25: Flussdiagramm Studiendesign und-methodik für Garber 2008	354
Abbildung 26: Flussdiagramm Studiendesign und-methodik für Pratley 2008	359
Abbildung 27: Flussdiagramm Studiendesign und-methodik für Lewin 2012	363

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ACCORD	Action to control Cardiovascular Risk in Diabetes
ADVANCE	Action in Diabetes and Vascular Disease
AkdÄ	Arzneimittelkommission der deutschen Ärzteschaft
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
AST	Aspartate aminotransferase
AUC	Area under the curve
BMI	Body-Mass-Index
BMS/AZ	Bristol-Myers Squibb / AstraZeneca
CONSORT	Consolidated Standards of Reporting Trials
DDG	Deutsche Diabetes-Gesellschaft
DIMDI	Deutsches Institut für Medizinische Dokumentation
DMP	Disease Management Programm
DPP-4	Dipeptidyl-Peptidase-4
DTSQ	Diabetes Treatment Satisfaction Questionnaire
EKG	Elektrokardiogramm
FBG	Nüchternblutzucker (Fasting Blood Glucose)
FEM	Fixed Effect Model
FFS	Freie Fettsäure
FPG	Nüchtern Plasmablutzucker (Fasting plasma glucose)
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
eGFR	Geschätzte Glomeruläre Filtrationsrate
GLP-1	Glucagon-like peptide-1
HbA1c	Glykiertes Hämoglobin
HDL	High density lipoprotein
HIV	Human immunodeficiency virus
HOMA	Homeostasis Model Assessment
HR	Hazard Ratio
ICTRP	International Clinical Trials Registry Platform

Abkürzung	Bedeutung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
IVRS	Interactive voice response system
KI	Konfidenzintervall
LOCF	Last observation carried forward
MeDRA	Medical Dictionary for Regulatory Activities
MCID	Minimal clinical important difference
MDG	Mean daily glucose
MWD	Mittelwertdifferenz
NBZ	Nüchternblutzucker
NIDDM	Non-insulin-dependent diabetes mellitus
OAD	Orales Antidiabetikum
OGTT	Oraler Glukose-Toleranztest
OR	Odds ratio
PBG	Postprandialer Blutzucker
PPI	Postprandiales Insulin
QTc	QT-Intervall, korrigiert gegen die Herzrate
RCT	Randomized Controlled Trial
REM	Random Effect Model
RR	Relatives Risiko
SAS	Statistical Analysis Software
SD	Standardabweichung
SE	Standardfehler
SGB	Sozialgesetzbuch
SGLT	Sodium Dependent Glucose Co-Transporter
SOC	System Organ Class
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SU	Sulfonylurea
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschte Ereignisse

Abkürzung	Bedeutung
UKPDS	United Kingdom Prospective Diabetes Study
ULN	Upper limit of the normal range
WHO	World Health Organization
XR	Extended release
C-Peptid	Connecting peptide

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 4.1)
- 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 C stellt den medizinischen Zusatznutzen von Dapagliflozin (Forxiga®) in der Zweifachkombination mit dem oralen Antidiabetikum (OAD) Sulfonylharnstoff (Anwendungsgebiet *Add-on* zu Sulfonylharnstoff) gegenüber der zweckmäßigen Vergleichstherapie, der Kombination aus Metformin und einem Sulfonylharnstoff sowie gegenüber der alternativen Vergleichstherapie Sulfonylharnstoff und einem Dipeptidyl-Peptidase-4 (DPP-4)-Inhibitor, dar.

Somit ergeben sich zwei Fragestellungen:

Primäre Fragestellung:

Ist für Sulfonylharnstoff+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 begleitende Therapie mit Schleifendiuretika, deren Blutzucker mit einem Sulfonylharnstoff allein zusammen mit Diät und Bewegung nicht ausreichend kontrolliert ist, basierend auf den Daten aus randomisierten kontrollierten Studien (RCT), ein medizinischer Zusatznutzen im Sinne des § 35a SGB V gegenüber der zweckmäßigen Vergleichstherapie nachweisbar?

Sekundäre Fragestellung:

Ist für 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 einem Sulfonylharnstoff allein zusammen mit Diät und Bewegung nicht ausreichend kontrolliert ist, und für die eine Behandlung mit Metformin nicht geeignet ist und die Anwendung von Insulin noch nicht angezeigt ist, basierend auf den Daten aus RCT, ein medizinischer Zusatznutzen gegenüber der alternativen Vergleichstherapie nachweisbar?

Der Bewertung des Zusatznutzens vorangestellt ist die Darstellung des medizinischen Nutzens von Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Placebo mittels der Methodik der evidenzbasierten Medizin und mittels patientenrelevanter Endpunkte. Dapagliflozin wurde im Rahmen der Erteilung der Zulassung bereits ein positives Nutzen-Risiko-Verhältnis auf Basis der in den Studien prädefinierten primären Endpunkte bestätigt.

Für die Bewertung des medizinischen Nutzens in der beschriebenen Indikation sind folgende patientenrelevante Endpunkte etabliert und wurden daher für die Zusatznutzenbewertung von Dapagliflozin, soweit 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 wurden 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 ein zusätzlicher Risikofaktor für kardiovaskuläre Erkrankungen ist, 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 (bzw. alternativen) Vergleichstherapie wurden anhand einer umfassenden Informationsbeschaffung identifiziert (Abschnitt 4.3.1.1 sowie 4.3.2.1.1).

Der Nachweis des **medizinischen Nutzens** von Sulfonylharnstoff+Dapagliflozin im Vergleich zu Sulfonylharnstoff+Placebo basiert auf einer doppelblinden RCT von Bristol-Myers Squibb / AstraZeneca (BMS / AZ) (Studie D1690C00005). Für den Zusatznutzen liegen im Anwendungsgebiet *Add-on* zu Sulfonylharnstoff keine Studien vor, die einen

direkten Vergleich zwischen dem zu bewertenden Arzneimittel und der zweckmäßigen Vergleichstherapie bzw. alternativen Vergleichstherapie ermöglichen. Deshalb wurden zur Bewertung des medizinischen Zusatznutzens von Dapagliflozin indirekte Vergleiche herangezogen.

Zur Beurteilung der klinischen Sicherheit und Wirksamkeit der ausgewählten zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Metformin), sowie der alternativen Vergleichstherapie (Sulfonylharnstoff+DPP-4-Inhibitor) im Vergleich mit dem zu bewertenden Arzneimittel (Sulfonylharnstoff+Dapagliflozin) wurde eine bibliografische Literaturrecherche in den Datenbanken MEDLINE, EMBASE und Cochrane unter der Fragestellung „Identifikation von Studien (RCT) zur Sicherheit und Wirksamkeit der Kombinationsbehandlung mit Metformin und Sulfonylharnstoffen bzw. DPP-4-Inhibitoren und Sulfonylharnstoffen zur Verbesserung der Blutzuckerkontrolle von Typ-2-Diabetikern“ durchgeführt. Ziel dieser Literaturrecherche war es, RCT mit der zweckmäßigen Vergleichstherapie (bzw. alternativen Vergleichstherapie) zu identifizieren, bei denen die Kontrollgruppe mit einem Sulfonylharnstoff und Placebo behandelt wurde.

Sowohl zum zu bewertenden Arzneimittel als auch zur zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Metformin) bzw. zur Kombinationstherapie Sulfonylharnstoff+DPP-4-Inhibitor wurde eine Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien durchgeführt. Die Suche erfolgte in den öffentlich verfügbaren Studienregistern *clinicaltrials.gov* sowie dem *International Clinical Trials Registry Platform Search Portal* (ICTRP Search Portal).

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

Ein-/Ausschlusskriterien für Studien

Die Ein- und Ausschlusskriterien der Studien zum Beleg des Nutzens und des Zusatznutzens der Behandlung mit Dapagliflozin in Kombination mit einem Sulfonylharnstoff 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 Sulfonylharnstoff-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker aufweisen. ¹	Patienten mit anderen Diabetes-Typen und Behandlungs-naive Patienten.
Intervention	Behandlung mit Dapagliflozin (5 mg oder 10 mg einmal täglich) + Sulfonylharnstoff ²	
Vergleichstherapie	Kombinationsbehandlung mit Metformin bzw. einem DPP-4-Inhibitor+Sulfonylharnstoff (für die Bewertung des Zusatznutzens) oder Sulfonylharnstoff +Placebo (für die Bewertung des Nutzens)	Andere Kombinationstherapien oder Monotherapien
Endpunkte	Mindestens einer der oben genannten patientenrelevanten Endpunkte muss berücksichtigt werden (HbA1c nur in Kombination mit Hypoglykämien)	
Studientypen	RCT	Nicht randomisierte Studien, Review-Artikel <i>Case Reports</i>
Studiendauer	≥24 Wochen	<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, XR: Extended release

¹ Unter besonderer Berücksichtigung von Patienten mit einer GFR ≥ 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, Glipizid, Glibenclamid

³ Bei geringer Evidenzlage und zwecks Durchführung indirekter Vergleiche können Studien mit einer Studiendauer von ≥ 18 Wochen eingeschlossen werden. Bei RCT mit *cross-over* Design, muss jede Periode dieses Kriterium erfüllen

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

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

Die für die Fragestellung relevanten Studien wurden hinsichtlich ihrer Aussagekraft bewertet. Basis der Bewertung waren Informationen aus klinischen Studienberichten (falls vorhanden) oder aus Volltextpublikationen der jeweiligen klinischen Studien.

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

Für randomisierte Studien wurde gemäß Verfahrensordnung 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 würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führte nicht zum Ausschluss aus der Bewertung. Die Klassifizierung diente vielmehr der Diskussion heterogener Studienergebnisse und beeinflusste die Sicherheit der Aussage.

Für die Darstellung des medizinischen Nutzens wurden keine Meta-Analysen durchgeführt, da für die Fragestellung nur eine relevante RCT von BMS / AZ existiert. Die Studien mit der zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Metformin) konnten aufgrund klinischer Heterogenität nicht zusammengefasst werden, während sich die Studien zur alternativen Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor für eine Metaanalyse eigneten.

Für dichotome Endpunkte wurde der Therapieeffekt in der Regel anhand des *odds ratio* (OR) zusammengefasst. Wenn in der Publikation oder dem Studienbericht keine *odds ratios* berichtet wurden, so wurden diese inkl. exakter 95% Konfidenzintervalle unter Verwendung der SAS-Software (SAS 9.1; SAS Institute Inc., Cary, NC, USA) selbst berechnet. Sofern eine Zelle in der Vierfeldertafel eine 0 aufwies, wurden der Schätzer und die Konfidenzintervalle unter Verwendung der sog. Woolf-Haldane Korrektur (Addition von 0.5 zu jeder Zelle der Vierfeldertafel) geschätzt. Wenn zwei oder mehr Zellen eine 0 aufwiesen, wurden keine Schätzer oder Konfidenzintervalle berechnet. Im Falle nicht hinreichend homogener Studienresultate oder 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.

Zur Abschätzung der klinischen Relevanz wurden, 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 – im Falle einer Meta-Analyse - 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.

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.

Da zur Bewertung des medizinischen Zusatznutzens des zu bewertenden Arzneimittels (Sulfonylharnstoff+Dapagliflozin) gegenüber der zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Metformin) bzw. gegenüber der alternativen Vergleichstherapie (Sulfonylharnstoff+DPP-4-Inhibitor) keine direkte *Head-to-Head*-Studie existiert, wurden adjustierte indirekte Vergleiche betrachtet. Die methodische Basis eines adjustierten indirekten Vergleiches von verschiedenen Behandlungsmöglichkeiten bildet ein gemeinsamer Vergleichsarm, anhand dessen sich die Ergebnisse zur Wirksamkeit und Sicherheit beider Interventionen quantifizieren lassen.

Der einfache adjustierte indirekte Vergleich wurde mittels eines frequentistischen Ansatzes durchgeführt. Er beruhte auf einem einfachen Netzwerk und wurde für Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin bzw. Sulfonylharnstoff+DPP-4-Inhibitor über den Brückenkomparator Sulfonylharnstoff+Placebo durchgeführt.

Grundlage für diesen Vergleich waren folgende Beziehungen der Effektmaße und ihrer Varianz:

$$- \ln \theta_{BC} = \ln \theta_{AB} - \ln \theta_{AC} \quad (1)$$

$$- \text{Var}(\ln \theta_{BC}) = \text{Var}(\ln \theta_{AB}) + \text{Var}(\ln \theta_{AC}) \quad (2)$$

- Das 95% Konfidenzintervall für $\ln \theta_{BC}$ berechnet sich als $\ln \theta_{BC} \pm 1.96 * \sqrt{\text{Var}(\ln \theta_{BC})}$ (3)

Effektmaße, die als θ in den Formeln (1) bis (3) geeignet sind, sind das RR, das OR, sowie das *Hazard Ratio* (HR). Für kontinuierliche, normalverteilte Variablen entfällt die Log-Transformation vor Berechnung des indirekten Effektschätzers und seiner Streuung. Grundannahme für die Gültigkeit dieser Zusammenhänge ist die Abwesenheit bedeutsamer Heterogenität zwischen den eingeschlossenen Studien (methodische Faktoren, Effektmodifikatoren).

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

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

Im Folgenden werden zunächst die Ergebnisse der berichteten Zielgrößen zum medizinischen Nutzen und dann im Anschluss die Ergebnisse zum Zusatznutzen dargestellt.

Die in der folgenden Beschreibung des Nutzens und des Zusatznutzens betrachteten Ergebnisse beziehen sich primär auf die für Dapagliflozin empfohlene Hauptdosis von 10 mg (in Ergänzung zur Anfangsdosis 5 mg für die Beschreibung des Nutzens) und die Zielpopulation gemäß Zulassung (Typ-2-Diabetiker, die unter einer Sulfonylharnstoff-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. Kreatininclarence <60 mL/min sowie von Patienten im Alter von ≥ 75 Jahren oder Patienten, die mit einem Schleifendiuretikum vorbehandelt waren).

Obwohl im Rahmen der Erteilung der Zulassung bereits bestätigt wurde, dass die Kombination von Dapagliflozin mit einem Sulfonylharnstoff bei der Behandlung von Typ-2-Diabetikern ein positives Nutzen-Risiko-Verhältnis aufweist, wurde der medizinische Nutzen in diesem Dossier erneut, jedoch mittels der Methodik der evidenzbasierten Medizin im Hinblick auf patientenrelevante Endpunkte, dargestellt.

Medizinischer Nutzen

Die **Ergebnisse zum Nutzen** des zu bewertenden Arzneimittels Dapagliflozin in Kombination mit Sulfonylharnstoff basieren auf einer RCT der Phase III (D1690C00005), die mit der freien Kombination von Glimepirid+Dapagliflozin durchgeführt wurde. Da hier nur eine relevante RCT vorliegt, wurde keine Meta-Analyse durchgeführt.

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 Monotherapie mit einem OAD nicht erreicht werden, wird

die Kombinationstherapie mit einem anderen 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. Eine optimale Blutzuckereinstellung stellt somit ein wichtiges Therapieziel dar. In der *United Kingdom Prospective Diabetes Study* (UKPDS), einer großen klinischen Studie, 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. Eine optimierte Blutzuckerkontrolle soll allerdings unter Vermeidung der Nebenwirkungen einer medikamentösen Therapie, insbesondere Hypoglykämien und Gewichtszunahmen, erreicht werden.

Für das **Ausmaß der Blutzuckerkontrolle**, gemessen am HbA1c, konnte gezeigt werden, dass die Behandlung mit Sulfonylharnstoff+Dapagliflozin sowohl nach 24 als auch nach 48 Wochen einen klinisch relevanten und statistisch signifikanten Unterschied im Vergleich zu Sulfonylharnstoff+Placebo (MWD [95%-KI]: -0,74% [-0,92;-0,55] / -0,77% [-1,02;-0,52]) bewirkt. Studienabbrüche aufgrund einer unzureichenden glykämischen Kontrolle, oder Patienten, die aufgrund einer unzureichenden glykämischen Kontrolle eine Notfallmedikation erhielten, gab es in der Sulfonylharnstoff+Placebo-Gruppe nach 24 als auch nach 48 Wochen signifikant häufiger (OR [95%-KI]: 0,07 [0,02;0,30] / 0,18 [0,10;0,33]) im Vergleich zur Sulfonylharnstoff+Dapagliflozin-Gruppe.

Hinsichtlich der Gesamtrate der **Hypoglykämien** (Anzahl der Patienten mit mindestens einer Hypoglykämie) gab es sowohl nach 24 als auch nach 48 Wochen keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen. Patienten, deren Blutzucker mit Sulfonylharnstoffen nicht mehr ausreichend kontrolliert werden kann, können somit mit Hilfe von Dapagliflozin effektiv und sicher weiter behandelt werden.

Für den Endpunkt **Gewichtsveränderung** zeigte sich sowohl nach 24 als auch nach 48 Wochen eine statistisch signifikant stärkere Gewichtsreduktion zugunsten von Sulfonylharnstoff+Dapagliflozin (MWD [95%-KI]: -1,72 kg [-2,43;-1,01]) / -1,80 kg [-2,80;-0,80]). Darüber hinaus erreichen statistisch signifikant mehr Patienten in der Sulfonylharnstoff+Dapagliflozin-Gruppe eine Gewichtsreduktion $\geq 5\%$ (OR [95%-KI]: 7,42 [2,79;19,78]) nach 24 Wochen / 3,47 [1,35;8,93] nach 48 Wochen).

Bei der Betrachtung der **Therapiezufriedenheit**, welche mithilfe des *Diabetes Treatment Satisfaction Questionnaire* (DTSQ) gemessen wurde, fällt auf, dass sich zwar beide Behandlungsgruppen in ihrer erreichten Punktzahl verbessern, d. h. die Therapiezufriedenheit steigt, die Verbesserung in der Sulfonylharnstoff +Dapagliflozin-Gruppe aber größer ist. Der Unterschied ist nicht statistisch signifikant.

Die **Blutdrucksenkung** war in der Sulfonylharnstoff+Dapagliflozin-Gruppe im Vergleich zur Sulfonylharnstoff+Placebo-Gruppe sowohl für den systolischen als auch für den diastolischen Blutdruck signifikant größer (MWD [95%-KI] systolisch: -4,4 mmHg [-7,2;-1,6] nach 24 Wochen / -4,46 mmHg [-8,62;-0,31] nach 48 Wochen / MWD [95%-KI]

diastolisch: -2,0 mmHg [-3,8;-0,2] nach 24 Wochen / -3,12 mmHg [-5,68;-0,56] nach 48 Wochen).

Für die Häufigkeit des Auftretens **unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse und Studienabbrüche aufgrund unerwünschter Ereignisse** gab es keine statistisch signifikanten Unterschiede. Auch für die Anzahl der kardialen Ereignisse war kein Unterschied zwischen den beiden Behandlungsgruppen zu sehen. Allgemein war das Auftreten dieser Ereignisse sehr selten. Ergänzend hierzu wurden Anzeichen für eine Genital- bzw. Harnwegsinfektion als unerwünschte Ereignisse von besonderem Interesse betrachtet. Für die Zielpopulation konnte zu keinem Zeitpunkt ein signifikanter Unterschied zwischen den Behandlungsgruppen (Sulfonylharnstoff+Dapagliflozin gegen Sulfonylharnstoff+Placebo) beobachtet werden.

Für keinen der betrachteten Endpunkte gab es einen Hinweis auf eine fazitrelevante Modifikation des Behandlungseffektes von Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Placebo durch die Faktoren Alter, Geschlecht oder Abstammung (asiatisch bzw. nicht-asiatisch). Daher kann man von einem über alle Gruppen hinweg konstanten Behandlungseffekt ausgehen.

Zusammenfassend kann man sagen, dass die Behandlung mit Dapagliflozin bei Patienten, deren Blutzucker mit Sulfonylharnstoffen alleine nicht mehr ausreichend kontrolliert werden kann, sicher und wirksam ist. Die Kombination von Sulfonylharnstoff+Dapagliflozin senkt den Blutzucker effektiv ohne die Gesamtrate an Hypoglykämien zu erhöhen. Darüber hinaus kommt es zu einer signifikanten Gewichtsreduktion und Blutdrucksenkung. Die Gesamtrate an unerwünschten Ereignissen ist in beiden Behandlungsgruppen ähnlich.

Medizinischer Zusatznutzen

Die zweckmäßige Vergleichstherapie, die vom Gemeinsamen Bundesausschuss (G-BA) festgelegt wurde, stellt die Kombination aus Sulfonylharnstoff und Metformin dar. Auch wenn BMS/AZ die vom G-BA bestimmte zweckmäßige Vergleichstherapie akzeptiert, gilt zu berücksichtigen, dass Dapagliflozin, zumindest zum jetzigen Zeitpunkt, in der Therapiesequenz nach Metformin zu sehen ist. Metformin ist eine Therapie, für welche es in der Monotherapie Nachweise für Überlebensvorteile gibt. Primär sehen wir daher den Einsatz von Dapagliflozin in Kombination mit Sulfonylharnstoff bei Patienten, für die Metformin nicht geeignet ist.

Die Ergebnisse zum **medizinischen Zusatznutzen** basieren auf einem adjustierten indirekten Vergleich von Sulfonylharnstoff+Dapagliflozin (aus der Studie D1690C00005) mit Studien zur zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin sowie gegenüber der alternativen Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor mit Sulfonylharnstoff+Placebo als Brückenkomparator. Die systematischen Literaturrecherchen identifizierten zwei relevante Studien mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin (Goldstein 2003 und DeFronzo 1995) und fünf relevante Studien mit dem Vergleich Sulfonylharnstoff+DPP-4-Inhibitor (Hermansen 2007, Chacra 2009, Garber 2008, Pratley 2009 und Lewin 2012). In den Sulfonylharnstoff+Metformin

Studien wurden die Sulfonylharnstoffe Glipizid (Goldstein 2003) und Glibenclamid (DeFronzo 1995) verwendet. In jeweils zwei der Studien wurde Glimperid (Hermansen et al. 2007, Garber et al. 2008) bzw. Glibenclamid (Chacra et al. 2009 und Pratley et al. 2009) als Sulfonylharnstoff benutzt. Die Studie von Lewin et al. 2012 verwendete diverse Sulfonylharnstoffe als Basistherapie. Im Gegensatz zu den Sulfonylharnstoff+DPP-4-Inhibitor Studien, waren die Sulfonylharnstoff+Metformin Studien zu heterogen, sodass kein Gesamtschätzer berechnet werden konnte. Der Vergleich Sulfonylharnstoff+Metformin war daher nur auf Grundlage der Einzelstudien möglich, während für den Vergleich Sulfonylharnstoff+DPP-4-Inhibitor sowohl ein Vergleich zur Gruppe Sulfonylharnstoff+DPP-4-Inhibitor, als auch zur Leitsubstanz Sulfonylharnstoff+Sitagliptin durchgeführt werden konnte.

Der indirekte Vergleich zu Sulfonylharnstoff+Metformin zeigt, dass das **Ausmaß der Blutzuckerkontrolle (HbA1c)** mit Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Metformin sehr ähnlich ist. Beide Behandlungen konnten den HbA1c-Wert effektiv senken. Lediglich der Vergleich mit der Studie von DeFronzo zeigte eine signifikant stärkere HbA1c-Senkung mit der zweckmäßigen Vergleichstherapie (MWD [95%-KI]: 1,14 [0,79;1,49]). Die Ergebnisse müssen aber hinterfragt werden, da der Ausgangs-HbA1c deutlich höher war in den älteren Sulfonylharnstoff+Metformin Studien und *Carry-over*-Effekte aufgrund der hohen eingesetzten Sulfonylharnstoff-Dosis nicht auszuschließen sind. Der indirekte Vergleich von Sulfonylharnstoff+Dapagliflozin mit Sulfonylharnstoff+DPP-4-Inhibitor ergab für den Endpunkt Blutzuckerkontrolle keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen (MWD [95%-KI]: -0,12 [-0,33;0,09]) bzw. Sulfonylharnstoff+Sitagliptin (MWD [95%-KI]: -0,17 [-0,48;0,14]).

Für den Endpunkt **Hypoglykämien** konnte lediglich die Publikation von DeFronzo (1995) für einen indirekten adjustierten Vergleich zwischen Sulfonylharnstoff+Dapagliflozin und Sulfonylharnstoff+Metformin herangezogen werden. Bezüglich der Gesamtrate der Patienten mit mindestens einer Hypoglykämie ergaben sich keine signifikanten Unterschiede zwischen den Behandlungsgruppen. Der indirekte Vergleich von Sulfonylharnstoff+Dapagliflozin zu Sulfonylharnstoff+DPP-4-Inhibitor ergab für den Endpunkt Hypoglykämien (Gesamtrate) keine statistisch signifikanten Unterschiede zwischen der zu bewertenden Therapie und der Sulfonylharnstoff+einem DPP-4-Inhibitor-Gruppe (OR [95%-KI]: 1,65 [0,52;5,18]) bzw. Sulfonylharnstoff+Sitagliptin (OR [95%-KI]: 0,80 [0,14;4,52]).

Gemessen an dem Ziel, eine adäquate Blutzuckerkontrolle bei gleichzeitiger Hypoglykämievermeidung zu erreichen, ergibt sich kein Hinweis auf eine Über- oder Unterlegenheit weder für Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Metformin noch für Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+DPP-4-Inhibitor. Alle Behandlungen senken den Blutzucker effektiv ohne die Gesamtrate an Hypoglykämien zu erhöhen.

Für den Endpunkt **Gewichtsveränderungen** konnte der indirekte Vergleich mit der Studie DeFronzo durchgeführt werden. Der Vergleich zeigt, dass die Patienten, die mit Sulfonylharnstoff+Dapagliflozin behandelt wurden eine statistisch signifikant höhere

Gewichtsreduktion hatten, als Patienten, die mit der zweckmäßigen Vergleichstherapie behandelt wurden (MWD [95%-KI] des indirekten Vergleiches -2,55 [-3,41;-1,69]). Ein indirekter Vergleich mit der Studie von Goldstein (2003) war aufgrund des fehlenden Streuungsmaßes nicht möglich. Die Ergebnisse der Studie zeigen jedoch, dass kein Unterschied zwischen der Behandlungs- und Placebogruppe bestand, während die Patienten aus der Studie D1690C00005, die mit Sulfonylharnstoff+Dapagliflozin behandelt wurden, signifikant mehr Gewicht verloren. Der indirekte Vergleich von Sulfonylharnstoff+Dapagliflozin zu Sulfonylharnstoff+einem DPP-4-Inhibitor im Endpunkt Gewichtsveränderung ergab statistisch signifikante Unterschiede zwischen Sulfonylharnstoff+Dapagliflozin und Sulfonylharnstoff+DPP-4-Inhibitor als Gruppe (MWD [95%-KI]: -2,36 [-3,14;-1,58]), wie auch der Kombination Sulfonylharnstoff+Sitagliptin (MWD [95%-KI]: -2,82 [-3,98;-1,66]).

Die signifikant größere Gewichtsreduktion unter Sulfonylharnstoff+Dapagliflozin trägt somit zu einer erfolgreichen Diabetestherapie bei. Damit liegt eine Verringerung nicht schwerwiegender Symptome bzw. eine relevante Vermeidung von Nebenwirkungen vor, sodass sich ein Hinweis auf einen geringen Zusatznutzen für die Behandlung mit Sulfonylharnstoff+Dapagliflozin sowohl im Vergleich zur zweckmäßigen wie auch zur alternativen Vergleichstherapie ergibt.

Der adjustierte indirekte Vergleich für den Endpunkt **Blutdrucksenkung** versus Sulfonylharnstoff+Metformin zeigte, dass die Patienten, die mit Sulfonylharnstoff+Dapagliflozin behandelt wurden, eine signifikant stärkere Senkung des systolischen Blutdrucks haben als Patienten, die mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin behandelt wurden (MWD [95%-KI]: -4,84 [-7,66;-2,02]). Die Absenkung des diastolischen Blutdrucks war ebenfalls stärker im Sulfonylharnstoff+Dapagliflozin-Arm, der Unterschied war aber statistisch nicht signifikant.

Die statistisch signifikante Absenkung des systolische Blutdruck unter Sulfonylharnstoff+Dapagliflozin stellt eine Verringerung nicht schwerwiegender Symptome dar, sodass sich ein Hinweis auf einen geringen Zusatznutzen ergibt.

Der indirekte Vergleich von Sulfonylharnstoff +Dapagliflozin mit Sulfonylharnstoff+DPP-4-Inhibitor im Endpunkt Blutdrucksenkung konnte aufgrund mangelnder Daten für die alternative Vergleichstherapie nicht durchgeführt werden.

Für die Betrachtung der **unerwünschten Ereignisse** (Gesamtrate) konnte mit beiden Studien zur zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin und mit allen Studien zur alternativen Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor ein indirekter Vergleich durchgeführt werden. Es ergaben sich keine signifikanten Unterschiede zwischen Sulfonylharnstoff+Dapagliflozin und der zweckmäßigen Vergleichstherapie bzw. der alternativen Vergleichstherapie. Der indirekte Vergleich für die Gesamtrate schwerwiegender unerwünschter Ereignisse wie auch Studienabbrüche aufgrund von unerwünschten Ereignissen ergaben ebenfalls keine signifikanten Unterschiede weder für den Vergleich

Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin noch für Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Die Ergebnisse zum Zusatznutzen zeigen, dass es für die Endpunkte „Ausmaß der Blutzuckerkontrolle“, „Hypoglykämien“, und „unerwünschte Ereignisse“ keinen Hinweis auf eine Über- oder Unterlegenheit für eine der Behandlungsoptionen (Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Metformin bzw. Sulfonylharnstoff+DPP-4-Inhibitor) gibt.

Für den Endpunkt „Gewichtsveränderungen“ konnte gezeigt werden, dass die Behandlung mit Sulfonylharnstoff+Dapagliflozin im Vergleich zu Sulfonylharnstoff+Metformin und Sulfonylharnstoff+DPP-4-Inhibitoren zu einer signifikant größeren Gewichtsreduktion führt. Im Vergleich zur zweckmäßigen Vergleichstherapie konnte darüber hinaus auch der Blutdruck signifikant gesenkt werden, was beides auf den neuartigen Wirkmechanismus von Dapagliflozin zurückzuführen ist. Gemessen an dem Ziel Gewichtszunahmen zu vermeiden und unter Berücksichtigung der Tatsache, dass eine Verminderung des Körpergewichts für die Patienten eine ähnlich wichtige Rolle spielt wie die Blutzuckereinstellung zusammen mit der Tatsache, dass die Blutdruckeinstellung besonders für Diabetiker eine wichtige Rolle in der Prävention kardiovaskulärer Ereignisse spielt, ergibt sich eine Verringerung nicht schwerwiegender Symptome bzw. eine relevante Vermeidung von Nebenwirkungen und damit insgesamt ein Hinweis auf einen geringen Zusatznutzen von Sulfonylharnstoff+Dapagliflozin.

Tabelle 4-2: Zusammenfassende Darstellung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit

Endpunkt	Quantifizierung des Zusatznutzens	Wahrscheinlichkeit
Im Vergleich zu Sulfonylharnstoff+Metformin		
Hypoglykämien unter Berücksichtigung des HbA1c	Kein Zusatznutzen	Hinweis
Gewichtsveränderung	Gering	Hinweis
Blutdrucksenkung	Gering	Hinweis
Unerwünschte Ereignisse	Kein Schaden	Hinweis
Im Vergleich zu Sulfonylharnstoff+DPP-4-Inhibitor		
Hypoglykämien unter Berücksichtigung des HbA1c	Kein Zusatznutzen	Hinweis
Gewichtsveränderung	Gering	Hinweis
Unerwünschte Ereignisse	Kein Schaden	Hinweis

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 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.

Die unter Sulfonylharnstoff+Dapagliflozin signifikant größere Gewichts- und Blutdruckreduktion ist auf den Wirkmechanismus von Dapagliflozin zurückzuführen. Dapagliflozin hemmt den Natrium-Glukose-Cotransporter 2 (*Sodium Dependent Glucose Co-Transporter 2*, SGLT-2) im proximalen Tubulus der Niere selektiv und fördert so die konzentrationsabhängige Ausscheidung von Glukose über den Urin, was nicht nur den Blutzucker senkt, sondern auch zu einem direkten Abtransport von Kalorien führt. Dies bewirkt einen Gewichtsverlust. Der Gewichtsverlust ist vor allem auf die Reduktion von Körperfettmasse – inklusive viszerale Fettmasse – zurückzuführen, welche der wichtigste Risikofaktor für die Entwicklung und das Fortschreiten der Insulinresistenz ist. Der diuretische Effekt führt zur Blutdruckreduktion.

Aufgrund dieser Tatsachen ergibt sich sowohl gegenüber Sulfonylharnstoff+Metformin als auch gegenüber Sulfonylharnstoff+DPP-4-Inhibitoren insgesamt ein **Hinweis auf einen geringen Zusatznutzen.**

4.2 Methodik

Abschnitt 4.1 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.

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 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 C stellt den medizinischen Zusatznutzen von Dapagliflozin in der Zweifachkombination mit einem oralen Antidiabetikum aus der Substanzgruppe der Sulfonylharnstoffe (*Add-on* zu Sulfonylharnstoff) dar.

Fragestellung

Die primäre Fragestellung dieser Nutzenbewertung adressiert den Zusatznutzen von Dapagliflozin gegenüber der vom G-BA festgelegten zweckmäßigen Vergleichstherapie, der Kombinationsbehandlung mit Metformin+Sulfonylharnstoff. Zusätzlich wird der medizinische Zusatznutzen von Dapagliflozin im Vergleich zu einer alternativen

Vergleichstherapie (DPP-4-Inhibitoren+Sulfonylharnstoff) beurteilt. Somit ergeben sich zwei Fragestellungen.

Primäre Fragestellung:

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

Der Bewertung des Zusatznutzens vorangestellt ist die Darstellung des medizinischen Nutzens von Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Placebo mittels der Methodik der evidenzbasierten Medizin und mittels patientenrelevanter Endpunkte. Dapagliflozin wurde im Rahmen der Erteilung der Zulassung bereits ein positives Nutzen-Risiko-Verhältnis auf Basis der in den Studien prädefinierten primären Endpunkte bestätigt (European Medicines Agency, 2012).

Vergleichstherapie

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

Für die Patienten, für die eine Kombinationsbehandlung aus Metformin mit einem Sulfonylharnstoff nicht geeignet ist und die Insulintherapie noch nicht angezeigt ist, ist die alternative Vergleichstherapie eine Kombination aus Sulfonylharnstoff und einem DPP-4-Inhibitor. Die Gruppe der DPP-4-Inhibitoren umfassen die Substanzen Sitagliptin, Vildagliptin, Linagliptin, Saxagliptin und Alogliptin. Obwohl für Alogliptin derzeit noch keine Arzneimittelzulassung vorliegt, wurde dieser DPP-4-Inhibitor in die vorliegenden Analysen aufgenommen um eine best mögliche Evidenz für die alternative Vergleichstherapie zu erhalten.

Patientenpopulation

Die primäre Fragestellung dieses Nutzendossiers betrifft die Patientenpopulation gemäß Zulassung. Dapagliflozin ist in Kombination mit einem Sulfonylharnstoff indiziert für

erwachsene Patienten mit manifestem Diabetes mellitus Typ 2, die trotz Diät und Bewegung sowie einer Sulfonylharnstoff-Monotherapie einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen. Im Rahmen des Zulassungsverfahrens 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 wird ergänzend die Gesamtpopulation der relevanten Studien betrachtet (im Folgenden Studienpopulation genannt).

Patientenrelevante Endpunkte in RCT zum Nachweis des Nutzens und Zusatznutzens

Die für die Bewertung des medizinischen Nutzens in der beschriebenen Indikation etablierten patientenrelevante Endpunkte (siehe Abschnitt 4.2.2) sollen, soweit vorhanden, für die Zusatznutzenbewertung von Dapagliflozin berücksichtigt werden.

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 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, 2012b): erwachsene Patienten mit manifestem Diabetes mellitus Typ 2, die unter einer Sulfonylharnstoff-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker aufweisen.

Intervention: Behandlung mit Dapagliflozin in der zugelassenen Wirkstärke (1 x täglich Dapagliflozin 5 mg oder 10 mg) in Kombination mit einem Sulfonylharnstoff.

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 3C, Abschnitt 3.1.1 und 3.1.2.

Für die Patienten, für die eine Kombinationsbehandlung aus Metformin mit einem Sulfonylharnstoff nicht geeignet und die Insulintherapie noch nicht angezeigt ist, ist die alternative Vergleichstherapie eine Kombination aus Sulfonylharnstoff und einem DPP-4-Inhibitor.

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 minimalen Behandlungsdauer von 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 4-3: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen

	Einschlusskriterien	Ausschlusskriterien
Patientenpopulation	Erwachsene Patienten mit manifestem Typ-2-Diabetes, die unter einer Sulfonylharnstoff-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker aufweisen. ¹	Patienten mit anderen Diabetes-Typen und Behandlungs-naive Patienten.
Intervention	Behandlung mit Dapagliflozin (5 mg oder 10 mg einmal täglich) + Sulfonylharnstoff	
Vergleichstherapie	Kombinationsbehandlung mit Metformin bzw. DPP-4-Inhibitor +Sulfonylharnstoff ² (für die Bewertung des Zusatznutzens) oder Sulfonylharnstoff +Placebo (für die Bewertung des Nutzens)	Andere Kombinationstherapien oder Monotherapien.
Endpunkte	Mindestens einer der oben genannten patientenrelevanten Endpunkte muss berücksichtigt werden (HbA1c 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 GFR ≥60 mL/min/1,73 m² oder Kreatininclearance ≥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

³ Bei geringer Evidenzlage und zwecks Durchführung indirekter Vergleiche können Studien mit einer Studiendauer von ≥18 Wochen eingeschlossen werden. Bei RCT mit cross-over Design, muss jede Periode dieses Kriterium erfüllen

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.

Literaturrecherche für das zu bewertende Arzneimittel

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

Literaturrecherche zur Durchführung eines indirekten Vergleichs – Studien mit der zweckmäßigen Vergleichstherapie bzw. der alternativen Vergleichstherapie

Da für das Anwendungsgebiet *Add-on* zu Sulfonylharnstoff keine Studien vorliegen, die einen direkten Vergleich zwischen dem zu bewertenden Arzneimittel und der zweckmäßigen (bzw. alternativen) Vergleichstherapie ermöglichen, werden zur Bewertung des medizinischen Zusatznutzens von Dapagliflozin indirekte Vergleiche herangezogen.

Zur Beurteilung der klinischen Sicherheit und Wirksamkeit der ausgewählten zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Metformin) bzw. alternativen Vergleichstherapie (Sulfonylharnstoff+DPP-4-Inhibitor) im Vergleich mit dem zu bewertenden Arzneimittel (Sulfonylharnstoff+Dapagliflozin) wird eine bibliografische Literaturrecherche unter der Fragestellung „Identifikation von Studien (RCT) zur Sicherheit und Wirksamkeit der *Add-on*-Therapie mit Metformin bzw. DPP-4-Inhibitoren zusätzlich zur Sulfonylharnstoff-Monotherapie zur Verbesserung der Blutzuckerkontrolle von Typ-2-Diabetikern“ durchgeführt. Ziel dieser Literaturrecherche ist es, RCT mit der zweckmäßigen Vergleichstherapie (bzw. alternativen Vergleichstherapie) zu identifizieren, bei denen die Kontrollgruppe mit einem Sulfonylharnstoff und Placebo behandelt wurde.

Beschreibung der Suche

Die bibliografischen Literaturrecherchen werden in den Datenbanken MEDLINE, EMBASE und Cochrane (*Cochrane Central Register of Controlled Trials*) durchgeführt. Für jede einzelne Datenbank wird eine sequentielle und individuell adaptierte Suchstrategie entwickelt und separat angewendet. Die individuellen Suchstrategien sind in Anhang 4-A/Anhang 4-A dokumentiert.

Bei allen Recherchen werden nur Humanstudien mit englischem oder deutschem Volltext berücksichtigt. Eine Einschränkung auf RCT wird mithilfe von validierten Filtern vorgenommen. Für die Suche wird die Plattform des Deutschen Instituts für Medizinische Dokumentation und Information (DIMDI), sowie die Oberfläche von OVID verwendet.

Die Ergebnisse der systematischen bibliografischen Literaturrecherche sind 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.

Studienregistersuche für das zu bewertende Arzneimittel

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 durchgeführt. Die Suche erfolgt entsprechend den Vorgaben in den Studienregistern clinicaltrials.gov und dem *International Clinical Trials Registry Platform Search Portal* (ICTRP Search Portal, Suchportal der WHO).

Da das Studienregister clinicalstudyresults.org nicht mehr zur Verfügung steht, wurde hier keine Suche vorgenommen.

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

Studienregistersuche zur Durchführung eines indirekten Vergleiches

Um sicherzustellen, dass alle verfügbaren Studiendaten zu den Vergleichstherapien für die Bewertung berücksichtigt werden, werden weitere Studienregistersuchen nach abgeschlossenen, abgebrochenen und laufenden Studien mit einer Kombinationstherapie mit Metformin und Sulfonylharnstoffen sowie mit einer Kombinationstherapie mit DPP-4-Inhibitoren und Sulfonylharnstoffen durchgeführt. Die Studienregistersuchen für die Vergleichstherapien werden in denselben Datenbanken wie die für das zu bewertende Arzneimittel durchgeführt.

Die detaillierte Suchstrategie ist in Anhang 4-B dokumentiert.

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 sowie die in der systematischen Literaturrecherche und in der Studienregistersuche identifizierten Studien für die zweckmäßige Vergleichstherapie (bzw. alternative Vergleichstherapie) wird unter Anwendung der in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien nach Titel und Abstract 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.

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 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-F):

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 für die Fragestellung relevanten Studien werden hinsichtlich ihrer Aussagekraft bewertet. Die Bewertung des Verzerrungspotenzials erfolgt entsprechend der Vorgaben der Dossievorlage: Basis der Bewertung sind Informationen aus klinischen Studienberichten (falls vorhanden) oder aus Volltextpublikationen der jeweiligen klinischen Studien.

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 Verzerrungspotentials der Einzelstudien werden tabellarisch zusammengefasst. Gemäß Verfahrensordnung wurde 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 würden.

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“ führt nicht zum Ausschluss aus der Bewertung. Die Klassifizierung dient vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

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

Eine zusammenfassende Bewertung der Verzerrungsaspekte erfolgt 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)¹. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-² bzw. STROBE-Statements³ 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-A). Der Patientenfluss wird gemäß CONSORT-Flow-Chart dargestellt.

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 eingeschlossenen Patientenpopulation werden sowohl demographische 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

¹ Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.

² 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.

³ 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.

werden anhand des Mittelwertes und der Standardabweichung 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 werde für dichotome Zielvariablen das OR als das vom G-BA präferierte Effektmaß berechnet. Im Fall, dass in einem Behandlungsarm kein Ereignis beobachtet wurde (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} .

Folgende 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, 1998a; 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 4-4 ist die für diesen Zweck herangezogene Methodik dargestellt.

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

Ergebnis für die Einzelendpunkte		Ergebnis der gemeinsamen Betrachtung
Blutzuckerkontrolle (HbA1c)	Hypoglykämie-Rate	
		-Bewertung des Zusatznutzens
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 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 Ereigniss 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 Nichtunterlegenheitsgrenze (>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 Therapieverträglichkeit 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.

Zur Analyse der Verträglichkeit und Sicherheit wurden 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. Für den Endpunkt kardiale Ereignisse 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. Da im Verlauf einer Dapagliflozin-Behandlung vermehrt Ereignisse, die ein Anzeichen für Genital- bzw. Harnwegsinfektionen sein können, auftraten, wird auch die Gesamtrate dieser Ereignisse als unerwünschte Ereignisse vom speziellen Interesse in Ergänzung zur Gesamtrate unerwünschter Ereignisse dargestellt. Bei einer erhöhten Gesamtrate unerwünschter Ereignisse, werden die unerwünschten Ereignisse vom speziellen Interesse mit in die Gesamtbeurteilung herangezogen. Wenn die Vergleichstherapie mit anderweitigen unerwünschten Ereignissen vergesellschaftet ist, welche die Gesamtrate ausgleichen, wird auf die Beurteilung einzelner Ereignisse verzichtet, da eine Gewichtung unterschiedlicher einzelner Ereignisse untereinander nicht operationalisierbar ist.

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 Faktoren des sogenannten metabolischen Syndroms, d. h. Übergewicht, Hypertonie und Fettstoffwechselstörung, assoziiert ist und ein viszerales Übergewicht einen gemeinsamen Risikofaktor für die Grunderkrankung (Diabetes mellitus) und kardiovaskuläre Erkrankungen darstellt, gilt ein gewichtsneutraler bzw. ein gewichtsreduzierender Effekt einer antidiabetischen Therapie des Diabetes mellitus Typ 2 in den Leitlinien als vorteilhaft und stellt somit ein patientenrelevantes Therapieziel dar. Daher wurden Gewichtsveränderungen als patientenrelevant betrachtet.

Bezüglich der Verbesserung der Lebensqualität wird die Therapiezufriedenheit anhand des *Diabetes Treatment Satisfaction Questionnaire Status* (DTSQs) berücksichtigt. Der DTSQs ist ein Messinstrument, das die Therapiezufriedenheit anhand von drei Dimensionen widerspiegelt: Behandlungszufriedenheit, Hyperglykämiewahrnehmung und Hypoglykämiewahrnehmung. Der DTSQs ist ein validierter Fragebogen, welcher häufig zur Erfassung der Therapiezufriedenheit bei Diabetes Mellitus Typ 2 angewendet wird und kann zur Nutzenbewertung herangezogen werden. Der DTSQs von der WHO und der internationalen Diabetesgesellschaft zu Messung der Therapiezufriedenheit empfohlen (Bradley C., 1994).

Alle in der Zusatznutzenbeurteilung betrachteten Endpunkte werden als patientenrelevant eingestuft, da sie einerseits die krankheitsbedingte Morbidität wiedergeben 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 und Gewichtszunahmen stellt ein relevantes Therapieziel einer antihyperglykämischen Therapie des Diabetes mellitus Typ 2 dar.

Weiterhin wird die Änderung im systolischen und diastolischen Blutdruck als Surrogatparameter für die kardiovaskuläre Morbidität und Mortalität 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.

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 Effekten⁴ 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

⁴ DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-188.

Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁵ 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 Heterogenität⁶ 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.

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 der *Software Review Manager* (RevMan, Version 5) durchgeführt. Sofern berichtet, werden primär die Resultate aus *Intention-to-treat*-Analysen kombiniert, adjustierte Werte werden, falls vorhanden, gegenüber „rohen“ Werten bevorzugt. Alle Meta-Analysen basieren auf einem Modell mit zufälligen Effekten (REM); Modelle mit festen Effekten (FEM) können im Rahmen von Sensitivitätsanalysen verwendet werden. 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, aus 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

⁵ 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.

⁶ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. *Measuring inconsistency in meta-analyses*. *BMJ* 2003;327(7414):557-560.

Studienresultate oder im Falle hoher Ereignisraten, werden auch andere Effektmaße berechnet (wie z.B. das RR) und für Sensitivitätsanalysen verwendet (siehe Abschnitt 4.2.5.4). Diese Effektmaße wurden jedoch nur dann im Dossier dargestellt, wenn es relevante Änderungen/Abweichungen im Vergleich zum OR gab. Für kontinuierliche Variablen werden Mittelwertdifferenzen oder standardisierte Mittelwertdifferenzen (Hedges' g) berechnet.

Zur Abschätzung der klinischen Relevanz wurden, 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 Responseudefinition 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 – im Falle einer Meta-Analyse - 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.

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 können zusammenfassende Schätzer berechnet werden (Skipka und Bender, 2010); die Interpretierbarkeit der Schätzer muss dann allerdings kritisch diskutiert werden 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, Endpunktoperalisierungen, 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.

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.

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, qualitative versus quantitative Interaktion) sowie hinsichtlich möglicher Auswirkungen auf das aus der Gesamtpopulation abgeleitete Fazit diskutiert. Diese Bewertung im Gesamtkontext ist insbesondere notwendig, da trotz einer großen Anzahl durchgeführter Subgruppenanalysen keine Korrektur für multiples Testen vorgenommen wurde. 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)
- asiatischer und nicht-asiatischer Abstammung

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 Analysemethoden dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich⁷. Komplexe Verfahren für den simultanen Vergleich von 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“⁸, „Multiple-Treatment-Meta-Analysen“⁹ oder auch „Netzwerk-Meta-Analysen“¹⁰, 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 ist¹¹. Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

⁷ Glenny AM, Altman DG, Song F, Sakarovich C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-148.

⁸ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

⁹ Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

¹⁰ Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17(3): 279-301.

¹¹ B. Schöttker, D. Lühmann, D. Boukhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

¹² 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.

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

- *Genauere 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.*

Da es keine direkte *Head-to-Head*-Studie Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin gibt, werden adjustierte indirekte Vergleiche betrachtet. Die methodische Basis eines adjustierten indirekten Vergleiches von verschiedenen Behandlungsmöglichkeiten bildet ein gemeinsamer Vergleichsarm, anhand dessen sich die Ergebnisse zur Wirksamkeit und Sicherheit beider Interventionen quantifizieren lassen (Glenny et al., 2005).

Der einfache adjustierte indirekte Vergleich wird mittels eines frequentistischen Ansatzes durchgeführt. Er beruht auf einem einfachen Netzwerk und wird für Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin bzw. Sulfonylharnstoff+DPP-4-Inhibitor mit Sulfonylharnstoff+Placebo als Brückenkomparator durchgeführt.

Grundlage für diesen Vergleich sind folgende Beziehungen der Effektmaße und ihrer Varianz:

- $\ln \theta_{BC} = \ln \theta_{AB} - \ln \theta_{AC}$ (1)
- $\text{Var}(\ln \theta_{BC}) = \text{Var}(\ln \theta_{AB}) + \text{Var}(\ln \theta_{AC})$ (2)
- Das 95% Konfidenzintervall für $\ln \theta_{BC}$ berechnet sich als $\ln \theta_{BC} \pm 1.96 * \sqrt{\text{Var}(\ln \theta_{BC})}$ (3)

Effektmaße, die als θ in den Formeln (1) bis (3) geeignet sind, sind das RR, das OR, sowie das HR. Für kontinuierliche, normalverteilte Variablen entfällt die Log-Transformation vor Berechnung des indirekten Effektschätzers und seiner Streuung. Grundannahme für die Gültigkeit dieser Zusammenhänge ist die Abwesenheit bedeutsamer Heterogenität zwischen

den eingeschlossenen Studien (methodische Faktoren, Effektmodifikatoren). Für den Fall, dass vor Durchführung des adjustierten indirekten Vergleiches eine Meta-Analyse durchgeführt werden muss, ist die Untersuchung der Heterogenität der eingeschlossenen Studien mittels statistischer Tests sowie gebräuchlicher Heterogenitätsmaße vorgesehen (siehe Abschnitt 4.2.5.3). Des Weiteren werden die eingeschlossenen Studien hinsichtlich wichtiger Eigenschaften (Ein- und Ausschlusskriterien; Patientencharakteristika etc.) charakterisiert und auf das Vorliegen von wesentlichen Unterschieden hin untersucht (klinische Heterogenität).

Alle Studien und Resultate zur primären Fragestellung (Vergleichstherapie gemäß Festlegung durch den G-BA) sowie zur sekundären Fragestellung (Vergleich der Arzneimittel der Wirkstoffgruppe der DPP-4-Inhibitoren) werden im Abschnitt 4.3.2.1 präsentiert.

Für den indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor wurden die verschiedenen DPP-4-Inhibitoren aufgrund ihres vergleichbaren Wirkungs- und Nebenwirkungsprofils in einer gemeinsamen Analyse zusammengefasst. Zusätzlich wurde der indirekte Vergleich gegenüber Sulfonylharnstoff+Sitagliptin einzeln dargestellt, da Sitagliptin als Leitsubstanz der DPP-4-Inhibitoren in Deutschland angesehen wird. Bzgl. der für den indirekten Vergleich relevanten Tagesdosierungen der DPP-4-Inhibitoren wurde in erster Linie die Tagesdosis gewählt, die in der Indikation *Add-on* zu Sulfonylharnstoff-Monotherapie in Deutschland zugelassen ist. Bestand in dieser Indikation keine Zulassung, wurden nah verwandte Zulassungsgebiete herangezogen (z.B. *Add-on* zur Metformin-Monotherapie). Bestand für den DPP-4-Inhibitor zum Zeitpunkt der Dossiererstellung keine Zulassung in Deutschland, so wurde aus Gründen der Konservativität die wirksamere Dosis als relevant für den indirekten Vergleich angesehen.

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 4-5: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
MB102029	ja	abgeschlossen	104 Wochen	Dapagliflozin 5 mg+Vorbehandlung Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00005	ja	abgeschlossen	48 Wochen	Dapagliflozin 2,5 mg+Glimepirid Dapagliflozin 5 mg+Glimepirid Dapagliflozin 10 mg+Glimepirid Placebo+ Glimepirid
MB102020	nein	abgebrochen	4 Wochen	Dapagliflozin 5 mg Dapagliflozin 20 mg
MB102035	nein	abgeschlossen	12 Wochen	Dapagliflozin 10 mg+Vorbehandlung (Metformin und/oder Sulfonylharnstoff) Placebo+Vorbehandlung (Metformin und/oder Sulfonylharnstoff)
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
MB102054	nein	abgeschlossen*	24 Wochen	Dapagliflozin 5 mg+Vorbehandlung Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
MB102073	nein	laufend	12 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
MB102077	nein	laufend	12 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung

* Es liegt noch kein Studienbericht vor.

Geben Sie an, welchen Stand die Information in Tabelle 4-5 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 Studienstatus in Tabelle 4-5 ist zum 29.10.2012 abgebildet.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-5 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-6: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
MB102029	Die Studie untersucht nur Patienten mit Nierenfunktionsstörung (eGFR zwischen 30 mL/min/1,73 m ² und 59 mL/min/1,73 m ²). Für diese Patienten wird eine Behandlung mit Dapagliflozin nicht empfohlen.
MB102020	Phase-1-Studie. Die Studiendauer beträgt nur 4 Wochen.
MB102035	Die Vergleichstherapie ist Hydrochlorothiazid. Die Studiendauer ist zu gering.
D1690C00018 (inkl. LT I und II)	Die Vorbehandlung der Patienten ist heterogen. Eine Subgruppenanalyse nach Vorbehandlung wird für die von uns betrachteten Sulfonylharnstoffe (Glimepirid, Glibenclamid und Glipizid) nicht berichtet.
D1690C00019 (inkl. LT I und II)	Die Vorbehandlung der Patienten ist heterogen. Eine Subgruppenanalyse nach Vorbehandlung wird für die von uns betrachteten Sulfonylharnstoffe (Glimepirid, Glibenclamid und Glipizid) nicht berichtet.
MB102054	Die Studie wurde kürzlich abgeschlossen. Es liegen noch keine Ergebnisse vor. Die Studie wurde daher in Anhang 4-E aufgenommen.
MB102073	Die Studie läuft noch. Es liegen keine Ergebnisse vor. Die Studiendauer ist zu gering.
MB102077	Die Studie läuft noch. Es liegen keine Ergebnisse vor. Die Studiendauer ist zu gering.

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.

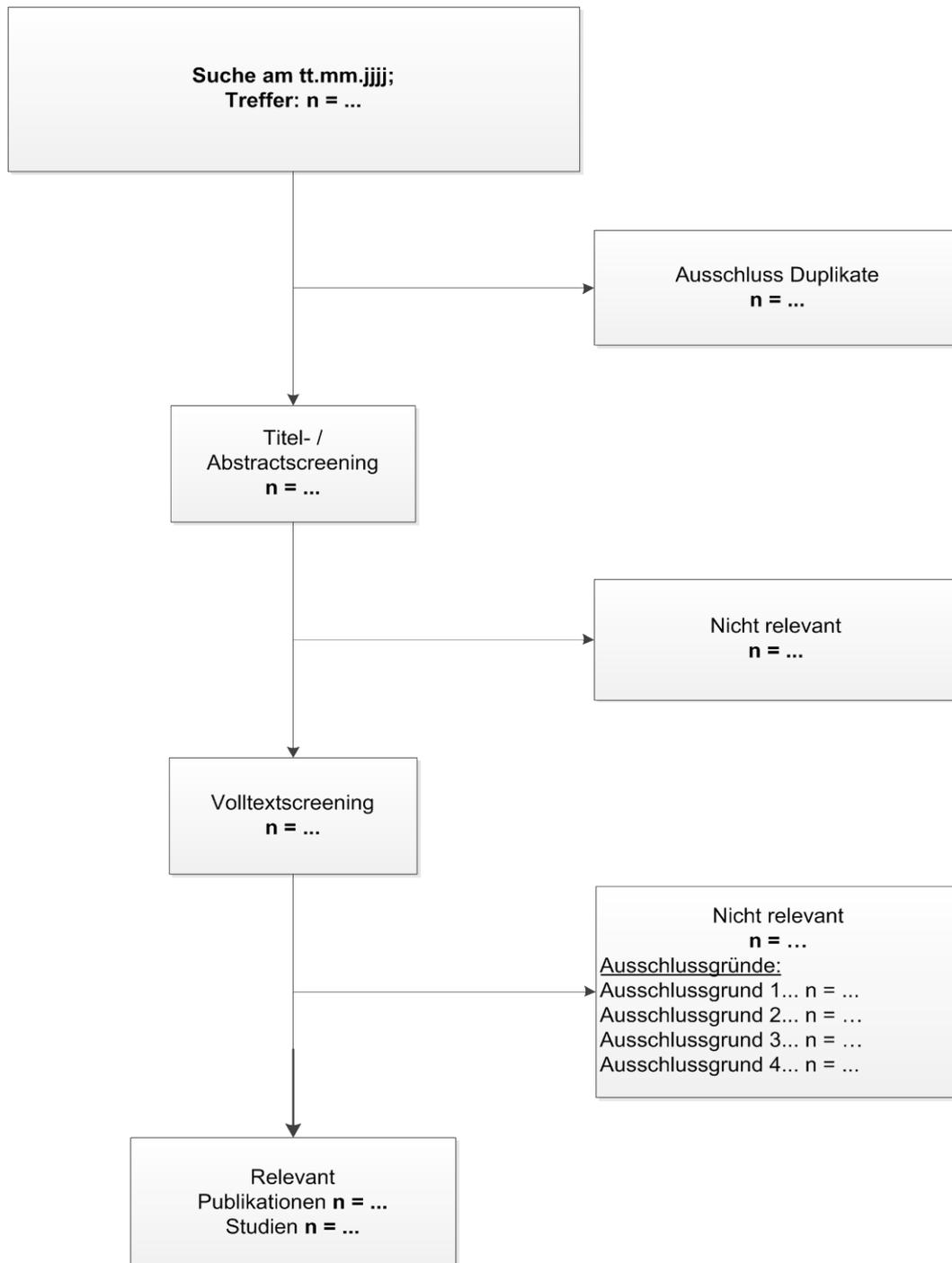


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Wie bereits in Abschnitt 4.2.3.2 beschrieben, wurde keine bibliographische 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 4-5) 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 4-7: 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)
D1690C00005	clinicaltrials.gov [NCT00680745] (Clinicaltrials.gov, 2012l) ICTRP Search Portal [EUCTR2007-005931-27-HU] (ICTRP Search Portal, 2012d) [NCT00680745] (ICTRP Search Portal, 2012c)	ja	n. d.
<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>			

4.3.1.1.4 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 beispiel-

haften 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 4-8: 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						
D1690C 00005	ja	ja	nein	ja (AstraZeneca, 2010a) (AstraZeneca, 2010b)	ja [NCT00680745] (Clinicaltrials.gov, 2012l) [EUCTR2007-005931-27-HU] (ICTRP Search Portal, 2012d) [NCT00680745] (ICTRP Search Portal, 2012c)	ja (Strojek et al., 2011)
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
--						
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.</p> <p>b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

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-A zu hinterlegen.

Tabelle 4-9: 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
D1690C00005	RCT, Phase-III-Studie, doppelblind, parallel, placebokontrolliert, multizentrisch (84 Zentren in 7 Ländern)	Erwachsene Patienten mit Typ-2-Diabetes, deren Blutzuckerwert alleine mit Glimepirid (Sulfonylharnstoff) nicht ausreichend kontrolliert werden kann (HbA1c \geq 7,0% und \leq 10%)	<ul style="list-style-type: none"> • Glimepirid+ Dapagliflozin 2,5 mg (n=154) • Glimepirid+ Dapagliflozin 5 mg (n=145) • Glimepirid+ Dapagliflozin 10 mg (n=151) • Glimepirid+ Placebo (n=146) 	Enrollment-Phase: 1 Woche Glimepirid (offene Gabe)-Lead-in Phase: 8 Wochen Qualifikationsphase: 1 Woche Behandlung: 24 Wochen Verlängerungsphase: 24 Wochen Follow-up-Phase nach Beenden der Studienmedikation: 3 Wochen	International 4/2008 – 5/2010	Primärer Endpunkt: HbA1c-Änderung in Woche 24 im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: HbA1c-Änderung in Woche 48 im Vergleich zum Ausgangswert Gewichtveränderung in Woche 24 und 48 im Vergleich zum Ausgangswert Veränderung des Blutdrucks Therapiezufriedenheit Anzahl Hypoglykämien Kardiale Ereignisse Unerwünschte Ereignisse

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

Studie	Zu bewertende Arzneimittel	Zweckmäßige Vergleichstherapie	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
D169C00005	Glimepirid ¹ (4 mg/Tage) + Dapagliflozin (2,5, 5 oder 10 mg)	Glimepirid ¹ (4 mg/Tag) + Placebo	Enrollment-Phase: 1 Woche Glimepirid (offene Gabe)-Lead-in Phase: 8 Wochen Qualifikationsphase: 1 Woche Behandlung: 24 Wochen Verlängerungsphase: 24 Wochen Follow-up-Phase nach Beenden der Studienmedikation: 3 Wochen Es gab während der Behandlungs- und Verlängerungsphase eine Medikation für den Notfall (Metformin, Pioglitazon oder Rosiglitazon), falls keine ausreichende Blutzuckerkontrolle erreicht wurde.

¹ Glimepirid konnte bis auf 2 mg heruntertitriert werden oder ganz abgesetzt werden, um Hypoglykämien zu vermeiden. Eine Hochtitrierung war nicht erlaubt. Bei Bedarf erhielt der Patient eine Rescuemedikation.

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

Studie Gruppe	N	Alter (Jahre)	Geschlecht w / m (%)	BMI (kg/m ²) Mittelwert (SD)	Gewicht (kg) Mittelwert (SD)
D169C00005					
Glimepirid+Dapagliflozin* 5 mg	142	60,2 (9,73)	50,0/50,0	29,84 (5,182)	81,00 (18,635)
Glimepirid+Dapagliflozin* 10 mg	151	58,9 (8,32)	56,3/43,7	29,75 (5,641)	80,56 (17,869)
Glimepirid+Placebo	145	60,3 (10,16)	51,0/49,0	29,74 (4,569)	80,94 (15,773)

*es werden nur die Gruppen mit den Dosierungen 5 und 10 mg Dapagliflozin betrachtet, da die Dosierung 2,5 mg nicht zugelassen ist und somit keine Aussagekraft für die Nutzenbewertung hat.

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

Studie Gruppe	N	Ausgangswert HbA1c (%) Mittelwert (SD)	Diabetesdauer (Jahre) Mittelwert (SD)
D169C00005			
Glimepirid+Dapagliflozin* 5 mg	142	8,12 (0,781)	7,35 (5,706)
Glimepirid+Dapagliflozin* 10 mg	151	8,07 (0,790)	7,15 (5,459)
Glimepirid+Placebo	145	8,15 (0,736)	7,38 (5,741)

*es werden nur die Gruppen mit den Dosierungen 5 und 10 mg Dapagliflozin betrachtet, da die Dosierung 2,5 mg nicht zugelassen ist und somit keine Aussagekraft für die Nutzenbewertung hat.

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

Studie Gruppe	N	Studienabbrecher	
		Nach 24 Wochen:	Nach 48 Wochen
D1690C00005			
Glimepirid+Dapagliflozin* 5 mg	145	13 (9,0)	17 (11,7)
Glimepirid+Dapagliflozin* 10 mg	151	10 (6,6)	18 (11,9)
Glimepirid+Placebo	146	13 (8,9)	19 (8,9)

*es werden nur die Gruppen mit den Dosierungen 5 und 10 mg Dapagliflozin betrachtet, da die Dosierung 2,5 mg nicht zugelassen ist und somit keine Aussagekraft für die Nutzenbewertung hat.

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

Glimepirid+Dapagliflozin versus Glimepirid+Placebo:

Studie D1690C00005: randomisierte, doppelblinde, placebokontrollierte, 24-wöchige Phase-III-Studie mit einer Verlängerungsphase von 24 Wochen, um die Wirksamkeit, Sicherheit und Verträglichkeit von Dapagliflozin in Kombination mit Glimepirid bei Patienten mit Typ-2-Diabetes mit einem HbA1c-Wert von $\geq 7\%$ und $\leq 10\%$ zu untersuchen. In die Studie eingeschlossen wurden Typ-2-Diabetes Patienten, die älter waren als 18 Jahre und mit einer Sulfonylharnstofftherapie (mind. halbmaximale Tagesdosis) unzureichend kontrolliert waren. Die Patienten wurden zunächst in einer 8-wöchigen Run-In Phase open-label mit Glimepirid 4 mg/Tag behandelt. Bei Erfüllung der Ein- und Ausschlusskriterien wurden die Patienten danach mit dem Zuteilungsverhältnis 1:1:1:1 randomisiert zu einer zusätzlichen Behandlung mit 2,5 mg, 5 mg, 10 mg Dapagliflozin oder Placebo. Eingeschlossene Patienten hatten Durchschnittswerte für Alter im Bereich von 58,9 – 60,3 Jahren, für BMI von 29,74 – 29,84, für HbA1c-Ausgangswert von 8,07 – 8,15% und für Diabetesdauer 7,15 – 7,38 Jahre. Während der Studie konnte Glimepirid (offene Behandlung) in der Dosis reduziert oder komplett abgesetzt werden (z.B. bei Auftreten von Hypoglykämien), eine Dosissteigerung war nicht erlaubt. Patienten mit inadäquater glykämischer Kontrolle (nach vordefinierten Kriterien) konnten während der Studienteilnahme *Rescue*-Medikation erhalten (Metformin, Pioglitazon oder Rosiglitazon).

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

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			
D1690C00005	ja	ja	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Studie D169C000005: Es handelt sich hierbei um eine doppelblinde, randomisierte, placebokontrollierte Studie. Die Randomisierungssequenz wurde adäquat durchgeführt. Die Randomisierung (Blockrandomisierung) erfolgte streng nach der Reihenfolge innerhalb eines Zentrums. Die Patienten und Behandler waren verblindet. Die Erhebung und Auswertung der Ergebnisse zu den einzelnen definierten Endpunkten erfolgte planmäßig und wie vorab definiert. Es ergaben 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	Hypoglykämien	Gewichtsveränderung	Un-erwünschte Ereignisse	Folgekomplikationen	Gesamtmortalität	Systolischer und diastolischer Blutdruck
D1690C00005	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 einer Überlebenszeitanalyse untersucht. Insgesamt gab es einen Todesfall in der Sulfonylharnstoff+Dapagliflozin-Gruppe und keinen Todesfall in der Sulfonylharnstoff+Placebo-Gruppe nach 48 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 Komata	Stationäre Behandlungen	Gesundheitsbezogene Lebensqualität	Therapiezufriedenheit	Ausmaß der Blutzuckersenkung (HbA1c)
D1690C00005	nein	nein	nein	ja	ja*

* Inklusive Studienabbrüche aufgrund unzureichender glykämischer Kontrolle

Nicht berichtete Zielgrößen

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

Verwendete Daten

Alle im Folgenden beschriebenen Ergebnisse werden jeweils für die Studienpopulation (Patienten mit manifestem Diabetes mellitus Typ 2, die unter einer Sulfonylharnstoff-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 GFR <60 mL/min/1,73 m² bzw. Kreatininclearance <60 mL/min sowie von Patienten im Alter von ≥75 Jahren und Patienten, die mit Schleifendiuretika vorbehandelt waren) dargestellt. Weiterhin wird jeweils die 5 mg Dosis, sowie die 10 mg Dosis von Dapagliflozin berichtet. Die Daten für die Zielpopulation wurden post-hoc analysiert (Bristol-Myers Squibb und AstraZeneca GmbH, 2012a).

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 abweichen.

Für die Wirksamkeitsendpunkte „HbA1c-Veränderung“ und „Gewichtsveränderungen“ und den Sicherheitsendpunkt „Hypoglykämien“ wurden die Daten bis zur Einnahme der Notfallmedikation verwendet. Für alle anderen sicherheitsrelevanten Endpunkte wurden auch Daten nach Einnahme der Notfallmedikation berücksichtigt.

4.3.1.3.1 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
D1690C00005	<ul style="list-style-type: none"> • Anhand der Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 24 (falls der Wert nicht erhoben wurde, wurde der letztmögliche vor Woche 24 (LOCF) betrachtet) bzw. bis Woche 48 (<i>Repeated Measurements</i> Modell) • Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

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
D1690C00005	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Ausmaß der Blutzuckerkontrolle (HbA1c) ist in dieser Studie der primäre Endpunkt und wurde als adjustierte Differenz der Mittelwerte der HbA1c-Werte in Woche 24 und zu Studienbeginn berechnet. Um den Erhalt der Wirksamkeit von jeder Dosis Dapagliflozin zu bewerten, wurde zusätzlich die Veränderung des HbA1c nach 48 Wochen erhoben. Auch Studienabbrüche und Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle wurden in beiden Behandlungsgruppen erhoben. Die Endpunkterheber waren zu jedem Zeitpunkt der Erhebung verblindet. In allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn gleiche 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 innerhalb der ersten 24 Wochen durch die LOCF-Methode ersetzt. Nach 48 Wochen wurde kein Ersatz mit der LOCF-Methode vorgenommen und eine *Repeated Measure Analysis* (Analyse für wiederholte Messungen) durchgeführt. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Somit kann das Verzerrungspotential für den Endpunkte „Ausmaß der Blutzuckerkontrolle“ 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 4-19: 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	Ausgangs- wert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert° % (SE)	N	Ausgangs- wert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert° % (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
	5 mg: 142	8,12 (0,781)	7,49 (0,853)	-0,63 (0,0627)	143	8,15 (0,741)	8,00 (0,928)	-0,13 (0,0625)	-0,49 (0,0885)*** [-0,67;-0,32]
	10 mg: 150	8,07 (0,792)	7,27 (0,843)	-0,82 (0,0610)					-0,68 (0,0873)*** [-0,86;-0,51]
24 Wochen**									
	5 mg: 128		7,40 (0,891)	-0,62 (0,0646)	110	7,80 (0,781)	-0,14 (0,0668)	-0,48 (0,0928) [-0,67;-0,30]	
	10 mg: 137		7,25 (0,821)	-0,82 (0,0624)				-0,68 (0,0914) [-0,86;-0,50]	
48 Wochen**									
	5 mg: 90		7,08 (0,742)	-0,56 (0,0764)	55	7,29 (0,657)	-0,04 (0,0873)	-0,53 (0,1155) [-0,75;-0,30]	
	10 mg: 107		7,11 (0,620)	-0,73 (0,0721)				-0,70 (0,1129) [-0,92;-0,47]	

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF; ** Repeated measure Analyse;*** p-Wert <0,0001(ANCOVA mit α -Korrektur nach Dunnett); °adjustiert bezüglich Ausgangswert

SD: Standardabweichung; SE: Standardfehler, KI: Konfidenzintervall

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 ^o % (SE)	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert ^o % (SE)	
D1690C000 05	Sulfonylharnstoff +Dapagliflozin				Sulfonylharnstoff +Placebo				
	24 Wochen*								
	5 mg: 113	8,19 (0,791)	7,54 (0,895)	-0,63 (0,0692)	111	8,17 (0,744)	8,03 (0,951)	-0,13 (0,0698)	-0,50 (0,0982) [-0,70;-0,31]
	10 mg: 129	8,09 (0,800)	7,25 (0,829)	-0,86 (0,0647)					-0,74 (0,0952) [-0,92;-0,55]
24 Wochen**									
	5 mg: 102		7,46 (0,889)	-0,63 (0,0718)	84		7,84 (0,834)	-0,12 (0,0748)	-0,51 (0,1036) [-0,71;-0,30]
10 mg: 120		7,25 (0,841)	-0,86 (0,0662)	-0,74 (0,0999) [-0,93;-0,54]					
48 Wochen**									
	5 mg: 73		7,13 (0,765)	-0,58 (0,0861)	41		7,32 (0,711)	-0,02 (0,1024)	-0,56 (0,1330) [-0,83;-0,30]
10 mg: 94		7,09 (0,626)	-0,79 (0,0776)	-0,77 (0,1279) [-1,02;-0,52]					

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

* LOCF; ** *Repeated measure* Analyse;

SD: Standardabweichung; SE: Standardfehler; KI: Konfidenzintervall

^oadjustiert bezüglich Ausgangswert

Als unterstützende Wirksamkeitsmessung wurden die Studienabbrüche (aufgrund von unzureichender Blutzuckerkontrolle) oder die Notfallmedikation (die aufgrund unzureichender Blutzuckerkontrolle verabreicht wurde) erhoben. Unzureichende Blutzuckerkontrolle, die den Abbruch trotz der maximalen Dosis des Notfallmedikaments notwendig machten, waren HbA1c-Werte von >8% für einen Zeitraum von 12 Wochen.

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

Studie	N	Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle, n (%)* SE
D1690C00005		
Nach 24 Wochen		
Sulfonylharnstoff +Dapagliflozin 5 mg	142	8 (5,6) 1,931
Sulfonylharnstoff +Dapagliflozin 10 mg	151	3 (2,0) 1,136
Sulfonylharnstoff +Placebo	145	24 (16,2) 2,952
5 mg vs.Placebo OR [95%-KI]		0,30 [0,13;0,70]
10 mg vs. Placebo OR [95%-KI]		0,10 [0,03;0,35]
Nach 48 Wochen		
Sulfonylharnstoff +Dapagliflozin 5 mg	142	42 (29,1) 3,610
Sulfonylharnstoff +Dapagliflozin 10 mg	151	27 (18,4) 3,030
Sulfonylharnstoff +Placebo	145	77 (52,1) 3,841
5 mg vs.Placebo OR [95%-KI]		0,37 [0,23;0,60]
10 mg vs. Placebo OR [95%-KI]		0,19 [0,11;0,33]

*Prozentangaben adjustiert bezüglich HbA1c -Ausgangswert
SE: Standardfehler; KI: Konfidenzintervall; OR: Odds Ratio

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

Studie	N	Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle, n (%*) SE
D1690C00005		
Nach 24 Wochen		
Sulfonylharnstoff +Dapagliflozin 5 mg	113	6 (5,2) 2,107
Sulfonylharnstoff +Dapagliflozin 10 mg	129	2 (1,6) 1,088
Sulfonylharnstoff+Placebo	113	21 (18,3) 3,527
5 mg vs. Placebo OR [95%-KI]		0,25 [0,10;0,63]
10 mg vs. Placebo OR [95%-KI]		0,07 [0,02;0,30]
Nach 48 Wochen		
Sulfonylharnstoff +Dapagliflozin 5 mg	113	33 (28,0) 4,042
Sulfonylharnstoff +Dapagliflozin 10 mg	129	23 (18,3) 3,3294
Sulfonylharnstoff +Placebo	113	61 (52,5) 4,227
5 mg vs. Placebo OR [95%-KI]		0,35 [0,20;0,61]
10 mg vs. Placebo OR [95%-KI]		0,18 [0,10;0,33]

* Prozentangaben adjustiert bezüglich Ausgangswert des HbA1c
SE: Standardfehler; KI: Konfidenzintervall; OR: Odds Ratio

Die Behandlung mit Sulfonylharnstoff +Dapagliflozin war bezogen auf die HbA1c-Reduktion in beiden Dosisgruppen nach 24 Wochen signifikant besser als die Behandlung mit Sulfonylharnstoff+Placebo. Auch nach 48 Wochen blieb der Effekt bestehen. Die Anzahl an Patienten, die aufgrund unzureichender glykämischer Kontrolle die Studie abbrachen oder ein Notfallmedikament einnahmen, war in der Sulfonylharnstoff+Placebo-Gruppe signifikant höher, als in den Sulfonylharnstoff+Dapagliflozin-Gruppen (5 mg und 10 mg).

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.

4.3.1.3.1.1 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
D1690C00005	<p>Anzahl von Patienten mit mindestens einer Hypoglykämie innerhalb der 24/48 Wochen (mittels <i>diary review</i>).</p> <p>Hypoglykämien wurden noch unterteilt in:</p> <p>Schwerwiegende Hypoglykämien: Anzahl der Patienten mit mind. einer symptomatischen Hypoglykämie, bei der externe Hilfe nötig war (Blutzuckerwert von ≤ 54 mg/dL)</p> <p>Leichte Hypoglykämien: Entweder Anzahl der Patienten mit mindestens einer symptomatischen Hypoglykämie mit Blutzuckerwert von < 63 mg/dL Hypoglykämie, bei der keine Hilfe nötig war oder Anzahl der Patienten mit einer asymptomatischen Hypoglykämie mit bestätigtem Blutzuckerwert < 63 mg/dL</p> <p>Andere Hypoglykämien: Anzahl der Patienten mit einer symptomatischen Hypoglykämie (patientenberichtet), keine bestätigte Messung durchgeführt</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

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
D1690C00005	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Auftreten einer Hypoglykämie war in dieser Studie ein Sicherheitsendpunkt. Es wurde die Anzahl der Patienten, bei denen mindestens eine Hypoglykämie beobachtet wurde (mittels einem von den Patienten geführten Tagebuch und einem Glucometer), erhoben. Die Endpunkterheber waren zu jedem Zeitpunkt verblindet, in allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn gleiche Ausfallraten. Der Endpunkt wurde bei allen Patienten erhoben, die mindestens eine Studienmedikation erhalten haben (*Safety Analysis Set*) inklusive der Patienten, die eine antidiabetische Notfallmedikation bekommen haben.

Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Somit konnte das Verzerrungspotential für den 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 D1690C000005	Zu bewertende Arzneimittel N / n (%)		Vergleichstherapie N / n (%)		Odds Ratio [95%-KI]
24 Wochen	Sulfonylharnstoff+Dapagliflozin		Sulfonylharnstoff+Placebo		
Patienten mit mind einer Hypoglykämie	5 mg: 145	10 (6,9)	146	7 (4,8)	1,47 [0,54;3,98]
	10 mg:151	12 (7,9)			1,71 [0,66;4,48]
schwer	5 mg: 145	0 (0,0)	146	0 (0,0)	Nicht berechenbar
	10 mg:151	0 (0,0)			Nicht berechenbar
leicht	5 mg: 145	8 (5,5)	146	3 (2,1)	2,78 [0,72;10,71]
	10 mg:151	10 (6,6)			3,38 [0,91;12,54]
andere	5 mg: 145	3 (2,1)	146	5 (3,4)	0,60 [0,14;2,54]
	10 mg:151	3 (2,0)			0,57 [0,13;2,44]
48 Wochen	Sulfonylharnstoff+Dapagliflozin		Sulfonylharnstoff+Placebo		
Patienten mit mind einer Hypoglykämie	5 mg: 145	15 (10,3)	146	10 (6,8)	1,57 [0,68;3,62]
	10 mg:151	17 (11,3)			1,73 [0,76;3,90]
schwer	5 mg: 145	0 (0,0)	146	0 (0,0)	Nicht berechenbar
	10 mg:151	0 (0,0)			Nicht berechenbar
leicht	5 mg: 145	12 (8,3)	146	4 (2,7)	3,20 [1,01;10,18]
	10 mg:151	13 (8,6)			3,34 [1,06;10,51]
andere	5 mg: 145	5 (3,4)	146	7 (4,8)	0,68 [0,21;2,19]
	10 mg:151	5 (3,3)			0,71 [0,22;2,29]
Abbruch aufgrund einer Hypoglykämie	5 mg: 145	0 (0,0)	146	0 (0,0)	Nicht berechenbar
	10 mg:151	0 (0,0)			Nicht berechenbar

KI: Konfidenzintervall

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

Studie D1690C00005	Zu bewertende Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
24 Wochen	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff +Placebo	
Patienten mit mind einer Hypoglykämie	5 mg: 116 8 (6,9)	114 5 (4,4)	1,61 [0,51;5,09]
	10 mg: 129 12 (9,3)		2,24 [0,76;6,55]
schwer	5 mg: 116 0	114 0	Nicht berechenbar
	10 mg: 129 0		Nicht berechenbar
leicht	5 mg: 116 6 (5,2)	114 2 (1,8)	3,05 [0,60;15,46]
	10 mg: 129 10 (7,8)		4,71 [1,01;21,95]
andere	5 mg: 116 3 (2,6)	114 4 (3,5)	0,73 [0,16;3,34]
	10 mg: 129 3 (2,3)		0,65 [0,14;2,99]
48 Wochen	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff+Placebo	
Patienten mit mind einer Hypoglykämie	5 mg: 116 12 (10,3)	114 8 (7,0)	1,53 [0,60;3,89]
	10 mg: 129 17 (13,2)		2,01 [0,83;4,86]
schwer	5 mg: 116 0	114 0	Nicht berechenbar
	10 mg: 129 0		Nicht berechenbar
leicht	5 mg: 116 9 (7,8)	114 3 (2,6)	3,11 [0,82;11,81]
	10 mg: 129 13 (10,1)		4,15 [1,15;14,94]
andere	5 mg: 116 4 (3,4)	114 6 (5,3)	0,64 [0,18;2,34]
	10 mg: 129 5 (3,9)		0,73 [0,22;2,45]
Abbruch aufgrund von Hypoglykämien	5 mg: 116 0 (0,0)	114 0 (0,0)	Nicht berechenbar
	10 mg: 129 0 (0,0)		Nicht berechenbar

KI: Konfidenzintervall

Leichte Hypoglykämien waren unter Dapagliflozin statistisch signifikant erhöht. Schwerwiegende Hypoglykämien traten innerhalb von 48 Wochen in keiner Behandlungsgruppe auf. In der Gesamtrate an Hypoglykämien (Patienten mit mindestens einer Hypoglykämie) zeigte sich kein statistisch signifikanter Unterschied. Ein Studienabbruch aufgrund von Hypoglykämien ereignete sich bis zum Studienende ebenfalls 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 3: 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 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 4-27: Operationalisierung von Gewichtsveränderung

Studie	Operationalisierung
D1680C00005	<ul style="list-style-type: none"> anhand der Änderung des Körpergewichts in kg von Studienbeginn bis Woche 24 / 48 (falls der Wert nicht erhoben wurde, wurde der letztmögliche vor Woche 24 (LOCF) betrachtet). Das Gewicht sollte ohne Schuhe, mit leichter Kleidung und „on a fasting stomach and empty bladder“ gemessen werden <p>Anhand einer Responder-Analyse:</p> <ul style="list-style-type: none"> Anzahl der Patienten mit einer Gewichtsabnahme von mehr als 3% oder 5% nach Woche 24 und 48

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-28: 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
D1690C00005	niedrig	ja	Ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Gewichtsveränderung war nach 24 Wochen in dieser Studie ein sekundärer, hierarchisch und konfirmatorisch getesteter Endpunkt und wurde als Differenz des mittleren Gewichts in Woche 24 und des Mittelwertes 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, 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. Auch wurde die Anzahl der Patienten mit einer Gewichtsabnahme von mehr als 3% und mehr als 5% in beiden Gruppen nach 24 und nach 48 Wochen untersucht. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Somit wurde das Verzerrungspotential für diesen Endpunkt mit niedrig bewertet.

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

Tabelle 4-29: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert 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)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
	5 mg: 142	81,00 (18,635)	79,44 (18,363)	-1,56 (0,2286)	145	80,94 (15,773)	80,22 (15,516)	-0,72 (0,2263)	-0,84 (0,3217) [-1,47;-0,21]^
	10 mg: 151	80,56 (17,869)	78,31 (17,804)	-2,26 (0,2217)					-1,54 (0,3168) [-2,17;-0,92]^^
24 Wochen**									
	5 mg: 129		78,63 (17,986)	-1,60 (0,2410)	111		80,56 (15,589)	-0,73 (0,2481)	-0,86 (0,3459) [-1,54;-0,19]
	10 mg: 138		78,60 (18,109)	-2,31 (0,2328)					-1,58 (0,3401) [-2,25;-0,92]
48 Wochen**									
	5 mg: 90		80,63 (18,553)	-1,54 (0,2926)	55		81,68 (13,821)	-0,77 (0,3315)	-0,76 (0,4421) [-1,63;0,11]
	10 mg: 107		79,11 (17,817)	-2,41 (0,2766)					-1,64 (0,4317) [-2,48;-0,79]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF; **Repeated measure Analyse; ^p-Wert :0,0091; ^^p-Wert: < 0,0001

SD Standardabweichung SE:Standardfehler; KI: Konfidenzintervall;

° adjustiert bezüglich Ausgangswert

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

Tabelle 4-30: 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 ^o kg (SE)	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert ^o kg (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
	5 mg: 113	80,42 (18,883)	78,92 (18,599)	-1,51 (0,2637)	113	80,72 (16,178)	80,11 (16,031)	-0,62 (0,2637)	-0,89 (0,3729) [-1,62;-0,15]
	10 mg: 129	80,76 (18,300)	78,42 (18,273)	-2,34 (0,2468)					-1,72 (0,3611) [-2,43;-1,01]
24 Wochen**									
	5 mg: 103		77,96 (17,950)	-1,54 (0,2777)	84		81,12 (16,214)	-0,56 (0,2894)	-0,98 (0,4011) [-1,77;-0,19]
	10 mg: 121		78,69 (18,406)	-2,36 (0,2566)					-1,80 (0,3868) [-2,56;-1,04]
48 Wochen**									
	5 mg: 73		79,42 (18,998)	-1,43 (0,3465)	41		82,31 (14,095)	-0,66 (0,4013)	-0,77 (0,5301) [-1,81;0,27]
	10 mg: 94		79,59 (18,313)	-2,46 (0,3146)					-1,80 (0,5096) [-2,80;-0,80]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF; **Repeated measure Analyse

SD Standardabweichung SE:Standardfehler; KI: Konfidenzintervall;

^o adjustiert bezüglich Ausgangswert

Tabelle 4-31: Ergebnisse für Gewichtsveränderung ($\geq 3\%$ und $\geq 5\%$ Gewichtsreduktion) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	Sulfonylharnstoff+ Dapagliflozin N / n (%*) SE	Sulfonylharnstoff+ Placebo N / n (%*) SE	Odds Ratio [95%-KI]
D16900C00005			
24 Wochen			
$\geq 3\%$	5 mg: 45 (31,7); 142 3,901	145 25 (17,4) 3,123	2,23 [1,28;3,89]
	10 mg: 70 (46,3) 151 4,054		4,15 [2,43;7,09]
$\geq 5\%$	5 mg: 20 (14,1) 142 2,919	145 9 (6,4) 1,996	2,48 [1,09;5,65]
	10 mg: 36 (23,7) 151 3,449		4,73 [2,19;10,23]
48 Wochen			
$\geq 3\%$	5 mg: 37 (26,1); 142 3,681	145 22 (15,3) 2,965	1,97 [1,09;3,55]
	10 mg: 69 (45,7) 151 4,054		4,740 [2,70;8,20]
$\geq 5\%$	5 mg: 22 (15,5) 142 3,036	145 12 (8,5) 2,275	2,03 [0,96;4,28]
	10 mg: 31 (20,4) 151 3,279		2,86 [1,41;5,83]

Daten sind inklusive der Patienten, die eine Notfallmedikation bekommen haben.

*Prozentangabe adjustiert bezüglich Ausgangswert

SE:Standardfehler; KI: Konfidenzintervall;

Tabelle 4-32: Ergebnisse für Gewichtsveränderung ($\geq 5\%$ Gewichtsreduktion) aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Sulfonylharnstoff + Dapagliflozin N / n (%*) SE		Sulfonylharnstoff + Placebo N / n (%*) SE		Odds Ratio [95%-KI]
D16900C00005					
24 Wochen					
$\geq 5\%$	5 mg: 113	16 (14,2) 3,278	113	5 (4,6) 1,928	3,56 [1,26;10,09]
	10 mg: 129	33 (25,5) 3,809			7,42 [2,79;19,78]
48 Wochen					
$\geq 5\%$	5 mg: 113	17 (15,0) 3,362	113	6 (5,5) 2,106	3,16 [1,20;8,34]
	10 mg: 129	21 (16,2) 3,246			3,47 [1,35;8,93]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

* Prozentangabe adjustiert bezüglich Ausgangswert

SE:Standardfehler; KI: Konfidenzintervall

Bezogen auf die Gewichtsreduktion ergab sich nach 24 Wochen in beiden Dosisgruppen (5 mg und 10 mg) ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten der Sulfonylharnstoff+Dapagliflozin-Behandlung. Mit dem *Repeated Measurements* Modell war nach 48 Wochen zwischen den beiden Behandlungsgruppen Sulfonylharnstoff+Dapagliflozin gegen Sulfonylharnstoff+Placebo in beiden Dosisgruppen immer noch einen Unterschied zugunsten der Sulfonylharnstoff+Dapagliflozin-Behandlung zu sehen, welcher für die 5 mg-Gruppe nicht mehr signifikant war.

Patienten mit einer Gewichtsreduktion von $\geq 5\%$ waren in den Sulfonylharnstoff+Dapagliflozin-Gruppen signifikant häufiger zu finden. Für die Studienpopulation war der Unterschied in der 5 mg-Gruppe nach 48 Wochen nicht mehr 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 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

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

4.3.1.3.1.3 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 4-33: Operationalisierung von Therapiezufriedenheit (DTSQ)

Studie	Operationalisierung
D1690C00005	Anhand des <i>Diabetes Treatment Satisfaction</i> Fragebogens (DTSQ-Status in Woche 24 / Woche 48) und der Veränderung zum Ausgangswert

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-34: 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
D1690C00005	niedrig	ja	unklar	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Erfassung der Therapiezufriedenheit mittels eines Fragebogens (DTSQ) war in dieser Studie in Woche 24 und 48 ein sekundärer Endpunkt. Die Endpunkterheber waren zu jedem Zeitpunkt bezüglich der Behandlung und Erhebung verblindet. In allen Behandlungsgruppen gab es nach Randomisierung und Medikationsbeginn geringe 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 in der 24-wöchigen Studie wurden mit der LOCF-Methode ersetzt, in der 48-wöchigen Studie fand keine Imputation statt. Es gab daher unterschiedliche Abbrecherraten zwischen den Gruppen, sodass eine adäquate Umsetzung des ITT-Prinzips fraglich ist. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Es wurde aber angenommen, dass das Verzerrungspotential für diesen Endpunkt niedrig ist.

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

Tabelle 4-35: Ergebnisse für Therapiezufriedenheit (DTSQ) aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert [95%-KI]
	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
	5 mg: 130	27,4 (5,93)	30,3 (5,82)	2,7 (0,430) [°]	113	28,3 (7,11)	30,6 (5,51)	2,6 (0,461) [°]	0,01 (0,631) [-1,1;1,3] [^]
	10 mg: 141	27,6 (6,43)	30,9 (4,95)	3,2 (0,412) [°]					0,6 (0,619) [-0,6;1,8] ^{^^}
24 Wochen**									
	5 mg: 127	27,4 (5,99)	30,6 (5,61)	3,2 (6,19)	111	28,4 (7,06)	30,6 (5,54)	2,3 (7,23)	0,90 (0,87960) [-0,82;2,62]
	10 mg: 137	27,6 (6,47)	31,0 (4,91)	3,4 (6,12)					1,10 (0,8627) [-0,59;2,79]
48 Wochen**									
	5 mg: 90	27,3 (5,87)	31,8 (4,82)	4,6 (6,47)	55	29,4 (6,91)	31,6 (5,05)	2,2 (6,34)	2,40 (1,0936) [0,26;4,54]
	10 mg: 105	27,7 (6,65)	32,2 (4,08)	4,4 (6,87)					2,20 (1,0864) [0,07;4,33]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF; ** beobachtete Werte; ° adjustierter Mittelwert + Standardfehler

[^]p-Wert:0,8695

^{^^}p-Wert:0,3575

SD Standardabweichung; KI: Konfidenzintervall

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

Tabelle 4-36: Ergebnisse für Therapiezufriedenheit aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert [95%-KI]
	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen**									
	5 mg: 102	27,0 (6,22)	30,5 (5,82)	3,4 (6,31)	84	27,9 (7,47)	30,1 (5,80)	2,2 (7,92)	1,20 [-0,89;3,29]
	10 mg: 120	27,5 (6,55)	31,1 (4,90)	3,6 (6,05)					1,40 [-0,43;3,23]
48 Wochen**									
	5 mg: 73	26,8 (6,11)	31,6 (5,17)	4,8 (6,93)	41	28,4 (7,35)	30,8 (5,42)	2,4 (6,99)	2,40 [-0,27;5,07]
	10 mg: 92	27,8 (6,75)	32,3 (4,17)	4,5 (7,03)					2,10 [-0,48;4,68]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

SD Standardabweichung; KI: Konfidenzintervall

** beobachtete Werte

Nach 24 sowie nach 48 Wochen der Behandlung mit Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Placebo lassen sich für den Endpunkt Therapiezufriedenheit keine signifikanten Unterschiede feststellen. Diese Analyse beruht jedoch auf den beobachteten Werten der Studienpopulation. In der Sulfonylharnstoff+Placebo-Gruppe füllten nach 48 Wochen über die Hälfte der Patienten den Fragebogen zur Therapiezufriedenheit nicht mehr aus, wohingegen es in der Sulfonylharnstoff+Dapagliflozin noch etwa 70% - 75% der Patienten waren. In der Studienpopulation war der Unterschied zwischen den Behandlungsgruppen nach 48 Wochen etwas größer, wodurch er hier signifikant wurde (MWD [95%-KI]: 2,40 [0,26;4,54] und 2,20 [0,07;4,33]). In der Zielpopulation zeigten sich keine Unterschiede zwischen den Behandlungsoptionen.

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.

4.3.1.3.1.4 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 4-37: Operationalisierung von Blutdrucksenkung

Studie	Operationalisierung
D1690C00005	Anhand der Änderung des systolischen und diastolischen Blutdrucks (mmHg) von Studienbeginn bis Woche 24 (falls der Wert nicht erhoben wurde, wurde der letztmögliche vor Woche 24 (LOCF) betrachtet) bzw. bis Woche 48 (<i>Repeated Measurements</i> Modell). Pro Visite wurden drei Blutdruckmessungen vorgenommen (Durchschnitt von den 3 Messungen wurde für die Analyse verwendet). Zwischen den Messungen wurde eine einminütige Pause eingelegt. Vor den Messungen verbrachte der Patient mindestens 5 Minuten in einer ruhenden Position. Es wurde immer das gleiche Blutdruckmessgerät verwendet.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-38: 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
D1690C00005	niedrig	ja	Ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Blutdrucksenkung war in dieser Studie ein sekundärer Endpunkt und wurde als adjustierte Differenz des mittleren systolischen / diastolischen Blutdrucks in Woche 24 und des Mittelwerts 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, 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. Es gilt zu berücksichtigen, dass eine Änderung der antihypertensive 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. Eine statistische Auswertung der Ergebnisse fand wie geplant statt. Das Verzerrungspotential für diesen Endpunkt wird trotz der ungenauen Begleitmedikation 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 4-39: Ergebnisse für Blutdrucksenkung aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

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)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
systolisch	5 mg: 142	130,9 (15,06)	127,8 (13,73)	-4,0 (0,958)	145	133,3 (13,86)	131,9 (13,49)	-1,2 (0,946)	-2,8 (1,347) [-5,5;-0,2]^
	10 mg: 150	132,4 (13,71)	127,6 (14,20)	-5,0 (0,930)					-3,8 (1,327) [-6,4;-1,2]^^
diastolisch	5 mg: 142	77,9 (8,48)	76,7 (7,85)	-1,7 (0,602)	145	79,8 (7,76)	77,9 (7,79)	-1,4 (0,596)	-0,3 (0,849) [-1,9;1,4]^
	10 mg: 150	78,7 (8,24)	76,0 (8,05)	-2,8 (0,585)					-1,4 (0,836) [-3,0;0,3]^o
24 Wochen**									
systolisch	5 mg: 129	127,41 (13,327)		-4,14 (0,9858)	111	131,86 (13,674)		-0,86 (1,0427)	-3,28 (1,4352) [-6,10;-0,46]
	10 mg: 138	127,38 (13,993)		-4,88 (0,9517)		-4,02 (1,4117) [-6,80;-1,25]			
diastolisch	5 mg: 129	76,61 (7,788)		-1,76 (0,6140)	111	77,67 (7,703)		-1,26 (0,6498)	-0,50 (0,8948) [-2,26;1,26]
	10 mg: 138	76,02 (8,051)		-2,78 (0,5933)		-1,52 (0,8800) [-3,25;0,21]			

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)	
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
48 Wochen**							
systolisch	5 mg: 90	128,89 (15,003)	-3,01 (1,2127)	55	135,06 (12,695)	1,81 (1,5081)	-4,81 (1,9365) [-8,62;-1,01]
	10 mg: 107	127,63 (14,065)	-4,21 (1,1179)				-6,02 (1,8778) [-9,71;-2,32]
diastolisch	5 mg: 90	76,89 (8,784)	-1,41 (0,7712)	55	78,69 (7,813)	0,79 (0,9589)	-2,20 (1,2298) [-4,62;0,22]
	10 mg: 107	76,69 (8,158)	-2,15 (0,7121)				-2,94 (1,1946) [-5,29;-0,59]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

** Repeated Measurements Analyse

^p-Wert : 0,0352

^^p-Wert: 0,0047

°p-Wert: 0,7478

°°p-Wert: 0,0993

SD: Standardabweichung SE: Standardfehler; KI: Konfidenzintervall

#adjustiert bezüglich Ausgangswert

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

Tabelle 4-40: Ergebnisse für Blutdrucksenkung 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 ^o mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert ^o mmHg (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
systolisch	5 mg: 113	129,9 (14,83)	126,6 (13,62)	-4,3 (1,047)	113	134,0 (14,42)	132,1 (13,8)	-1,1 (1,047)	-3,3 (1,484) [-6,2;-0,4]
	10 mg: 129	131,9 (13,52)	126,5 (13,14)	-5,5 (0,978)					-4,4 (1,432) [-7,2;-1,6]
diastolisch	5 mg: 113	77,8 (8,22)	76,7 (7,64)	-1,8 (0,660)	113	80,6 (7,55)	78,8 (7,42)	-1,0 (0,661)	-0,8 (0,937) [-2,6;1,1]
	10 mg: 129	79,1 (8,46)	76,1 (8,17)	-3,0 (0,616)					-2,0 (0,904) [-3,8;-0,2]
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen**									
systolisch	5 mg: 103		126,16 (13,123)	-4,59 (1,1063)	84		132,48 (13,995)	-0,46 (1,1964)	-4,13 (1,6336) [-7,34;-0,92]
	10 mg: 121		126,46 (13,227)	-5,38 (1,0186)					-4,92 (1,5713) [-8,02;-1,83]
diastolisch	5 mg: 103		76,56 (7,462)	-1,90 (0,6761)	84		78,65 (7,252)	-0,75 (0,7322)	-1,15 (0,9992) [-3,11;0,82]
	10 mg: 121		75,88 (8,250)	-3,13 (0,6222)					-2,38 (0,9608) [-4,27;-0,49]

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)	
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
48 Wochen**							
systolisch	5 mg: 73	127,95 (14,824)	-3,37 (1,3388)	41	133,80 (12,283)	0,35 (1,7450)	-3,73 (2,2044) [-8,07;0,62]
	10 mg: 94	127,69 (13,301)	-4,11 (1,1832)				-4,46 (2,1088) [-8,62;-0,31]
diastolisch	5 mg: 73	76,44 (8,160)	-2,0 (0,8338)	41	79,36 (7,729)	0,93 (1,0679)	-2,93 (1,3548) [-5,6;-0,26]
	10 mg: 94	76,95 (7,969)	-2,19 (0,7376)				-3,12 (1,2980) [-5,68;-0,56]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

** Repeated Measurements Analyse

SD: Standardabweichung; SE: Standardfehler; KI: Konfidenzintervall

° adjustiert bezüglich Ausgangswert

Patienten mit einer Sulfonylharnstoff+Dapagliflozin-Behandlung hatten im Mittel nach 24 Wochen eine höhere Blutdrucksenkung (systolisch: -4,3 mmHg in der 5 mg-Gruppe und -5,5 mmHg in der 10 mg-Gruppe /diastolisch: -1,8 mmHg in der 5 mg-Gruppe und -3,0 mmHg in der 10 mg-Gruppe) als Patienten mit einer Sulfonylharnstoff+Placebo-Behandlung (systolisch: -1,1 mmHg / diastolisch: -1,0 mmHg). Der Unterschied zwischen den Behandlungsoptionen war für den systolischen Blutdruck (Ausnahme: Sulfonylharnstoff+Dapagliflozin 5 mg-Gruppe nach 48 Wochen) und den diastolischen Blutdruck (Ausnahme: Sulfonylharnstoff+Dapagliflozin 5 mg-Gruppe nach 24 Wochen) 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.

4.3.1.3.1.5 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 4-41: Operationalisierung von Unerwünschte Ereignissen

Studie	Operationalisierung
D1690C00005	Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis Studienabbruch wegen unerwünschter Ereignisse Anzahl der Patienten mit mindestens einem Anzeichen für das Vorliegen einer Genitalinfektionen bzw. Harnwegsinfektionen jeweils erhoben innerhalb der 24 wöchigen / 48 wöchigen Behandlungsphase

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-42: 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
D1690C00005	niedrig	ja	Ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt „Unerwünschte Ereignisse“ in der Studie war niedrig. Die Erhebung der unerwünschten Ereignisse erfolgte fortlaufend bis zum letzten

definierten Zeitpunkt. 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*) inklusive der Patienten, die eine antidiabetische Notfallmedikation bekommen haben. Eine statistische Auswertung der Ergebnisse fand wie geplant statt.

Tabelle 4-43: 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 (%)
D1690C00005				
24 Wochen				
Sulfonylharnstoff+Dapagliflozin 5 mg	145	70 (48,3)	10 (6,9)	5 (3,4)
Sulfonylharnstoff Dapagliflozin 10 mg	151	76 (50,3)	9 (6,0)	4 (2,6)
Sulfonylharnstoff +Placebo	146	69 (47,3)	7 (4,8)	3 (2,1)
5 mg vs.Placebo OR [95%-KI]		1,04 [0,66;1,65]	1,47 [0,54;3,98]	1,70 [0,40;7,26]
10 mg vs. Placebo OR [95%-KI]		1,13 [0,72;1,78]	1,26 [0,46;3,47]	1,30 [0,29;5,90]
48 Wochen				
Sulfonylharnstoff+Dapagliflozin 5 mg	145	88 (60,7)	16 (11,0)	5 (3,4)
Sulfonylharnstoff+Dapagliflozin 10 mg	151	89 (58,9)	13 (8,6)	4 (2,6)
Sulfonylharnstoff +Placebo	146	81 (55,5)	13 (8,9)	5 (3,4)
5 mg vs.Placebo OR [95%-KI]		1,24 [0,78;1,98]	1,27 [0,59;2,74]	1,01 [0,29;3,56]
10 mg vs. Placebo OR [95%-KI]		1,15 [0,73;1,82]	0,96 [0,43;2,16]	0,77 [0,20;2,92]

KI: Konfidenzintervall; OR: Odds Ratio

Tabelle 4-44: 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 (%)
D1690C00005				
24 Wochen				
Sulfonylharnstoff+Dapagliflozin 5 mg	116	51 (44,0)	8 (6,9)	3 (2,6)
Sulfonylharnstoff+Dapagliflozin 10 mg	129	63 (48,8)	7 (5,4)	1 (0,8)
Sulfonylharnstoff +Placebo	114	53 (46,5)	5 (4,4)	2 (1,8)
5 mg vs.Placebo OR [95%-KI]		0,90 [0,54;1,52]	1,61 [0,51;5,09]	1,49 [0,24;9,07]
10 mg vs. Placebo OR [95%-KI]		1,10 [0,66;1,82]	1,25 [0,39;4,06]	0,44 [0,04;4,89]
48 Wochen				
Sulfonylharnstoff+Dapagliflozin 5 mg	116	67 (57,8)	13 (11,2)	3 (2,6)
Sulfonylharnstoff+Dapagliflozin 10 mg	129	74 (57,4)	11 (8,5)	1 (0,8)
Sulfonylharnstoff+Placebo	114	63 (55,3)	8 (7,0)	3 (2,6)
5 mg vs.Placebo OR [95%-KI]		1,11 [0,66;1,86]	1,67 [0,67;4,20]	0,98 [0,19;4,97]
10 mg vs. Placebo OR [95%-KI]		1,09 [0,66;1,81]	1,24 [0,48;3,19]	0,29 [0,03;2,82]

KI: Konfidenzintervall; OR: Odds Ratio

Die Gesamtrate der Patienten, die ein unerwünschtes Ereignis oder einen Studienabbruch aufgrund unerwünschter Ereignisse hatten war in allen Gruppen sowohl nach 24 als auch nach 48 Wochen vergleichbar. Schwerwiegende unerwünschte Ereignisse traten unter Sulfonylharnstoff+Dapagliflozin etwas häufiger auf. Es gab jedoch keine statistisch signifikanten Unterschiede.

Ergänzend zu den Daten der unerwünschten Ereignisse werden hier auch die Daten zu unerwünschten Ereignissen von besonderem Interesse dargestellt – Anzeichen einer Genital – bzw. Harnwegsinfektion. Die Symptome zu diesem Endpunkt wurden proaktiv von den Prüfern 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 4-45: Ergebnisse für Unerwünschte Ereignisse – Anzeichen einer Genital- bzw. Harnwegsinfektion aus RCT mit dem zu bewertenden Arzneimittel (Studienpopulation)

Studie	N	Anzahl der Patienten mit Anzeichen einer Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen einer Harnwegsinfektion, n (%)
D1690C00005			
24 Wochen			
Sulfonylharnstoff+Dapagliflozin 5 mg	145	9 (6,2)	10 (6,9)
Sulfonylharnstoff +Dapagliflozin 10 mg	151	10 (6,6)	8 (5,3)
Sulfonylharnstoff +Placebo	146	1 (0,7)	9 (6,2)
5 mg vs.Placebo OR [95%-KI]		9,60 [1,20;76,75]	1,13 [0,44;2,86]
10 mg vs. Placebo OR [95%-KI]		10,28 [1,30;81,39]	0,85 [0,32;2,27]
48 Wochen			
Sulfonylharnstoff +Dapagliflozin 5 mg	145	9 (6,2)	11 (7,6)
Sulfonylharnstoff +Dapagliflozin 10 mg	151	13 (8,6)	12 (7,9)
Sulfonylharnstoff +Placebo	146	2 (1,4)	11 (7,5)
5 mg vs.Placebo OR [95%-KI]		4,76 [1,01;22,45]	1,01 [0,42;2,40]
10 mg vs. Placebo OR [95%-KI]		6,78 [1,50;30,61]	1,06 [0,45;2,48]

KI: Konfidenzintervall; OR: Odds Ratio

Tabelle 4-46: Ergebnisse für Unerwünschte Ereignisse – Anzeichen einer Genital- bzw. Harnwegsinfektion aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Anzahl der Patienten mit Anzeichen einer Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen einer Harnwegsinfektion, n (%)
D1690C00005			
24 Wochen			
Sulfonylharnstoff +Dapagliflozin 5 mg	116	7 (6,0)	5 (4,3)
Sulfonylharnstoff +Dapagliflozin 10 mg	129	7 (5,4)	5 (3,9)
Sulfonylharnstoff +Placebo	114	1 (0,9)	6 (5,3)
5 mg vs.Placebo OR [95%-KI]		7,26 [0,88;59,96]	0,81 [0,24;2,74]
10 mg vs. Placebo OR [95%-KI]		6,48 [0,79;53,52]	0,73 [0,22;2,45]
48 Wochen			
Sulfonylharnstoff +Dapagliflozin 5 mg	116	7 (6,0)	6 (5,2)
Sulfonylharnstoff +Dapagliflozin 10 mg	129	8 (6,2)	8 (6,2)
Sulfonylharnstoff +Placebo	114	2 (1,8)	8 (7,0)
5 mg vs.Placebo OR [95%-KI]		3,60 [0,73;17,70]	0,72 [0,24;2,15]
10 mg vs. Placebo OR [95%-KI]		3,70 [0,77;17,81]	0,88 [0,32;2,42]

KI: Konfidenzintervall; OR: Odds Ratio

Bei der Anzahl der Patienten mit Anzeichen einer Genitalinfektion und der Anzahl der Patienten mit Anzeichen einer Harnwegsinfektion ergaben sich für die Zielpopulation sowohl nach 24 als auch nach 48 Wochen keine signifikanten Unterschiede. Auch bei der Betrachtung der gesamten Studienpopulation gab es bei den Anzeichen einer Harnwegsinfektion sowohl nach 24 als auch nach 48 Wochen keine signifikanten Unterschiede, Hinweise auf Genitalinfektionen traten allerdings hier signifikant häufiger unter Sulfonylharnstoff+Dapagliflozin-Behandlung auf.

Dabei gilt zu berücksichtigen, dass die Symptome in allen Fällen nur leicht bis mittelschwer in der Ausprägung waren und die Infektionen mit den gängigen Behandlungsmethoden einfach zu behandeln waren.

Für die Studienpopulation werden die aufgetretenen Anzeichen für Genital- und Harnwegsinfektionen nach 48 Wochen im Folgenden detailliert beschrieben:

In der mit Sulfonylharnstoff+Dapagliflozin (10 mg) behandelten Gruppe traten bei 13 Patienten (8,6%) 19 mal Anzeichen für Genitalinfektion auf, von denen 15 (78,9%) antimikrobiell behandelt wurden. In der Sulfonylharnstoff+Placebo-Gruppe traten bei 2 Patienten 2 mal Anzeichen für Genitalinfektion auf, von denen keine antimikrobiell behandelt wurde.

Die Anzahl an Anzeichen für Harnwegsinfektionen war zwischen den Gruppen etwa gleich. Es gab einen Studienabbruch aufgrund einer Harnwegsinfektion in der mit Sulfonylharnstoff+Dapagliflozin (10 mg) behandelten Gruppe.

In der mit Sulfonylharnstoff+Dapagliflozin (10 mg) behandelten Gruppe traten bei 12 Patienten (7,9%) 13 mal Anzeichen für Harnwegsinfektionen auf, von denen 10 (76,9%) antimikrobiell behandelt wurden. In der Sulfonylharnstoff+Placebo-Gruppe traten bei 11 Patienten (7,5%) 19 mal Anzeichen für Harnwegsinfektionen auf, von denen 18 (94,7%) antimikrobiell behandelt wurden.

Da es sich bei der Anzahl der 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.

In keiner der Behandlungsgruppen traten Infektionen der Nieren 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.

4.3.1.3.1.6 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 4-47: Operationalisierung von Kardiale Ereignisse

Studie	Operationalisierung ^a
D1690C00005	Anzahl der Patienten mit mindestens einem kardialen Ereignis innerhalb von 24 bzw. 48 Wochen. Die Erhebung der kardialen Ereignisse erfolgte im Rahmen der Erhebung der unerwünschten Ereignisse (SOC)

SOC: System Organ Class

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-48: 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
D1690C00005	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt „Kardiale Ereignisse“ in der Studie ist 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*) inklusive der Patienten, die eine antidiabetische Notfallmedikation bekommen haben. Eine statistische Auswertung der Ergebnisse fand wie geplant statt.

Tabelle 4-49: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
D1690C00005		
24 Wochen		
Sulfonylharnstoff +Dapagliflozin 5 mg	145	6 (4,1)
Sulfonylharnstoff +Dapagliflozin 10 mg	151	6 (4,0)
Sulfonylharnstoff +Placebo	146	7 (4,8)
5 mg vs. Placebo OR [95%-KI]		0,86 [0,28;2,62]
10 mg vs. Placebo OR [95%-KI]		0,82 [0,27;2,51]
48 Wochen		
Sulfonylharnstoff +Dapagliflozin 5 mg	145	9 (6,2)
Sulfonylharnstoff +Dapagliflozin 10 mg	151	7 (4,6)
Sulfonylharnstoff +Placebo	146	13 (8,6)
5 mg vs. Placebo OR [95%-KI]		0,68 [0,28;1,64]
10 mg vs. Placebo OR [95%-KI]		0,50 [0,19;1,28]

KI: Konfidenzintervall; OR: Odds Ratio

Es gibt keinen Unterschied in der Häufigkeit des Auftretens von kardialen Ereignissen zwischen den Behandlungsgruppen.

Tabelle 4-50: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Zielpopulation)

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
D1690C00005		
24 Wochen		
Sulfonylharnstoff +Dapagliflozin 5 mg	116	4 (3,4)
Sulfonylharnstoff +Dapagliflozin 10 mg	129	3 (2,3)
Sulfonylharnstoff +Placebo	114	3 (2,6)
5 mg vs.Placebo OR [95%-KI]		1,32 [0,29;6,04]
10 mg vs. Placebo OR [95%-KI]		0,88 [0,17;4,45]
48 Wochen		
Sulfonylharnstoff +Dapagliflozin 5 mg	116	7 (6,0)
Sulfonylharnstoff +Dapagliflozin 10 mg	129	4 (3,1)
Sulfonylharnstoff +Placebo	114	6 (5,3)
5 mg vs.Placebo OR [95%-KI]		1,16 [0,38;3,55]
10 mg vs. Placebo OR [95%-KI]		0,58 [0,16;2,09]

KI: Konfidenzintervall; OR: Odds Ratio

Bezüglich der Gesamtrate der kardialen Ereignisse ergeben sich keine signifikanten Unterschiede zwischen den Behandlungsgruppen sowohl nach 24 als auch nach 48 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 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 bezogen sich ausschließlich auf den Endpunkt HbA1c und auf zusammenfassende sicherheitsrelevante Endpunkte.

Im Folgenden werden zunächst die p-Werte der Interaktionstests der präspezifizierten Subgruppenanalysen gemäß dem Studienbericht für den Endpunkt „Ausmaß der Blutzuckerkontrolle (HbA1c)“ vom Ausgangswert bis Woche 24 angegeben:

Tabelle 4-51: Interaktionstest der präspezifizierten Subgruppenanalysen für den Endpunkt Ausmaß der Blutzuckerkontrolle

Merkmal	Subgruppen	Interaktion p-Wert
Geschlecht	(männlich / weiblich)	0,9164
Alter	(<65 / ≥65/ ≥75 Jahre)	0,9166
Alter und Geschlecht	(Alter der Frau ≤50 / Alter der Frau >50 Jahre)	n.a.
Ethnische Zugehörigkeit	(weiß / schwarz / asiatisch / anders)	0,4089
Ethnizität	(Latino / kein Latino)	n.a.
Body Mass Index (Ausgangswert)	(<25 kg/m ² , ≥25 kg/m ²)	0,6884
Geografische Herkunft	(Latein-Amerika / Europa)	0,4089
HbA1c (Ausgangswert)	(<8%, / ≥8-<9% / ≥9%)	0,2546
eGFR (Ausgangswert)	(<30 / ≥30-<60 / ≥60<90 / ≥90mL/min/1,73 m ²)	0,0115

n.a.: Nicht angegeben aufgrund von zu kleinen Subgruppen (weniger als 10 Patienten in einer Gruppe)

Hierbei fiel der statistisch signifikante Interaktionstest bezüglich der eGFR auf (p=0,0115). Im Rahmen des Zulassungsverfahrens ergab sich eine Zielpopulation von Dapagliflozin, 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 behandelt werden, einschließt (Bristol-Myers Squibb und AstraZeneca GmbH, 2012b).

Zusätzlich wurden für alle Endpunkte Subgruppenanalysen (*post-hoc*) nach

- Geschlecht
- Alter (<65 Jahre und ≥65 bis <75 Jahre)
- Asiatischer bzw. nicht-asiatischer Abstammung

durchgeführt (Bristol-Myers Squibb und AstraZeneca GmbH, 2012a).

Die Wahl der Alters-Cuts beruht darauf, dass die Deutsche Diabetes Gesellschaft (DDG) den älteren Patienten als in der Regel >65 Jahren definiert (Hader et al., 2004). Das entspricht annähernd den oben dargestellten a priori definierten Alters-Cuts. Dadurch konnten Subgruppenanalysen für erwachsene (<65 Jahre) und den älteren (≥ 65 bis <75 Jahre) Patienten dargestellt werden. Da es nach 48 Wochen weniger Werte gab, wurden Interaktionstests für stetige Endpunkte nach 24 Wochen berechnet. Für dichotome Endpunkte wurden Interaktionstests für beide Zeitpunkte angegeben.

Diabetes mellitus kann gemäß der Deutschen Diabetes Gesellschaft 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 konnte, war eine Einteilung nach dem Schweregrad nicht sinnvoll.

Die Bewertung des Nutzens beruhte nur auf einer Studie. Aus diesem Grund wurden keine Meta-Analysen durchgeführt. Die im Folgenden dargestellten Subgruppenanalysen bezogen sich auf die Zielpopulation gemäß Zulassung.

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

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

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

Tabelle 4-52: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert % (SE) [95%-KI]
	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert ^{oo} % (SE)	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert ^{oo} % (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
männlich	5 mg: 57	8,08 (0,728)	7,28 (0,862)	-0,78(0,1020)	56	8,11 (0,757)	7,88 (0,990)	-0,20 (0,1029)	-0,58 (0,1448) [-0,87;-0,30]
	10 mg:56	8,05 (0,713)	7,14 (0,776)	-0,91 (0,1029)					-0,71 (0,1455) [-0,99;-0,42]
weiblich	5 mg: 56	8,30 (0,844)	7,79 (0,861)	-0,48 (0,0921)	55	8,24 (0,733)	8,19 (0,893)	-0,04 (0,0929)	-0,43 (0,1308) [-0,69;-0,17]
	10 mg:73	8,13 (0,863)	7,33 (0,864)	-0,83 (0,0807)					-0,79 (0,1231) [-1,03;-0,55]
Interaktions- test	5 mg vs. Placebo								I ² =0%, p=0,47
	10 mg vs. Placebo								I ² =0%, p=0,67
24 Wochen^o									
männlich	5 mg: 50	7,17 (0,830)		-0,78 (0,1063)	41	7,62 (0,753)	-0,23 (0,1120)	-0,55 (0,1544) [-0,85;-0,24]	
	10 mg:51	7,13 (0,797)		-0,91 (0,1063)				-0,68 (0,1543) [-0,98;-0,37]	

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

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert % (SE) [95%-KI]
	N	Ausgangs- wert Mittel- wert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert ^{oo} % (SE)	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert ^{oo} % (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
weiblich	5 mg: 52	7,74 (0,862)	-0,48 (0,0945)	43	8,04 (0,863)	-0,01 (0,0976)	-0,47 (0,1357) [-0,74;-0,21]		
	10 mg:69	7,33 (0,867)	-0,82 (0,0811)				-0,82 (0,1270) [-1,07;-0,56]		
48 Wochen^o									
männlich	5 mg: 42	7,08 (0,884)	-0,65 (0,1239)	21	7,14 (0,613)	-0,23 (0,1576)	-0,43 (0,2004) [-0,82;-0,03]		
	10 mg:42	6,99 (0,581)	-0,82 (0,1236)				-0,59 (0,1989) [-0,98;-0,20]		
weiblich	5 mg: 31	7,20 (0,571)	-0,52 (0,1199)	20	7,52 (0,771)	0,19 (0,1340)	-0,70 (0,1785) [-1,05;-0,35]		
	10 mg:52	7,16 (0,655)	-0,76 (0,0986)				-0,94 (0,1659) [-1,27;-0,61]		

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

SD: Standardabweichung; SE: Standardfehler

^o Repeated Measurements Modell^{oo} adjustiert bezüglich Ausgangswert

Tabelle 4-53: Ergebnisse für Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach dem Geschlecht

Studie	N	Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle, n (%*)
D1690C00005		
Nach 24 Wochen		
Männlich		
Sulfonylharnstoff +Dapagliflozin 5 mg	57	3 (5,3)
Sulfonylharnstoff +Dapagliflozin 10 mg	56	1 (1,8)
Sulfonylharnstoff +Placebo	58	10 (17,2)
5 mg vs. Placebo OR [95%-KI]		0,27 [0,01;1,03]
10 mg vs. Placebo OR [95%-KI]		0,09 [0,01;0,71]
Nach 24 Wochen		
Weiblich		
Sulfonylharnstoff +Dapagliflozin 5 mg	56	3 (5,4)
Sulfonylharnstoff +Dapagliflozin 10 mg	73	1 (1,4)
Sulfonylharnstoff +Placebo	55	11 (20,0)
5 mg vs. Placebo OR [95%-KI]		0,23 [0,06;0,86]
10 mg vs. Placebo OR [95%-KI]		0,06 [0,01;0,45]
Interaktionstest		
5 mg vs. Placebo		I ² =0%; p=0,87
10 mg vs. Placebo		I ² =0%; p=0,76
Nach 48 Wochen		
Männlich		
Sulfonylharnstoff +Dapagliflozin 5 mg	57	13 (22,5) 5,493
Sulfonylharnstoff +Dapagliflozin 10 mg	56	10 (17,9) 5,074
Sulfonylharnstoff +Placebo	58	27 (43,7) 5,730
5 mg vs. Placebo OR [95%-KI]		0,34 [0,15;0,76]
10 mg vs. Placebo OR [95%-KI]		0,25 [0,11;0,59]
Nach 48 Wochen		
Weiblich		
Sulfonylharnstoff +Dapagliflozin 5 mg	56	20 (33,1) 5,738
Sulfonylharnstoff +Dapagliflozin 10 mg	73	13 (18,9) 4,321
Sulfonylharnstoff +Placebo	55	34 (60,9) 6,103
5 mg vs. Placebo OR [95%-KI]		0,34 [0,20;0,60]
10 mg vs. Placebo OR [95%-KI]		0,13 [0,06;0,30]
Interaktionstest		
5 mg vs. Placebo		I ² =0%; p=0,98
10 mg vs. Placebo		I ² =7%; p=0,30

* Prozentangabe adjustiert bezüglich Ausgangswert des HbA1c (falls vorhanden mit SE)
SE: Standardfehler

Hinsichtlich der Blutzuckerkontrolle ergaben sich keine Hinweise auf eine Effektmodifikation durch das Geschlecht. Der Interaktionstest für die Veränderung des HbA1c nach 24 Wochen

betrug für die einzelnen Dosisgruppen $I^2=0\%$; $p=0,47$ / $I^2=0\%$; $p=0,67$ und der Interaktionstest für Studienabbrüche aufgrund unzureichender Blutzuckerwerte betrug $I^2=0\%$; $p=0,87$ / $I^2=0\%$; $p=0,76$ nach 24 Wochen und $I^2=0\%$; $p=0,98$ / $I^2=7\%$; $p=0,30$ nach 48 Wochen.

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

Tabelle 4-54: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert % (SE) [95%-KI]
	N	Ausgangs- wert Mittel- wert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert ^{oo} % (SE)	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert ^{oo} % (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
Alter <65 Jahre	5 mg: 85	8,23 (0,801)	7,55 (0,933)	-0,65 (0,0804)	81	8,19 (0,689)	8,03 (0,865)	-0,15 (0,0823)	-0,50 (0,1150) [-0,73;-0,27]
	10 mg: 106	8,07 (0,832)	7,22 (0,837)	-0,88 (0,0721)					-0,73 (0,1095) [-0,95;-0,51]
Alter ≥65 und <75 Jahre	5 mg: 28	8,06 (0,730)	7,50 (0,784)	-0,57 (0,1372)	31	8,13 (0,889)	8,05 (1,168)	-0,07 (0,1326)	-0,50 (0,1908) [-0,88;-0,12]
	10 mg: 23	8,20 (0,636)	7,34 (0,804)	-0,83 (0,1517)					-0,76 (0,2012) [-1,16;-0,36]
Interaktions- test	5 mg vs. Placebo								I ² =0%; p=1,00
	10 mg vs. Placebo								I ² =0%; p=0,90
24 Wochen°									
Alter <65 Jahre	5 mg: 76		7,46 (0,932)	-0,65 (0,0824)	58		7,85 (0,734)	-0,13 (0,0875)	-0,53 (0,1201) [-0,76;-0,29]

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

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppenunterschied Mittelwert % (SE) [95%-KI]	
	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert ^{oo} % (SE)	N	Ausgangswert Mittelwert % (SD)		Endwert Mittelwert % (SD)
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo				
	10 mg: 98	7,22 (0,848)	-0,87 (0,0726)				-0,75 (0,1138) [-0,97;-0,52]	
Alter ≥65 und <75 Jahre	5 mg: 26	7,47 (0,765)	-0,56 (0,1492)	26		7,81 (1,039)	-0,10 (0,1475)	-0,46 (0,2097) [-0,88;-0,05]
	10 mg: 22	7,36 (0,817)	-0,81 (0,1634)					-0,71 (0,2203) [-1,15;-0,28]
48 Wochen^o								
Alter <65 Jahre	5 mg: 54	7,07 (0,718)	-0,65 (0,0954)	25		7,5 (0,658)	0,05 (0,1196)	-0,70 (0,1524) [-1,00;-0,40]
	10 mg: 78	7,09 (0,637)	-0,82 (0,0821)					-0,86 (0,1446) [-1,15;-0,58]
Alter ≥65 und <75 Jahre	5 mg: 19	7,29 (0,886)	-0,32 (0,2102)	16		7,05 (0,724)	0,07 (0,2339)	-0,39 (0,3096) [-1,02;0,23]
	10 mg: 16	7,07 (0,588)	-0,67 (0,2264)					-0,75 (0,3292) [-1,41;-0,08]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

SD: Standardabweichung; SE: Standardfehler

^o Repeated Measurements Modell^{oo} adjustiert bezüglich Ausgangswert

Tabelle 4-55: Ergebnisse für Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach dem Alter

Studie	N	Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle, n (%*)
D1690C00005		
Nach 24 Wochen		
Alter <65 Jahre		
Sulfonylharnstoff +Dapagliflozin 5 mg	85	5 (5,8) 2,552
Sulfonylharnstoff +Dapagliflozin 10 mg	106	2 (1,9) 1,322
Sulfonylharnstoff +Placebo	82	18 (21,9) 4,474
5 mg vs. Placebo OR [95%-KI]		0,22 [0,08;0,63]
10 mg vs. Placebo OR [95%-KI]		0,07 [0,02;0,30]
Nach 24 Wochen		
Alter ≥65 und <75 Jahre		
Sulfonylharnstoff +Dapagliflozin 5 mg	28	1 (3,6)
Sulfonylharnstoff +Dapagliflozin 10 mg	23	0 (0,0)
Sulfonylharnstoff +Placebo	31	3 (9,7)
5 mg vs. Placebo OR [95%-KI]		0,35 [0,03;3,53]
10 mg vs. Placebo OR [95%-KI]		0,17 [0,01;3,53]
Interaktionstest		
5 mg vs. Placebo		I ² =0%; p=0,73
10 mg vs. Placebo		I ² =0%; p=0,59
Nach 48 Wochen		
Alter <65 Jahre		
Sulfonylharnstoff +Dapagliflozin 5 mg	85	27 (29,9) 4,740
Sulfonylharnstoff +Dapagliflozin 10 mg	106	17 (16,8) 3,459
Sulfonylharnstoff +Placebo	82	49 (58,4) 5,129
5 mg vs. Placebo OR [95%-KI]		0,31 [0,17;0,59]
10 mg vs. Placebo OR [95%-KI]		0,13 [0,07;0,25]
Nach 48 Wochen		
Alter ≥65 und <75 Jahre		
Sulfonylharnstoff +Dapagliflozin 5 mg	28	6 (21,9) 7,555
Sulfonylharnstoff +Dapagliflozin 10 mg	23	6 (25,3) 9,124
Sulfonylharnstoff +Placebo	31	12 (34,2) 6,076
5 mg vs. Placebo OR [95%-KI]		0,43 [0,14;1,37]
10 mg vs. Placebo OR [95%-KI]		0,56 [0,17;1,82]
Interaktionstest		
5 mg vs. Placebo		I ² =0%; p=0,63
10 mg vs. Placebo		I ² =77,6%; p=0,03

* Prozentangabe adjustiert bezüglich Ausgangswert des HbA1c (falls vorhanden mit SE)

SE: Standardfehler

Hinsichtlich der Blutzuckerkontrolle ergaben sich keine Hinweise auf eine Effektmodifikation durch das Alter (Interaktionstest: I²=0%; p=1,00 / I²=0%; p=0,90 nach 24 Wochen). Nach

48 Wochen zeigte der Interaktionstest bezüglich der Studienabbrüche aufgrund unzureichender glykämischer Kontrolle eine Signifikanz ($I^2=77,6\%$; $p=0,03$) für den Vergleich 10 mg vs. Placebo. Da es für die 5 mg-Dosis keinen Hinweis auf Interaktion gab, die Gruppe der über 65-jährigen Patienten relativ klein war, und alle berechneten OR tendenziell in dieselbe Richtung wiesen (<1 , d.h. Vorteil für Dapagliflozin gegenüber Placebo), handelt es sich hierbei höchstwahrscheinlich um einen nicht relevanten Zufallsbefund, der nicht als Hinweis auf eine fazitrelevante Modifikation des Behandlungseffektes zu interpretieren ist.

4.3.1.3.2.3 Ausmaß der Blutzuckerkontrolle (HbA1c) stratifiziert nach asiatischer bzw. nicht-asiatischer Abstammung - RCT

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

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)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
Asiaten	5 mg: 42	8,22 (0,777)	7,59 (0,788)	-0,63 (0,1094)	38	8,37 (0,741)	8,40 (1,004)	0,07 (0,1155)	-0,70 (0,1591) [-1,02;-0,39]
	10 mg:40	8,02 (0,802)	7,20 (0,788)	-0,87 (0,1130)					-0,95 (0,1627) [-1,27;-0,62]
Nicht-Asiaten	5 mg: 71	8,16 (0,804)	7,50 (0,957)	-0,63 (0,0880)	73	8,07 (0,731)	7,84 (0,870)	-0,24 (0,0867)	-0,39 (0,1236) [-0,63;-0,15]
	10 mg:89	8,13 (0,801)	7,27 (0,850)	-0,85 (0,0785)					-0,61 (0,1170) [-0,84;-0,38]
Interaktions- test	5 mg vs. Placebo								I ² =53,3%; p=0,14
	10 mg vs. Placebo								I ² =63,5%; p=0,09
24 Wochen°									
Asiaten	5 mg: 40		7,56 (0,797)	-0,61 (0,1125)			8,19 (0,976)	0,13 (0,1254)	-0,74 (0,1681) [-1,07;-0,41]
	10 mg:38		7,23 (0,798)	-0,87 (0,1155)					-1,00 (0,1715) [-1,34;-0,66]

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert % (SE) [95%-KI]
	N	Ausgangs- wert Mittel- wert % (SD)	Endwert Mittelwert °° % (SD)	Adjustierte Veränderung Mittelwert % (SE)	N	Ausgangswert Mittelwert % (SD)	Endwert Mittelwert % (SD)	Adjustierte Veränderung Mittelwert°° % (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
Nicht-Asiaten	5 mg: 62	7,39 (0,943)	-0,64 (0,0923)	56	7,66 (0,696)	-0,26 (0,0930)		-0,38 (0,1311) [-0,63;-0,12]	
	10 mg:89	7,26 (0,864)	-0,85 (0,0808)					-0,59 (0,1232) [-0,83;-0,35]	
40 Wochen°									
Asiaten	5 mg:: 29	7,35 (0,648)	-0,47 (0,1263)	13	7,64 (0,901)	0,32 (0,1583)		-0,79 (0,2018) [-1,19;-0,39]	
	10 mg:32	7,22 (0,589)	-0,77 (0,1271)					-1,09 (0,2032) [-1,50;-0,69]	
Nicht-Asiaten (48 Wochen)	5 mg: 46	7,03 (0,677)	-0,69 (0,0955)	32	7,26 (0,655)	-0,26 (0,1064)		-0,44 (0,1424) [-0,72;-0,15]	
	10 mg: 67	7,11 (0,640)	-0,80 (0,0815)					-0,54 (0,1339) [-0,81;-0,28]	

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

SD: Standardabweichung; SE: Standardfehler

° Repeated Measurements Modell

°° adjustiert bezüglich Ausgangswert

Tabelle 4-57: Ergebnisse für Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach asiatischer und nicht-asiatischer Abstammung

Studie	N	Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle, n (%)* SE
D1690C00005		
Nach 24 Wochen		
Asiaten		
Sulfonylharnstoff +Dapagliflozin 5 mg	42	0
Sulfonylharnstoff +Dapagliflozin 10 mg	40	0
Sulfonylharnstoff +Placebo	39	7 (17,9)
5 mg vs. Placebo OR [95%-KI]		0,05 [0,00;0,93]
10 mg vs. Placebo OR [95%-KI]		0,05 [0,00;0,97]
Nach 24 Wochen		
Nicht-Asiaten		
Sulfonylharnstoff +Dapagliflozin 5 mg	71	6 (8,2) 3,293
Sulfonylharnstoff +Dapagliflozin 10 mg	89	2 (2,3) 1,571
Sulfonylharnstoff +Placebo	74	14 (19,5) 4,419
5 mg vs. Placebo OR [95%-KI]		0,40 [0,14;1,10]
10 mg vs. Placebo OR [95%-KI]		0,10 [0,02;0,45]
Interaktionstest		
5 mg vs. Placebo		I ² =41,5%; p=0,19
10 mg vs. Placebo		I ² =0%; p=0,71
Nach 48 Wochen		
Asiaten		
Sulfonylharnstoff +Dapagliflozin 5 mg	42	13 (30,5) 6,569
Sulfonylharnstoff +Dapagliflozin 10 mg	40	8 (24,8) 5,685
Sulfonylharnstoff +Placebo	39	24 (56,8) 7,274
5 mg vs. Placebo OR [95%-KI]		0,28 [0,11;0,70]
10 mg vs. Placebo OR [95%-KI]		0,16 [0,06;0,43]
Nach 48 Wochen		
Nicht-Asiaten		
Sulfonylharnstoff +Dapagliflozin 5 mg	71	20 (26,9) 5,123
Sulfonylharnstoff +Dapagliflozin 10 mg	89	15 (16,7) 3,952
Sulfonylharnstoff +Placebo	74	37 (50,7) 5,172
5 mg vs. Placebo OR [95%-KI]		0,39 [0,20;0,78]
10 mg vs. Placebo OR [95%-KI]		0,20 [0,10;0,42]
Interaktionstest		
5 mg vs. Placebo		I ² =0%; p=0,57
10 mg vs. Placebo		I ² =0%; p=0,68

* Prozentangabe adjustiert bezüglich Ausgangswert des HbA1c (falls vorhanden mit SE)
SE: Standardfehler

Bezüglich des Ausmaßes der Blutzuckerkontrolle anhand des HbA1c ergab sich ein Hinweis auf eine Interaktion durch die Abstammung der Patienten (Interaktionstest: I²=53,3%; p=0,14)

und $I^2=63,5\%$; $p=0,09$) sowohl in der 5 mg als auch in der 10 mg Dosisgruppe. Auch für die Operationalisierung Studienabbrüche oder Notfallmedikation aufgrund unzureichender glykämischer Kontrolle ergab sich ein Hinweis auf eine Effektmodifikation durch die Abstammung der Patienten ($I^2=41,5\%$; $p=0,19$), jedoch nur nach 24 Wochen und zwischen Sulfonylharnstoff+Dapagliflozin 5 mg versus Sulfonylharnstoff+Placebo. Der Effekt war allerdings trotz dieses Befundes in jeder Subgruppe gleichgerichtet (eine statistisch signifikante Verbesserung der Blutzuckersenkung zugunsten von Sulfonylharnstoff+Dapagliflozin), qualitative Unterschiede zwischen den Gruppen fanden sich nicht. Als Hinweis auf eine fazitrelevante Interaktion kann dieser Befund daher nicht interpretiert werden.

4.3.1.3.2.4 Hypoglykämien stratifiziert nach dem Geschlecht – RCT

Tabelle 4-58: Ergebnisse für Hypoglykämien stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertende Arzneimittel		Vergleichstherapie		Odds Ratio
D1690C00005	N / n (%)		N / n (%)		[95%-KI]
24 Wochen	Sulfonylharnstoff+Dapagliflozin		Sulfonylharnstoff+Placebo		
Männlich					
Patienten mit mind. einer Hypoglykämie	5 mg: 58	6 (10,3)	58	5 (8,6)	1,22 [0,35;4,26]
	10 mg:56	7 (12,5)			1,51 [0,45;5,09]
schwerwiegend	5 mg: 58	0	58	0	Nicht berechenbar
	10 mg:56	0			Nicht berechenbar
leicht	5 mg: 58	5 (8,6)	58	2 (3,4)	2,64 [0,49;14,21]
	10 mg:56	7 (12,5)			4,00 [0,79;20,16]
andere	5 mg: 58	1 (1,7)	58	4 (6,9)	0,24 [0,03;2,19]
	10 mg:56	0			0,11 [0,01;2,04]
24 Wochen	Sulfonylharnstoff +Dapagliflozin		Sulfonylharnstoff+Placebo		
Weiblich					
Patienten mit mind. einer Hypoglykämie	5 mg: 58	2 (3,4)	56	0	5,00 [0,23;106,50]
	10 mg:73	5 (6,8)			9,07 [0,49;167,62]
schwerwiegend	5 mg: 58	0	56	0	Nicht berechenbar
	10 mg:73	0			Nicht berechenbar
leicht	5 mg: 58	1 (1,7)	56	0	2,95 [0,12;73,90]
	10 mg:73	3 (4,1)			5,61 [0,28;110,87]
andere	5 mg: 58	2 (3,4)	56	0	5,00 [0,23;106,50]

Studie D1690C00005	Zu bewertende Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
	10 mg: 73 3 (4,1)		5,61 [0,28;110,87]
Interaktionstest¹ 5 mg vs. Placebo 10 mg vs. Placebo	I ² =0%, p=0,40 I ² =19%, p=0,27		
48 Wochen Männlich	Sulfonylharnstoff +Dapagliflozin	Sulfonylharnstoff+ Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 58 7 (12,1)	58 7 (12,1)	1,00 [0,33;3,06]
	10 mg: 56 10 (17,9)		1,58 [0,56;4,50]
schwerwiegend	5 mg: 58 0	58 0	Nicht berechenbar
	10 mg: 56 0		Nicht berechenbar
leicht	5 mg: 58 6 (10,3)	58 2 (3,4)	3,23 [0,62;16,73]
	10 mg: 56 8 (14,3)		4,67 [0,95;23,04]
andere	5 mg: 58 1 (1,7)	58 6 (10,3)	0,15 [0,02;1,31]
	10 mg: 56 2 (3,6)		0,32 [0,06;1,66]
48 Wochen Weiblich	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff+ Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 58 5 (8,6)	56 1 (1,8)	5,19 [0,59;45,90]
	10 mg: 73 7 (9,6)		5,83 [0,70;48,877]
schwerwiegend	5 mg: 58 0	56 0	Nicht berechenbar
	10 mg: 73 0		Nicht berechenbar
leicht	5 mg: 58 3 (5,2)	56 1 (1,8)	3,00 [0,30;29,74]
	10 mg: 73 5 (6,8)		4,04 [0,46;35,64]
andere	5 mg: 58 3 (5,2)	56 0	7,13 [0,36;141,18]
	10 mg: 73 3 (4,1)		5,61 [0,28;110,87]
Interaktionstest¹ 5 mg vs. Placebo 10 mg vs. Placebo	I ² =42,4%, p=0,19 I ² =14,1%; p=0,28		

¹ Der Interaktionstest bezieht sich auf Patienten mit mindestens einer Hypoglykämie

Der Interaktionstest ergab nach 24 Wochen keine signifikanten Interaktionen (Sulfonylharnstoff+Dapagliflozin 5 mg versus Sulfonylharnstoff+Placebo: I²=0%, p=0,40 und Sulfonylharnstoff+Dapagliflozin 10 mg versus Sulfonylharnstoff+Placebo I²=19%, p=0,27).

Nach Woche 48 zeigte der Interaktionstest ein signifikantes Ergebnis im Vergleich Sulfonylharnstoff+Dapagliflozin 5 mg versus Sulfonylharnstoff+Placebo ($I^2=42,4\%$, $p=0,19$), aber aufgrund des gleichgerichteten Effekts (kein Unterschied) und der geringen Ereignisse in den Subgruppen ergab sich daraus keine Änderung des aus der Bewertung der Gesamtpopulation abgeleiteten Fazits. Im Vergleich Sulfonylharnstoff+Dapagliflozin 10 mg versus Sulfonylharnstoff+Placebo zeigte sich keine signifikante Interaktion bezüglich des Geschlechts ($I^2=14,1\%$; $p=0,28$).

Es wurden keine weiteren Interaktionstest für die Aufspaltung in schwerwiegend, leicht und andere aufgrund der kleinen Ereigniszahlen durchgeführt.

4.3.1.3.2.5 Hypoglykämien stratifiziert nach dem Alter– RCT

Tabelle 4-59: Ergebnisse für Hypoglykämien stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertende Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
D1690C00005			
24 Wochen Alter <65 Jahre	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff+ Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 87 7 (8,0)	82 4 (4,9)	1,71 [0,48;6,06]
	10 mg:106 11 (10,4)		2,26 [0,69;7,37]
schwerwiegend	5 mg: 87 0	82 0	Nicht berechenbar
	10 mg:106 0		Nicht berechenbar
leicht	5 mg: 87 6 (6,9)	82 1 (1,2)	6,00 [0,71;50,96]
	10 mg:106 9 (8,5)		7,52 [0,93;60,58]
andere	5 mg: 87 2 (2,3)	82 3 (3,7)	0,62 [0,10;3,81]
	10 mg:106 3 (2,8)		0,77 [0,15;3,90]
24 Wochen Alter ≥65 und <75 Jahre	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff+ Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 29 1 (3,4)	32 1 (3,1)	1,11 [0,07;18,55]
	10 mg:23 1 (4,3)		1,41 [0,08;23,76]
schwerwiegend	5 mg: 29 0	32 0	Nicht berechenbar
	10 mg:23 0		Nicht berechenbar
leicht	5 mg: 29 0	32 1 (3,1)	0,36 [0,01;9,09]
	10 mg:23 1 (4,3)		1,41 [0,08;23,76]
andere	5 mg: 29 1 (3,4)	32 1 (3,1)	1,11 [0,07;18,55]
	10 mg:23 0		0,45 [0,02;11,46]
Interaktionstest¹ 5 mg vs.Placebo 10 mg vs. Placebo		I ² =0%, p=0,78 I ² =0%, p=0,83	

Studie D1690C00005	Zu bewertende Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
48 Wochen Alter <65 Jahre	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff+ Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 87 11 (12,6)	82 7 (8,5)	1,55 [0,57;4,21]
	10 mg: 106 13 (12,3)		1,50 [0,57;3,94]
schwerwiegend	5 mg: 87 0	82 0	Nicht berechenbar
	10 mg: 106 0		Nicht berechenbar
leicht	5 mg: 87 9 (10,3)	82 2 (2,4)	4,62 [0,97;22,04]
	10 mg: 106 9 (8,5)		3,71 [0,78;17,67]
andere	5 mg: 87 3 (3,4)	82 5 (6,1)	0,55 [0,13;2,38]
	10 mg: 106 5 (4,7)		0,76 [0,21;2,73]
48 Wochen Alter ≥65 und <75 Jahre	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff+ Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 29 1 (3,4)	32 1 (3,1)	1,11 [0,07;18,55]
	10 mg: 23 4 (17,4)		6,53 [0,68;62,82]
schwerwiegend	5 mg: 29 0	32 0	Nicht berechenbar
	10 mg: 23 0		Nicht berechenbar
leicht	5 mg: 29 0	32 1 (3,1)	0,36 [0,01;9,09]
	10 mg: 23 4 (17,4)		6,53 [0,68;62,82]
andere	5 mg: 29 1 (3,4)	32 1 (3,1)	1,11 [0,07;18,55]
	10 mg: 23 0		0,45 [0,02;11,46] ⁺
Interaktionstest¹ 5 mg vs. Placebo 10 mg vs. Placebo	I ² =0%, p=0,76 I ² =27,1%;p=0,24		

¹ Der Interaktionstest bezieht sich auf Patienten mit mindestens einer Hypoglykämie

Es ergaben sich hier keine signifikanten Interaktionstest nach 24 Wochen / 48 Woche für die Gesamtrate der Hypoglykämien (Sulfonylharnstoff+Dapagliflozin 5 mg versus Sulfonylharnstoff+Placebo I²=0%, p=0,78 / I²=0%, p=0,76). Auch für Sulfonylharnstoff+Dapagliflozin 10 mg versus Sulfonylharnstoff+Placebo ergaben sich keine Interaktionen. (I²=0%, p=0,83 und I²=27,1%;p=0,24).

Es wurden keine weiteren Interaktionstest für die Aufspaltung in schwerwiegend, leicht und andere aufgrund der kleinen Ereigniszahlen durchgeführt.

4.3.1.3.2.6 Hypoglykämien stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT

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

Studie	Zu bewertende Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
D1690C00005			
24 Wochen Asiaten	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff+ Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 43 4 (9,3)	39 2 (5,1)	1,90 [0,33;10,98]
	10 mg:40 4 (10,0)		2,06 [0,35;11,93]
schwerwiegend	5 mg: 43 0	39 0	Nicht berechenbar
	10 mg:40 0		Nicht berechenbar
leicht	5 mg: 43 3 (7,0)	39 1 (2,6)	2,85 [0,28;28,61]
	10 mg:40 3 (7,5)		3,08 [0,31;30,98]
andere	5 mg: 43 2 (4,7)	39 1 (2,6)	1,85 [0,16;21,28]
	10 mg:40 1 (2,5)		0,97 [0,06;16,14]
24 Wochen Nicht-Asiaten	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff +Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 73 4 (5,5)	75 3 (4,0)	1,39 [0,30;6,44]
	10 mg:89 8 (9,0)		2,37 [0,61;9,28]
schwerwiegend	5 mg: 73 0	75 0	Nicht berechenbar
	10 mg:89 0		Nicht berechenbar
leicht	5 mg: 73 3 (4,1)	75 1 (1,3)	3,17 [0,32;31,21]
	10 mg:89 7 (7,9)		6,32 [0,76;52,56]
andere	5 mg: 73 1 (1,4)	75 3 (4,0)	0,33 [0,03;3,28]
	10 mg:89 2 (2,2)		0,55 [0,09;3,39]
Interaktionstest¹ 5 mg vs. Placebo 10 mg vs. Placebo		I ² =0%; p=0,79 I ² =0%; p=0,90	

Studie D1690C00005	Zu bewertende Arzneimittel N / n (%)	Vergleichstherapie N / n (%)	Odds Ratio [95%-KI]
48 Wochen Asiaten	Sulfonylharnstoff+Dapagliflozin	Sulfonylharnstoff +Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 43 7 (16,3)	39 4 (10,3)	1,70 [0,46;6,33]
	10 mg:40 6 (15,0)		1,54 [0,40;5,96]
schwerwiegend	5 mg: 43 0	39 0	Nicht berechenbar
	10 mg:40 0		Nicht berechenbar
leicht	5 mg: 43 5 (11,6)	39 2 (5,1)	2,43 [0,44;13,34]
	10 mg:40 4 (10,0)		2,06 [0,35;11,93]
andere	5 mg: 43 3 (7,0)	39 2 (5,1)	1,39 [0,22;8,77]
	10 mg:40 2 (5,0)		0,97 [0,13;7,28]
48 Wochen Nicht-Asiaten	Sulfonylharnstoff +Dapagliflozin	Sulfonylharnstoff+ Placebo	
Patienten mit mind. einer Hypoglykämie	5 mg: 73 5 (6,8)	75 4 (5,3)	1,31 [0,34;5,07]
	10 mg:89 11 (12,4)		2,50 [0,76;8,22]
schwerwiegend	5 mg: 73 0	75 0	Nicht berechenbar
	10 mg:89 0		Nicht berechenbar
leicht	5 mg: 73 4 (5,5)	75 1 (1,3)	4,29 [0,47;39,33]
	10 mg:89 9 (10,1)		8,32 [1,03;67,31]
andere	5 mg: 73 1 (1,4)	75 4 (5,3)	0,25 [0,03;2,26]
	10 mg:89 3 (3,4)		0,62 [0,13;2,86]
Interaktionstest¹ 5 mg vs. Placebo 10 mg vs. Placebo		$I^2=0\%$; $p=0,78$ $I^2=0\%$; $p=0,60$	

¹ Der Interaktionstest bezieht sich auf Patienten mit mindestens einer Hypoglykämie

Der Interaktionstest nach 24 Wochen zeigte in den Gruppen Sulfonylharnstoff+Dapagliflozin 5mg / 10mg versus Sulfonylharnstoff+Placebo bezüglich der Gesamtrate von Hypoglykämien kein signifikantes Ergebnis ($I^2=0\%$; $p=0,79$ / $I^2=0\%$; $p=0,90$); auch nach 48 Wochen zeigte sich ebenfalls kein signifikantes Resultat ($I^2=0\%$; $p=0,78$ / $I^2=0\%$; $p=0,60$).

Es wurden keine weiteren Interaktionstest für die Aufspaltung in schwerwiegend, leicht und andere aufgrund der kleinen Ereigniszahlen durchgeführt.

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

4.3.1.3.2.7 Gewichtsveränderung stratifiziert nach dem Geschlecht– RCT

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

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert kg (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)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
männlich	5 mg: 57	88,62 (19,141)	87,10 (18,756)	-1,51 (0,4161)	58	85,73 (13,627)	85,22 (13,754)	-0,53 (0,4137)	-0,98 (0,5869) [-2,14;0,18]
	10 mg 56	89,47 (15,370)	86,93 (15,486)	-2,53 (0,4200)					-2,00 (0,5903) [-3,16;-0,83]
weiblich	5 mg: 56	72,08 (14,582)	70,59 (14,363)	-1,53 (0,3269)	55	75,44 (17,072)	74,72 (16,605)	-0,70 (0,3295)	-0,83 (0,4647) [-1,75;0,08]
	10 mg 73	74,07 (17,621)	71,89 (17,563)	-2,18 (0,2858)					-1,49 (0,4363) [-2,35;-0,63]
Interaktions- test	5 mg versus Placebo								I ² =0%; p=0,84
	10 mg versus Placebo								I ² =0%; p=0,48
24 Wochen**									
männlich	5 mg: 51		86,81 (17,857)	-1,50 (0,4504)	41		86,39 (13,372)	-0,52 (0,4737)	-0,89 (0,6540) [-2,27;0,31]
	10 mg: 52		87,13 (15,741)	-2,58 (0,4511)					-2,06 (0,6553) [-3,35;-0,76]
weiblich	5 mg: 52		69,29 (13,300)	-1,59 (0,3369)	43		76,09 (17,213)	-0,60 (0,3472)	-0,99 (0,4853) [-1,95;-0,03]

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

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppenunterschied Mittelwert kg (SE) [95%-KI]
	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert kg (SE)	N	Ausgangswert Mittelwert kg (SD)	
D1690C00005	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
	10 mg: 69	72,32 (17,789)	-2,17 (0,2893)				
48 Wochen**							
männlich	5 mg: 42	87,51 (18,588)	-1,30 (0,500)	21	88,96 (11,116)	-0,59 (0,5916)	-0,72 (0,7747) [-2,25;0,81]
	10 mg: 42	86,76 (16,753)	-2,96 (0,5002)				
weiblich	5 mg: 31	68,45 (13,337)	-1,65 (0,4734)	20	75,34 (13,713)	-0,41 (0,5316)	-1,24 (0,7146) [-2,65;0,17]
	10 mg: 52	73,80 (17,588)	-2,12 (0,3904)				

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

**Repeated measure Analyse;

SD Standardabweichung SE:Standardfehler

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

Studie	Sulfonylharnstoff + Dapagliflozin N / n (%*) SE	Sulfonylharnstoff+ Placebo N / n (%*) SE	Odds Ratio [95%-KI]
D16900C00005			
24 Wochen Männlich			
$\geq 5\%$	5 mg: 57 7 (12,2) 4,342	58 2 (4,6) 2,354	3,92 [0,78;19,75]
	10 mg: 56 13 (23,1) 5,649		8,47 [1,81;39,52]
24 Wochen Weiblich			
$\geq 5\%$	5 mg: 56 9 (16,7) 4,898	55 3 (5,3) 3,060	3,32 [0,85;13,00]
	10 mg: 73 20 (27,1) 5,105		6,54 [1,83;23,35]
Interaktionstest	5 mg		$I^2=0\%$, $p=0,88$
	10 mg		$I^2=0\%$, $p=0,80$
48 Wochen Männlich			
$\geq 5\%$	5 mg: 57 9 (15,8) 4,824	58 1 (1,7) 1,709	10,96 [1,31;87,39]
	10 mg: 56 9 (16,7) 4,887		10,91 [1,33;89,29]
48 Wochen Weiblich			
$\geq 5\%$	5 mg: 56 8 (14,2) 4,675	55 5 (8,5) 3,825	1,67 [0,51;5,45]
	10 mg: 73 12 (16,4) 4,337		1,97 [0,65;5,96]
Interaktionstest	5 mg versus Placebo		$I^2=56,1\%$; $p=0,13$
	10 mg versus Placebo		$I^2=49,9\%$; $p=0,16$

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

* Prozentangabe adjustiert bezüglich Ausgangswert

SE:Standardfehler

Hinsichtlich der Gewichtsveränderung nach 24 Wochen ($I^2=0\%$; $p=0,84$ / $I^2=0\%$; $p=0,48$) und der Anzahl der Patienten mit einer Gewichtsveränderung von mehr als 5% nach 24 Wochen ergaben sich keine Hinweise auf eine Effektmodifikation durch das Geschlecht. Nach Woche 48 waren jedoch die Interaktionstests statistisch signifikant ($I^2=56,1\%$; $p=0,13$ und

$I^2=49,9\%$; $p=0,16$). Da die Effekte jedoch in beiden betrachteten Subgruppen in dieselbe Richtung wiesen (>1 , d.h. Vorteil für Dapagliflozin) und die Konfidenzintervalle der Effektschätzer in den Subgruppen aufgrund der relativ niedrigen Anzahl von betrachteten Ereignissen sehr breit waren, handelt es sich hier höchstwahrscheinlich um einen klinisch nicht relevanten Zufallsbefund, der in der Gesamtbewertung nicht berücksichtigt wurde.

4.3.1.3.2.8 Gewichtsveränderung stratifiziert nach dem Alter– RCT

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

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppen- unterschied Mittelwert 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)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
Alter <65 Jahre	5 mg: 85	82,25 (19,832)	80,85 (19,506)	-1,40 (0,3117)	82	81,51 (16,938)	81,06 (16,619)	-0,46 (0,3174)	-0,94 (0,4448) [-1,82;-0,07]
	10 mg: 106	80,74 (18,566)	78,49 (18,512)	-2,27 (0,2794)					-1,82 (0,4226) [-2,65;-0,99]
Alter ≥65 und <75 Jahre	5 mg: 28	74,87 (14,597)	73,07 (14,284)	-1,80 (0,4825)	31	78,64 (14,010)	77,57 (14,304)	-1,06 (0,4575)	-0,74 (0,6663) [-2,06;0,58]
	10 mg: 23	80,83 (17,414)	78,12 (17,302)	-2,71 (0,5338)					-1,65 (0,7011) [-3,03;-0,26]
Interaktions- test	5 mg vs. Placebo								I ² =0%;p=0,80
	10 mg vs. Placebo								I ² =0%;p=0,85
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen°									
Alter <65 Jahre	5 mg: 77		79,73 (18,683)	-1,41 (0,3291)	58		82,76 (16,543)	-0,33 (0,3505)	-1,08 (0,4808) [-2,03;-0,13]

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

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppen- unterschied Mittelwert 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)
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo				
	10 mg: 98		78,82 (18,738)	-2,31 (0,2915)				-1,98 (0,4559) [-2,88;-1,08]
Alter ≥65 und <75 Jahre	5 mg: 26		72,73 (14,673)	-1,85 (0,5050)	26	77,46 (15,123)	-1,04 (0,4952)	-0,82 (0,7081) [-2,23;0,59]
	10 mg: 23		78,12 (17,302)	-2,67 (0,5436)				-1,64 (0,7348) [-3,10;-0,17]
48 Wochen°								
Alter <65 Jahre	5 mg: 54		81,19 (20,079)	-1,20 (0,4169)	25	82,64 (15,036)	-0,25 (0,5106)	-0,95 (0,6590) [-2,25;0,35]
	10 mg: 78		80,26 (18,752)	-2,30 (0,3613)				-2,05 (0,6254) [-3,28;-0,82]
Alter ≥65 und <75 Jahre	5 mg: 19		74,37 (14,842)	-2,05 (0,5854)	16	81,81 (12,946)	-1,36 (0,6192)	-0,69 (0,8561) [-2,40;1,01]
	10 mg: 16		76,33 (16,143)	-3,16 (0,6283)				-1,79 (0,8786) [-3,54;-0,04]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

SD: Standardabweichung; SE: Standardfehler

° Repeated Measurements Modell

°° adjustiert bezüglich Ausgangswert

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

Studie	Sulfonylharnstoff + Dapagliflozin N / n (%*) SE	Sulfonylharnstoff + Placebo N / n (%*) SE	Odds Ratio [95%-KI]
D16900C00005			
24 Wochen			
Alter <65 Jahre			
$\geq 5\%$	5 mg: 85 11 (12,9) 3,638	82 2 (3,1) 1,676	5,95 [1,28;27,72]
	10 mg: 25 (23,0) 106 4,078		12,35 [2,83;53,86]
24 Wochen			
Alter ≥ 65 und <75 Jahre			
$\geq 5\%$	5 mg: 28 5 (18,3) 7,227	31 3 (9,8) 5,310	2,03 [0,44;9,41]
	10 mg: 23 8 (35,2) 9,924		4,98 [1,15;21,60]
Interaktionstest			$I^2=0\%$; $p=0,33$
			$I^2=0\%$; $p=0,39$
48 Wochen			
Alter <65 Jahre			
$\geq 5\%$	5 mg: 85 10 (11,8) 3,493	82 3 (4,1) 2,071	3,51 [0,93;13,25]
	10 mg: 18 (16,8) 106 3,645		5,39 [1,53;18,98]
48 Wochen			
Alter ≥ 65 und <75 Jahre			
$\geq 5\%$	5 mg: 28 7 (25,0)	31 3(9,7)	3,11 [0,72;13,48]
	10 mg: 23 3(13,0)		1,40 [0,26;7,66]
Interaktionstest	5 mg versus Placebo		$I^2=0\%$; $p=0,90$
	10 mg versus Placebo		$I^2=35,8\%$; $p=0,21$

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

* Prozentangabe adjustiert bezüglich dem Ausgangswert (falls durchschnittlich $n \leq 5$ in den Behandlungsarmen wurde keine Adjustierung vorgenommen)

SE:Standardfehler

Hinsichtlich der Gewichtsveränderung nach 24 Wochen ($I^2=0\%$; $p=0,80$ / $I^2=0\%$; $p=0,85$) und der Anzahl der Patienten mit einer Gewichtsreduktion von mehr als 5% nach 24 Wochen ($I^2=0\%$; $p=0,33$ / $I^2=0\%$; $p=0,39$) und nach 48 Wochen ($I^2=0\%$; $p=0,90$ und $I^2=35,8\%$; $p=0,21$) ergaben sich keine Hinweise auf eine Effektmodifikation durch das Alter.

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

4.3.1.3.2.9 Gewichtsveränderung stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT

Tabelle 4-65: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach asiatischer und nicht-asiatischer Abstammung

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert kg (SE) [95%-KI]
	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert ^{oo} kg (SE)	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert ^{oo} kg (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen*									
Asiaten	5 mg: 42	64,60 (11,470)	63,42 (11,119)	-1,23 (0,3662)	39	66,77 (10,700)	66,36 (10,595)	-0,38 (0,3795)	-0,85 (0,5279) [-1,89;0,20]
	10 mg 40	66,92 (13,238)	63,92 (12,527)	-2,97 (0,3748)					-2,59 (0,5330) [-3,64;-1,53]
Nicht-Asiaten	5 mg: 71	89,78 (15,947)	88,09 (15,859)	-1,68 (0,3541)	74	88,08 (13,529)	87,35 (13,478)	-0,73 (0,3466)	-0,94 (0,4956) [-1,92;0,03]
	10 mg 74	86,98 (16,830)	84,94 (16,602)	-2,06 (0,3164)					-1,32 (0,4692) [-2,25;-0,40]
Interaktions- test	5 mg versus Placebo								I ² =0%; p=0,89
	10 mg versus Placebo								I ² =61,3%; p=0,11
24 Wochen**									
Asiaten	5 mg: 40		63,46 (11,393)	-1,29 (0,3880)	28		66,94 (11,202)	-0,17 (0,4158)	-1,12 (0,5694) [-2,24;0,01]
	10 mg: 38		64,32 (12,722)	-2,97 (0,3953)					-2,80 (0,5734) [3,93;-1,66]
Nicht-Asiaten	5 mg: 63		87,17 (15,049)	-1,70 (0,3736)	56		88,21 (13,474)	-0,75 (0,3796)	-0,95 (0,5326) [-2,00;0,10]

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

Studie	Zu bewertendes Arzneimittel			Vergleichstherapie			Gruppenunterschied Mittelwert kg (SE) [95%-KI]
	N	Ausgangswert Mittelwert kg (SD)	Endwert Mittelwert kg (SD)	Adjustierte Veränderung Mittelwert ^{oo} kg (SE)	N	Ausgangswert Mittelwert kg (SD)	
D1690C00005	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
	10 mg: 83	85,27 (16,831)	-2,10 (0,3270)				-1,35 (0,5010) [-2,34;-0,36]
40 Wochen**							
Asiaten	5 mg: 29	63,33 (11,650)	-1,61 (0,4058)	13	67,70 (10,422)	-0,14 (0,4873)	-1,47 (0,6349) [-2,73;-0,22]
	10 mg: 32	64,77 (12,763)	-3,21 (0,4086)				-3,08 (0,6353) [-4,33;1,82]
Nicht-Asiaten (48 Wochen)	5 mg: 46	88,69 (15,996)	-1,60 (0,4661)	32	86,69 (11,371)	-0,90 (0,5073)	-0,70 (0,6889) [-2,06;0,66]
	10 mg: 67	85,37 (16,860)	-2,14 (0,3989)				-1,24 (0,6453) [-2,51;0,03]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

**Repeated measure Analyse;

SD Standardabweichung SE:Standardfehler

^{oo} adjustiert bezüglich Ausgangswert

Tabelle 4-66: Ergebnisse für Gewichtsverlust ($\geq 5\%$) aus RCT mit dem zu bewertenden Arzneimittel, stratifiziert nach asiatischer und nicht-asiatischer Herkunft

Studie	Sulfonylharnstoff+ Dapagliflozin N / n (%*) SE	Sulfonylharnstoff+ Placebo N / n (%*) SE	Odds Ratio [95%-KI]
D16900C00005			
24 Wochen Asiaten			
$\geq 5\%$	5 mg: 42 6 (15,0) 5,326	39 1 (2,7) 2,539	6,00 [0,73;55,22]
	10 mg: 40 14 (35,5) 7,448		20,46 [2,53;165,29]
24 Wochen Nicht-Asiaten			
$\geq 5\%$	5 mg: 71 10 (14,1) 4,128	74 4 (5,5) 2,624	2,91 [0,87;9,75]
	10 mg: 89 19 (21,1) 4,339		4,82 [1,56;14,88]
Interaktionstest			$I^2=0\%$; $p=0,54$
			$I^2=29,9\%$; $p=0,23$
48 Wochen Asiaten			
$\geq 5\%$	5 mg: 42 7 (15,9) 5,707	39 1 (3,0) 2,559	7,60 [0,89;64,92]
	10 mg: 40 11 (27,0) 7,013		14,41 [1,76;118,10]
48 Wochen Nicht-Asiaten			
$\geq 5\%$	5 mg: 71 10 (13,9) 4,125	74 5 (6,8) 2,917	2,26[0,73;6,99]
	10 mg: 89 10 (11,2) 3,348		1,75 [0,57;5,36]
Interaktionstest	5 mg		$I^2=0\%$; $p=0,33$
	10 mg		$I^2=66,8\%$; $p=0,08$

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

* Prozentangabe adjustiert bezüglich dem Ausgangswert

SE:Standardfehler

Bezüglich des Ausmaßes der Gewichtsveränderung ergab sich ein Hinweis auf eine signifikante Interaktion durch die Abstammung der Patienten (Interaktionstest: $I^2=61,3\%$; $p=0,11$ zwischen Sulfonylharnstoff+Dapagliflozin 10 mg versus Sulfonylharnstoff+Placebo). Der Effekt der Gewichtsveränderung war jedoch in allen Gruppen gleichgerichtet und zeigte einen Vorteil von Dapagliflozin gegenüber Placebo. Zudem wurde für den Vergleich Sulfonylharnstoff+Dapagliflozin 5 mg versus Sulfonylharnstoff+Placebo keine Interaktion beobachtet. Für die Operationalisierung „Gewichtsverlust ($\geq 5\%$)“ ergab sich ein Hinweis auf eine Effektmodifikation durch die Abstammung der Patienten ($I^2=66,8\%$; $p=0,08$), jedoch nur nach 48 Wochen (nicht nach 24 Wochen) und zwischen Sulfonylharnstoff+Dapagliflozin 10 mg versus Sulfonylharnstoff+Placebo (nicht für die 5 mg Gruppe). Insgesamt weisen diese Ergebnisse damit nicht auf eine fazitrelevante Modifikation des Behandlungseffektes hin.

4.3.1.3.2.10 Therapiezufriedenheit stratifiziert nach dem Geschlecht – RCT

Tabelle 4-67: Ergebnisse für Therapiezufriedenheit stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen** Männlich									
	5 mg: 51	28,0 (5,81)	30,9 (5,41)	2,9 (5,53)	41	29,0 (5,28)	30,4 (4,57)	1,4 (6,26)	1,50 [-0,94;3,94]
	10 mg: 51	28,5 (6,94)	31,1 (4,82)	2,6 (5,95)					1,20 [-1,32;3,72]
24 Wochen ** Weiblich									
	5 mg: 51	26,1 (6,53)	30,0 (6,23)	3,9 (7,02)	43	26,8 (9,02)	29,8 (6,81)	3,0 (9,25)	0,90 [-2,41;4,27]
	10 mg: 69	26,7 (6,18)	31,1 (4,99)	4,4 (6,05)					1,40 [-1,71;4,51]
Interaktions- test	5 mg versus Placebo								I ² =0%; p=0,78
	10 mg versus Placebo								I ² =0%; p=0,92
48 Wochen** Männlich									
	5 mg: 42	27,6 (5,89)	32,2 (4,40)	4,6 (6,03)	21	28,3 (5,66)	31,0 (3,57)	2,7 (5,54)	1,90 [-1,09;4,89]

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

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	
D1690C00005	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				1,00 [-2,17;4,17]
	10 mg: 42	29,0 (6,90)	32,7 (3,64)	3,7 (6,94)					
48 Wochen** Weiblich									
	5 mg: 31	25,7 (6,32)	30,7 (6,03)	5,0 (8,10)	20	28,5 (8,95)	30,7 (6,96)	2,2 (8,39)	2,80 [-1,85;7,45]
	10 mg: 50	26,8 (6,51)	31,9 (4,57)	5,1 (7,12)					2,10 [-2,41;6,61]
Interaktions- test	5 mg versus Placebo:								I ² =0%; p=0,75
	10 mg versus Placebo								I ² =0%; p=0,70

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

SD Standardabweichung, SE Standardfehler

LOCF wird nicht betrachtet

** beobachtete Werte

Es gab für die Subgruppenanalyse bezüglich der Therapiezufriedenheit keine signifikanten Interaktionen ($I^2=0\%$; $p=0,78$ / $I^2=0\%$; $p=0,92$ nach 24 Wochen und $I^2=0\%$; $p=0,75$ / $I^2=0\%$; $p=0,70$ nach 48 Wochen) bezüglich des Geschlechts.

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

4.3.1.3.2.11 Therapiezufriedenheit stratifiziert nach dem Alter – RCT

Tabelle 4-68: Ergebnisse für Therapiezufriedenheit stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen** Alter <65 Jahre									
	5 mg: 76	27,6 (6,52)	30,8 (5,59)	3,2 (6,27)	58	28,3 (7,61)	30,1 (6,08)	1,8 (8,29)	1,40 [-1,16;3,96]
	10 mg : 97	27,1 (6,63)	31,0 (4,90)	4,0 (6,19)					2,20 [-0,26;4,66]
24 Wochen** Alter ≥65 und <75 Jahre									
	5 mg: 26	25,3 (4,97)	29,4 (6,47)	4,1 (6,50)	26	26,9 (7,21)	30,1 (5,25)	3,2 (7,10)	0,90 [-2,80;4,60]
	10 mg : 23	29,2 (6,04)	31,4 (5,01)	2,2 (5,28)					-1,00 [-4,48;2,48]
Interaktionstest	5 mg versus Placebo								I ² =0%; p=0,83
	10 mg versus Placebo								I ² =53,8%; p=0,14
48 Wochen** Alter <65 Jahre									
	5 mg: 54	27,3 (6,57)	32,3 (4,31)	5,0 (7,29)	25	29,6 (6,71)	31,2 (5,54)	1,6 (7,75)	3,40 [-0,21;7,01]

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

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	
D1690C00005	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				3,20 [-0,25;6,65]
	10 mg : 76	27,3 (6,84)	32,1 (4,43)	4,8 (7,28)					
48 Wochen** Alter ≥65 und <75 Jahre									
	5 mg: 19	25,5 (4,41)	29,6 (6,84)	4,2 (5,94)	16	26,6 (8,14)	30,3 (5,36)	3,7 (5,62)	0,50 [-3,34;4,34]
	10 mg : 16	30,1 (5,94)	33,2 (2,56)	3,1 (5,74)					-0,60 [-4,54;3,34]
Interaktionstest	5 mg versus Placebo:								I ² =14,2%; p=0,28
	10 mg versus Placebo								I ² =50,6%; p=0,15

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

SD Standardabweichung, SE Standardfehler

LOCF wird nicht berichtet

** beobachtete Werte

Der Interaktionstest war in der Sulfonylharnstoff+Dapagliflozin 10 mg versus Sulfonylharnstoff+Placebo-Gruppe nach dem Alter signifikant ($I^2=53,8\%$; $p=0,14$ nach 24 Wochen / $I^2=50,6\%$; $p=0,15$ nach 48 Wochen), kein Hinweis auf Interaktion gab es aber für den Vergleich Sulfonylharnstoff+Dapagliflozin 5 mg versus Sulfonylharnstoff+Placebo. Angesichts der geringen Anzahl in die Auswertung einbezogener Patienten ist dieses Ergebnis daher nicht als fazitrelevante Modifikation des Behandlungseffektes von Sulfonylharnstoff+Dapagliflozin zu beurteilen.

4.3.1.3.2.12 Therapiezufriedenheit stratifiziert nach asiatischer und nicht nicht-asiatischer Abstammung – RCT

Tabelle 4-69: Ergebnisse für Therapiezufriedenheit stratifiziert nach asiatischer und nicht-asiatischer Abstammung - RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen** Asiaten									
	5 mg: 40	26,3 (5,91)	30,3 (5,76)	4,0 (5,07)	28	25,8 (7,51)	28,3 (6,67)	2,5 (6,79)	1,50 [-1,47;4,47]
	10 mg: 38	25,6 (6,14)	29,6 (5,72)	4,0 (6,63)					1,50 [-1,78;4,78]
24 Wochen ** Nicht-Asiaten									
	5 mg: 62	27,5 (6,42)	30,6 (5,91)	3,0 (7,00)	56	28,9 (7,30)	31,0 (5,14)	2,1 (8,49)	0,90 [-1,92;3,72]
	10 mg: 82	28,3 (6,59)	31,8 (4,33)	3,5 (5,80)					1,40 [-1,15;3,95]
Interaktions- test	5 mg versus Placebo								I ² =0%; p=0,77
	10 mg versus Placebo								I ² =0%; p=0,96
48 Wochen** Asiaten									
	5 mg: 27	26,3 (5,86)	29,1 (6,36)	2,8 (7,78)	9	27,2 (8,56)	29,6 (6,86)	2,3 (4,39)	0,50 [-3,00;4,60]

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

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert (SE) [95%-KI]
	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	N	Ausgangswert Mittelwert (SD)	Endwert Mittelwert (SD)	Veränderung Mittelwert (SD)	
D1690C00005	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				3,20 [-0,56;6,96]
	10 mg: 26	25,7 (6,05)	31,2 (3,84)	5,5 (6,34)					
48 Wochen** Nicht-Asiaten									
	5 mg: 46	27,1 (6,30)	33,1 (3,67)	5,9 (6,17)	32	28,7 (7,09)	31,2 (5,02)	2,5 (7,62)	3,40 [0,21;6,59]
	10 mg: 66	28,6 (6,88)	32,7 (4,25)	4,1 (7,29)					1,60 [-1,83;5,03]
Interaktions- test	5 mg versus Placebo:								I ² =11,3%; p=0,29
	10 mg versus Placebo								I ² =0%; p=0,54

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

SD Standardabweichung, SE Standardfehler

LOCF wird nicht berichtet

** beobachtete Werte

Bezüglich der Therapiezufriedenheit ergab sich kein statistisch signifikanter Interaktionstest ($I^2=0\%$; $p=0,77$ und $I^2=0\%$; $p=0,96$ nach 24 Wochen).

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

4.3.1.3.2.13 Blutdrucksenkung stratifiziert nach dem Geschlecht – RCT

Tabelle 4-70: Ergebnisse für Blutdrucksenkung stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel

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)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24* Wochen Männlich									
systolisch	5 mg: 57	131,4 (14,18)	128,5 (14,07)	-4,2 (1,529)	58	135,4 (12,28)	133,7 (12,86)	-1,0 (1,510)	-3,2 (2,154) [-7,4;1,1]
	10 mg: 56	134,2 (12,84)	129,4 (13,89)	-4,7 (1,535)					-3,7 (2,153) [-8,0;0,5]
diastolisch	5 mg: 57	79,0 (8,72)	77,3 (8,49)	-2,4 (0,947)	58	80,9 (7,33)	79,1 (7,67)	-1,6 (0,936)	-0,8 (1,333) [-3,4;1,8]
	10 mg 56	80,5 (8,48)	77,3 (7,83)	-3,2 (0,952)					-1,6 (1,335) [-4,2;1,1]
24* Wochen Weiblich									
systolisch	5 mg: 56	128,5 (15,45)	124,7 (12,99)	-4,6 (1,432)	55	132,6 (16,37)	130,3 (14,64)	-1,1 (1,448)	-3,4 (2,041) [-7,5;0,6]
	10 mg: 73	130,2 (13,85)	124,3 (12,17)	-5,9 (1,253)					-4,7 (1,914) [-8,5;-0,9]
diastolisch	5 mg: 56	76,7 (7,58)	76,1 (6,68)	-1,2 (0,924)	55	80,2 (7,83)	78,5 (7,20)	-0,4 (0,939)	-0,8 (1,324) [-3,4;1,8]

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)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
	10 mg: 73	78,1 (8,34)	75,2 (8,35)	-2,8 (0,808)					-2,4 (1,237) [-4,8;0,1]
Interaktions- test systolisch	5 mg versus Placebo								I ² =0%; p=0,92
	10 mg versus Placebo								I ² =0%; p=0,70
diastolisch	5 mg versus Placebo								I ² =0%; p=1,00
	10 mg versus Placebo								I ² =0%; p=0,66
24 Wochen** Männlich									
systolisch	5 mg: 51		128,55 (13,779)	-4,28 (1,6809)	41		133,58 (13,107)	0,21 (1,8323)	-4,50 (2,4892) [-9,41;0,42]
	10 mg 52		129,41 (14,353)	-4,32 (1,6676)					
diastolisch	5 mg: 51		77,22 (8,053)	-2,54 (0,9527)	41		78,95 (7,345)	-0,68 (1,0465)	-1,86 (1,4151) [-4,66;0,93]
	10 mg 52		76,97 (7,914)	-2,97 (0,9457)					
24 Wochen** Weiblich									
systolisch	5 mg: 52		123,81 (12,123)	-5,09 (1,4574)	43		131,43 (14,871)	-0,91 (1,5673)	-4,18 (2,1493) [-8,42;0,06]

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)	
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
	10 mg 69	124,24 (11,942)	-6,01 (1,2576)				-5,10 (2,0099) [-9,07;-1,13]
diastolisch	5 mg: 52	75,92 (6,852)	-1,40 (0,9697)	43	78,36 (7,237)	-0,71 (1,0400)	-0,69 (1,4338) [-3,52;2,14]
	10 mg 69	75,06 (8,459)	-3,12 (0,8327)				-2,41 (1,3329) [-5,04;0,22]
48 Wochen** Männlich							
systolisch	5 mg: 42	129,25 (14,023)	-3,99 (1,6956)	21	135,49 (7,985)	2,07 (2,3070)	-6,07 (2,8664) [-11,74;-0,39]
	10 mg 42	127,29 (12,349)	-5,94 (1,6904)				-8,02 (2,8589) [-13,68;-2,35]
diastolisch	5 mg: 42	77,05 (8,595)	-2,54 (1,1500)	21	80,84 (8,665)	1,65 (1,5654)	-4,19 (1,9407) [-8,03;-0,35]
	10 mg 42	76,68 (7,599)	-2,93 (1,1469)				-4,58 (1,9405) [-8,43;-0,74]

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)	
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
48 Wochen** Weiblich							
systolisch	5 mg: 31	126,19 (15,911)	-2,76 (2,1731)	20	132,02 (15,621)	-1,19 (2,6922)	-1,57 (3,4746) [-8,46;5,32]
	10 mg 52	128,01 (14,134)	-2,41 (1,6760)				-1,22 (3,1749) [-7,51;5,07]
diastolisch	5 mg: 31	75,61 (7,590)	-1,57 (1,2270)	20	77,80 (6,461)	0,15 (1,4719)	-1,72 (1,9182) [-5,52;2,08]
	10 mg 52	77,17 (8,323)	-1,43 (0,9529)				-1,58 (1,7537) [-5,05;1,89]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

** Repeated Measurements Analyse

SD: Standardabweichung SE: Standardfehler

Der Interaktionstest nach 24 Wochen bezüglich des Geschlechts zeigte keine Interaktion weder in der systolischen ($I^2=0\%$; $p=0,92$, $I^2=0\%$; $p=0,70$) noch in der diastolischen Blutdrucksenkung ($I^2=0\%$; $p=1,00$ $I^2=0\%$; $p=0,66$).

4.3.1.3.2.14 Blutdrucksenkung stratifiziert nach dem Alter – RCT

Tabelle 4-71: Ergebnisse für Blutdrucksenkung stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert mmHg (SE) [95%-KI]
	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert ^{oo} mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert ^{oo} mmHg (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24 Wochen* Alter <65 Jahre									
systolisch	5 mg: 85	129,8 (14,48)	125,6 (13,37)	-4,9 (1,176)	82	133,7 (15,15)	131,8 (13,82)	-0,6 (1,200)	-4,3 (1,683) [-7,6;-0,9]
	10 mg 106	130,7 (13,54)	125,3 (11,94)	-5,7 (1,052)					-5,0 (1,596) [-8,2;-1,9]
diastolisch	5 mg: 85	78,9 (7,63)	76,9 (7,35)	-2,4 (0,742)	82	81,0 (7,71)	79,8 (7,07)	-0,5 (0,758)	-1,9 (1,062) [-4,0;0,2]
	10 mg 106	79,4 (8,52)	76,2 (7,63)	-3,3 (0,664)					-2,8 (1,008) [-4,7;-0,8]
24 Wochen* Alter ≥65 und <75 Jahre									
systolisch	5 mg: 28	130,3 (16,13)	129,7 (14,15)	-2,5 (2,282)	31	135,1 (12,45)	132,8 (13,95)	-2,2 (2,140)	-0,3 (3,130) [-6,5;5,9]
	10 mg 23	137,5 (12,17)	132,3 (16,79)	-4,2 (2,493)					-1,9 (3,285) [-8,5;4,6]

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 ^{oo} mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert ^{oo} mmHg (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
diastolisch	5 mg: 28	74,5 (9,17)	76,1 (8,58)	-0,1 (1,426)	31	79,5 (7,11)	76,1 (7,75)	-2,3 (1,343)	2,3 (1,978) [-1,7;6,2]
	10 mg 23	78,0 (8,25)	75,5 (10,46)	-2,2 (1,550)					0,1 (2,048) [-4,0;4,1]
Interaktions- test systolisch	5 mg versus Placebo:								I ² =3,7%; p=0,31
	10 mg versus Placebo								I ² =0%; p=0,40
diastolisch	5 mg versus Placebo:								I ² =70,5%; p=0,07
	10 mg versus Placebo								I ² =37,9%; p=0,20
24 Wochen** Alter <65 Jahre									
systolisch	5 mg: 77		125,11 (12,547)	-5,31 (1,2312)	58		132,22 (14,156)	-0,13 (1,3766)	-5,17 (1,8499) [-8,81;-1,53]
	10 mg 98		125,09 (11,937)	-5,70 (1,0894)		-5,57 (1,7570) [-9,03;-2,11]			
diastolisch	5 mg: 77		76,66 (7,036)	-2,63 (0,7280)	58		80,03 (6,878)	-0,14 (0,8192)	-2,49 (1,0970) [-4,65;-0,33]
	10 mg 98		75,97 (7,704)	-3,50 (0,6447)		-3,37 (1,0433) [-5,42;-1,31]			

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 ^{oo} mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
24 Wochen** Alter ≥65 und <75 Jahre							
systolisch	5 mg: 26	129,24 (14,519)	-2,34 (2,4539)	26	133,05 (13,888)	-1,48 (2,4068)	-0,86 (3,4485) [-7,72;6,01]
	10 mg 23	132,30 (16,793)	-3,86 (2,6091)				-2,38 (3,5400) [-9,43;4,67]
diastolisch	5 mg: 26	76,26 (8,751)	0,40 (1,6199)	26	75,58 (7,246)	-1,76 (1,5945)	2,16 (2,3030) [-2,43;6,74]
	10 mg 23	75,51 (10,460)	-1,93 (1,7076)				-0,17 (2,3250) [-4,81;4,46]
48 Wochen** Alter <65 Jahre							
systolisch	5 mg: 54	128,72 (15,668)	-2,71 (1,5568)	25	132,29 (11,469)	-1,25 (2,2124)	-1,46 (2,7071) [-6,80;3,88]
	10 mg 78	126,64 (13,696)	-4,23 (1,3075)				-2,98 (2,5721) [-8,05;2,10]
diastolisch	5 mg: 54	76,57 (8,629)	-2,90 (0,9674)	25	80,83 (7,693)	0,83 (1,3605)	-3,73 (1,6694) [-7,02;-0,44]
	10 mg 78	77,28 (7,633)	-2,59 (0,8125)				-3,43 (1,5850) [-6,55;-0,30]

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 ^{oo} mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert ^{oo} mmHg (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
48 Wochen** Alter ≥65 und <75 Jahre									
systolisch	5 mg: 19	125,75 (12,222)	-6,18 (2,6799)	16	136,15 (13,498)	1,90 (2,8720)	-8,08 (3,9489) [-16,01;-0,14]		
	10 mg 16	132,79 (10,027)	-2,36 (2,8746)				-4,26 (4,0487) [-12,39;3,88]		
diastolisch	5 mg: 19	76,07 (6,847)	0,84 (1,7228)	16	77,06 (7,441)	1,42 (1,7887)	-0,58 (2,5029) [-5,59;4,43]		
	10 mg 16	75,38 (9,563)	-0,13 (1,8288)				-1,55 (2,5571) [-6,67;3,57]		

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

** Repeated Measurements Analyse

SD: Standardabweichung SE: Standardfehler

^{oo}adjustiert bezüglich Ausgangswert

Der Interaktionstest bezüglich des Alters zeigte nach 24 Wochen einen Hinweis auf Interaktion für die diastolischen Blutdrucksenkung ($I^2=70,5\%$; $p=0,07$, $I^2=37,9\%$; $p=0,20$). Da die Anzahl der in dieser Analyse berücksichtigten >65 jährigen Patienten jedoch gering war und sich zudem keine Hinweise für eine Interaktion bzgl. systolischer Blutdrucksenkung fanden ($I^2=3,7\%$; $p=0,31$, $I^2=0\%$; $p=0,40$), ist nicht von einer fazitrelevante Modifikation des Behandlungseffektes auszugehen.

4.3.1.3.2.15 Blutdrucksenkung stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT

Tabelle 4-72: Ergebnisse für Blutdrucksenkung stratifiziert nach asiatischer und nicht-asiatischer Abstammung aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied Mittelwert mmHg (SE) [95%-KI]
	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert ^{oo} mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert ^{oo} mmHg (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
24* Wochen Asiaten									
systolisch	5 mg: 40	123,9 (15,35)	119,7 (12,21)	-4,7 (1,442)	28	126,4 (12,94)	124,7 (12,80)	-1,1 (1,724)	-3,7 (2,249) [-8,1;0,8]
	10 mg: 38	123,5 (12,82)	120,3 (11,49)	-3,9 (1,480)					-2,8 (2,275) [-7,3;1,7]
diastolisch	5 mg: 40	75,9 (7,96)	75,3 (6,15)	-1,5 (1,027)	28	79,3 (9,02)	77,7 (8,10)	-0,9 (1,223)	-0,7 (1,606) [-3,9;2,5]
	10 mg: 38	76,6 (8,11)	76,7 (8,13)	-0,5 (1,049)					0,4 (1,617) [-2,8;3,6]
24* Wochen Nicht-Asiaten									
systolisch	5 mg: 63	133,6 (14,17)	130,2 (12,07)	-4,2 (1,478)	56	136,7 (12,54)	136,4 (12,99)	0,6 (1,568)	-4,8 (2,158) [-9,0;-0,5]
	10 mg: 83	135,4 (11,65)	129,3 (13,07)	-6,0 (1,286)					-6,6 (2,027) [-10,6;-2,6]
diastolisch	5 mg: 63	79,0 (8,19)	77,4 (8,14)	-1,9 (0,891)	56	79,9 (5,72)	79,1 (6,82)	-0,6 (0,945)	-1,3 (1,299) [-3,9;1,3]

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 ^{oo} mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	Endwert Mittelwert mmHg (SD)	Adjustierte Veränderung Mittelwert ^{oo} mmHg (SE)	
	Sulfonylharnstoff+Dapagliflozin				Sulfonylharnstoff+Placebo				
D1690C00005	10 mg: 83	80,1 (8,20)	75,5 (8,33)	-4,3 (0,777)					-3,7 (1,222) [-6,1;-1,3]
Interaktions- test systolisch	5 mg versus Placebo:								I ² =0%; p=0,70
	10 mg versus Placebo								I ² =35,8%; p=0,21
diastolisch	5 mg versus Placebo:								I ² =0%; p=0,73
	10 mg versus Placebo								I ² =75,7%; p=0,04
24 Wochen** Asiaten									
systolisch	5 mg: 40		119,71 (12,213)	-4,80 (1,3625)	28		124,68 (12,796)	-1,68 (1,5966)	-3,11 (2,1018) [-7,28;1,05]
	10 mg 38		120,32 (11,493)	-3,67 (1,3969)					-1,99 (2,1250) [-6,20;2,22]
diastolisch	5 mg: 40		75,31 (6,147)	-1,44 (1,0104)	28		77,74 (8,102)	-0,89 (1,1728)	-0,55 (1,5597) [-3,64;2,54]
	10 mg 38		76,68 (8,129)	-0,44 (1,0302)					0,45 (1,5662) [-2,65;3,56]

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 ^{oo} mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
24 Wochen** Nicht-Asiaten							
systolisch	5 mg: 63	130,25 (12,069)	4,33 (1,4949)	56	136,38 (12,989)	0,40 (1,5605)	-4,73 (2,1661) [-9,00 ; -0,46]
	10 mg 83	129,28 (13,073)	-6,32 (1,3018)				-6,72 (2,0318) [-10,73; -2,72]
diastolisch	5 mg: 63	77,35 (8,136)	-2,11 (0,8923)	56	79,11 (6,819)	-0,73 (0,9309)	-1,39 (1,2902) [-3,93; 1,16]
	10 mg 83	75,51 (8,328)	-4,44 (0,7779))				-3,71 (1,2130) [-6,10; -1,32]
48 Wochen** Asiaten							
systolisch	5 mg: 29	119,69 (11,662)	-6,22 (1,6243)	13	119,15 (12,503)	-5,04 (2,4068)	-1,18 (2,9039) [-6,96; 4,60]
	10 mg 32	119,33 (10,932)	-4,18 (1,5607)				0,86 (2,8667) [-4,85; 6,57]
diastolisch	5 mg: 29	74,22 (6,844)	-2,86 (1,1024)	13	75,20 (8,875)	-2,69 (1,6008)	-0,16 (1,9472) [-4,03; 3,71]
	10 mg 32	75,65 (7,021)	-1,38 (1,0577)				1,31 (1,9212) [-2,51; 5,13]

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 ^{oo} mmHg (SE)	N	Ausgangswert Mittelwert mmHg (SD)	
	Sulfonylharnstoff+Dapagliflozin			Sulfonylharnstoff+Placebo			
48 Wochen** Nicht-Asiaten							
systolisch	5 mg: 46	130,51 (14,795)	-3,68 (1,7367)	32	135,09 (11,025)	-1,47 (2,0327)	-2,21 (2,6823) [-7,51;3,09]
	10 mg 67	130,51 (12,983)	-4,89 (1,4453 ^o)				-3,43 (2,4952) [-8,35;1,50]
diastolisch	5 mg: 46	77,28 (8,104)	-1,86 (1,0457)	32	79,94 (7,459)	1,29 (1,2152)	-3,14 (1,6014) [-6,30;0,02]
	10 mg 67	76,76 (8,189)	-3,24 (0,8752)				-4,53 (1,4989) [-7,48;-1,57]

Daten sind exklusive der Patienten, die eine Notfallmedikation bekommen haben.

*LOCF

** Repeated Measurements Analyse

SD: Standardabweichung SE: Standardfehler

^{oo} adjustiert bezüglich Ausgangswert

Der Interaktionstest bezüglich der Abstammung zeigte keine Interaktion in der systolischen Blutdrucksenkung ($I^2=0\%$; $p=0,70$, $I^2=35,8\%$; $p=0,21$). In der diastolischen Blutdrucksenkung gab es einen signifikanten Interaktionstest für die Patientengruppe, die 10 mg Dapagliflozin erhielten ($I^2=75,7\%$; $p=0,04$). Da die Interaktion wiederum nur für den diastolischen Blutdruck beobachtet werden konnte, ist dies insgesamt nicht als fazitrelevant zu beurteilen.

4.3.1.3.2.16 Unerwünschte Ereignisse stratifiziert nach dem Geschlecht – RCT

Tabelle 4-73: Ergebnisse für Unerwünschte Ereignisse stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate unerwünschter Ereignisse, n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
D1690C00005				
24 Wochen				
Männlich				
Sulfonylharnstoff+Dapagliflozin 5 mg	58	24 (41,4)	4 (6,9)	2 (3,4)
Sulfonylharnstoff+Dapagliflozin 10 mg	56	26 (46,4)	4 (7,1)	0
Sulfonylharnstoff +Placebo	58	21 (36,2)	4 (6,9)	1 (1,7)
5 mg vs.Placebo OR [95%-KI]		1,24 [0,59;2,63]	1,00[0,24;4,21]	2,04 [0,18;23,019]
10 mg vs. Placebo OR [95%-KI]		1,53 [0,72;3,23]	1,04 [0,25;4,37]	0,34 [0,01;8,50]
24 Wochen				
Weiblich				
Sulfonylharnstoff +Dapagliflozin 5 mg	58	27 (46,6)	4 (6,9)	1 (1,7)
Sulfonylharnstoff +Dapagliflozin 10 mg	73	37 (50,7)	3 (4,1)	1 (1,4)
Sulfonylharnstoff +Placebo	56	32 (57,1)	1 (1,8)	1 (1,8)
5 mg vs.Placebo OR [95%-KI]		0,65 [0,31;1,37]	4,07[0,44;37,63]	0,96 [0,06;15,81]
10 mg vs. Placebo OR [95%-KI]		0,77 [0,38;1,55]	2,36 [0,24;23,29]	0,76 [0,05;12,49]
Interaktionstest				
5 mg vs.Placebo		I ² =30,6%; p=0,23	I ² =7,6%; p=0,30	I ² =0%;p=0,69
10 mg vs. Placebo		I ² =41,3%; p=0,19	I ² =0%; p=0,55	I ² =0%;p=0,71
48 Wochen				
Männlich				
Sulfonylharnstoff +Dapagliflozin 5 mg	58	32 (55,2)	7 (12,1)	2 (3,4)
Sulfonylharnstoff +Dapagliflozin 10 mg	56	27 (48,2)	5 (8,9)	0
Sulfonylharnstoff +Placebo	58	27 (46,6)	4 (6,9)	2 (3,4)
5 mg vs.Placebo OR [95%-KI]		1,41 [0,68;2,93]	1,85 [0,51;6,71]	1,00 [0,14;7,35]
10 mg vs. Placebo OR [95%-KI]		1,07 [0,51;2,23]	1,32 [0,34;5,21]	0,20 [0,01;4,26]
48 Wochen				
Weiblich				
Sulfonylharnstoff +Dapagliflozin 5 mg	58	35 (60,3)	6 (10,3)	1 (1,7)
Sulfonylharnstoff +Dapagliflozin 10 mg	73	47 (64,4)	6 (8,2)	1 (1,4)
Sulfonylharnstoff +Placebo	56	36 (64,3)	4 (7,1)	1 (1,8)
5 mg vs.Placebo OR [95%-KI]		1,18 [0,55;2,53]	1,56 [0,42;5,84]	0,96 [0,06;15,81]
10 mg vs. Placebo OR [95%-KI]		1,19 [0,58;2,42]	1,16 [0,31;4,34]	0,76 [0,05;12,49]
Interaktionstest				
5 mg vs.Placebo		I ² =0%; p=0,74	I ² = 0%; p=0,85	I ² = 0%; p=0,98
10 mg vs. Placebo		I ² =0%; p=0,84	I ² = 0%; p=0,89	I ² =0%; p=0,53

Bei der Gesamtrate unerwünschter Ereignisse zeigte sich eine positive Interaktion im Vergleich Sulfonylharnstoff+Dapagliflozin 10 mg gegenüber Sulfonylharnstoff+Placebo ($I^2=41,3\%$; $p=0,19$) nach Woche 24. Statistisch signifikante Effekte waren aber in keiner der Subgruppen zu beobachten, und zudem zeigte sich nach Woche 48 keine Interaktion mehr ($I^2=0\%$; $p=0,84$). Es ist daher nicht von einer fazitrelevanten Modifikation des Behandlungseffektes auszugehen.

Tabelle 4-74: Ergebnisse für Anzahl der Patienten mit Anzeichen einer Genital- oder Harnwegsinfektion stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Anzahl der Patienten mit Anzeichen auf Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen auf Harnwegsinfektion, n (%)
D1690C00005			
24 Wochen			
Männlich			
Sulfonylharnstoff+Dapagliflozin 5 mg	58	2 (3,4)	1 (1,7)
Sulfonylharnstoff +Dapagliflozin 10 mg	56	3 (5,4)	1 (1,8)
Sulfonylharnstoff +Placebo	58	0	0
5 mg vs.Placebo OR [95%-KI]		5,18 [0,24;110,22]	3,05 [0,12;76,48]
10 mg vs. Placebo OR [95%-KI]		7,65 [0,39;151,65]	3,16 [0,13;79,27]
24 Wochen			
Weiblich			
Sulfonylharnstoff+Dapagliflozin 5 mg	58	5 (8,6)	4 (6,9)
Sulfonylharnstoff+Dapagliflozin 10 mg	73	4 (5,5)	4 (5,5)
Sulfonylharnstoff +Placebo	56	1 (1,8)	6 (10,7)
5 mg vs.Placebo OR [95%-KI]		5,19 [0,59;45,90]	0,62 [0,16;2,32]
10 mg vs. Placebo OR [95%-KI]		3,19 [0,35;29,35]	0,48 [0,13;1,80]
Interaktionstest			
5 mg vs.Placebo		I ² =0%; p=1,00	I ² =0%; p=0,37
10 mg vs. Placebo		I ² =0%; p=0,64	I ² =10,7%; p=0,29
48 Wochen			
Männlich			
Sulfonylharnstoff +Dapagliflozin 5 mg	58	2 (3,4)	1 (1,7)
Sulfonylharnstoff +Dapagliflozin 10 mg	56	3 (5,4)	3 (5,4)
Sulfonylharnstoff +Placebo	58	0	1 (1,7)
5 mg vs.Placebo OR [95%-KI]		5,18 [0,24;110,22] ⁺	1,00 [0,06;16,38]
10 mg vs. Placebo OR [95%-KI]		7,65 [0,39;151,65] ⁺	3,23 [0,33;31,98]
48 Wochen			
Weiblich			
Sulfonylharnstoff+Dapagliflozin 5 mg	58	5 (8,6)	5 (8,6)
Sulfonylharnstoff+Dapagliflozin 10 mg	73	5 (6,8)	5 (6,8)
Sulfonylharnstoff +Placebo	56	2 (3,6)	7 (12,5)
5 mg vs.Placebo OR [95%-KI]		2,55 [0,47;13,71]	0,66 [0,20;2,22]
10 mg vs. Placebo OR [95%-KI]		1,99 [0,37;10,63]	0,51 [0,15;1,72]
Interaktionstest			
5 mg vs.Placebo		I ² =0%; p=0,69	I ² =0%; p=0,79
10 mg vs. Placebo		I ² =0%; p=0,44	I ² =48,1%; p=0,17

Bei den Ereignissen, die auf eine Harnwegsinfektion hinweisen, gab es hinsichtlich des Geschlechts eine signifikante Interaktion nach Woche 48 (I²=48,1%; p=0,17). Die Effektschätzer für die einzelnen Subgruppen waren aber alle statistisch nicht signifikant verändert. Die Anzahl der Ereignisse war gering, und auch ansonsten gab es bezüglich der

Ereignisse, die auf Harnwegs- oder Genitalinfektionen hinweisen, keine weiteren signifikanten Interaktionen. Es ist daher nicht von einer fazitrelevanten Modifikation des Behandlungseffektes auszugehen.

4.3.1.3.2.17 Unerwünschte Ereignisse stratifiziert nach dem Alter – RCT

Tabelle 4-75: Ergebnisse für Unerwünschte Ereignisse stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate unerwünschter Ereignisse* n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
D1690C00005				
24 Wochen				
Alter <65 Jahre				
Sulfonylharnstoff +Dapagliflozin 5 mg	87	38 (43,7)	7 (8,0)	2 (2,3)
Sulfonylharnstoff +Dapagliflozin 10 mg	106	51 (48,1)	6 (5,7)	1 (0,9)
Sulfonylharnstoff +Placebo	82	40 (48,8)	5 (6,1)	2 (2,4)
5 mg vs.Placebo OR [95%-KI]		0,18 [0,44;1,49]	1,35 [0,41;4,43]	0,94 [0,13;6,84]
10 mg vs. Placebo OR [95%-KI]		0,97 [0,55;1,73]	0,92 [0,27;3,14]	0,38 [0,03;4,28]
24 Wochen				
Alter ≥65 und <75 Jahre				
Sulfonylharnstoff +Dapagliflozin 5 mg	29	13 (44,8)	1 (3,4)	1 (3,4)
Sulfonylharnstoff +Dapagliflozin 10 mg	23	12 (52,2)	1 (4,3)	0
Sulfonylharnstoff +Placebo	32	13 (40,6)	0	0
5 mg vs.Placebo OR [95%-KI]		1,19 [0,43;3,28]	3,42 [0,13;87,34] ⁺	3,42 [0,13;87,34]
10 mg vs. Placebo OR [95%-KI]		1,59 [0,54;4,70]	4,33 [0,17;111,25] ⁺	nicht berechenbar
Interaktionstest				
5 mg vs.Placebo		I ² =0%;p=0,53	I ² =0%;p=0,60	I ² =0%;p=0,51
10 mg vs. Placebo		I ² =0%; p=0,43	I ² =0%; p=0,38	nicht berechenbar
48 Wochen				
Alter <65 Jahre				
Sulfonylharnstoff +Dapagliflozin 5 mg	87	49 (56,3)	11 (12,6)	2 (2,3)
Sulfonylharnstoff +Dapagliflozin 10 mg	106	59 (55,7)	8 (7,5)	1 (0,9)
Sulfonylharnstoff +Placebo	82	47 (57,3)	7 (8,5)	2 (2,4)
5 mg vs.Placebo OR [95%-KI]		0,96 [0,52;1,77]	1,55 [0,57;4,21]	0,94 [0,13;6,84]
10 mg vs. Placebo OR [95%-KI]		0,93 [0,52;1,67]	0,87 [0,30;2,52]	0,38 [0,03;4,28]
48 Wochen				
Alter ≥65 und <75 Jahre				
Sulfonylharnstoff +Dapagliflozin 5 mg	29	18 (62,1)	2 (6,9)	1 (3,4)
Sulfonylharnstoff +Dapagliflozin 10 mg	23	15 (65,2)	3 (13,0)	0
Sulfonylharnstoff +Placebo	32	16 (50,0)	1 (3,1)	1 (3,1)
5 mg vs.Placebo OR [95%-KI]		1,64 [0,59;4,54]	2,30 [0,20;26,75]	1,11 [0,07;18,55]
10 mg vs. Placebo OR [95%-KI]		1,88 [0,62;5,65]	4,65 [0,45;47,89]	0,45 [0,02;11,46]
Interaktionstest				
5 mg vs.Placebo		I ² =0%; p=0,38	I ² =0%; p=0,77	I ² =0%; p=0,93
10 mg vs. Placebo		I ² =16,5%; p=0,27	I ² =38,9%;p=0,20	I ² =0%; p=0,94

Hinsichtlich des Alters gab es bezüglich der unerwünschten Ereignisse keine signifikanten Interaktionstests.

Tabelle 4-76: Ergebnisse für Ergebnisse für Anzahl der Patienten mit Anzeichen einer Genital- oder Harnwegsinfektion stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Anzahl der Patienten mit Anzeichen auf Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen auf Harnwegsinfektion, n (%)
D1690C00005			
24 Wochen			
Alter <65 Jahre			
Sulfonylharnstoff +Dapagliflozin 5 mg	87	5 (5,7)	4 (4,6)
Sulfonylharnstoff +Dapagliflozin 10 mg	106	5 (4,7)	4 (3,8)
Sulfonylharnstoff +Placebo	82	1 (1,2)	5 (6,1)
5 mg vs.Placebo OR [95%-KI]		4,94 [0,56;43,21]	0,74 [0,19;2,81]
10 mg vs. Placebo OR [95%-KI]		4,01 [0,46;35,01]	0,60 [0,16;2,32]
24 Wochen			
Alter ≥65 und <75 Jahre			
Sulfonylharnstoff +Dapagliflozin 5 mg	29	2 (6,9)	1 (3,4)
Sulfonylharnstoff +Dapagliflozin 10 mg	23	2 (8,7)	1 (4,3)
Sulfonylharnstoff +Placebo	32	0	1 (3,1)
5 mg vs.Placebo OR [95%-KI]		5,91 [0,27;128,38]	1,11 [0,07;18,55]
10 mg vs. Placebo OR [95%-KI]		7,56 [0,35;126,25]	1,41 [0,08;23,76]
Interaktionstest			
5 mg vs.Placebo		I ² =0%; p=0,93	I ² =0%; p=0,80
10 mg vs. Placebo		I ² =0%;p=0,74	I ² =0%; p=0,60
48 Wochen			
Alter <65 Jahre			
Sulfonylharnstoff +Dapagliflozin 5 mg	87	5 (5,7)	5 (5,7)
Sulfonylharnstoff +Dapagliflozin 10 mg	106	6 (5,7)	6 (5,7)
Sulfonylharnstoff +Placebo	82	1 (1,2)	7 (8,5)
5 mg vs.Placebo OR [95%-KI]		4,94 [0,56;43,21]	0,65 [0,20;2,15]
10 mg vs. Placebo OR [95%-KI]		4,86 [0,57;41,19]	0,64 [0,21;1,99]
48 Wochen			
Alter ≥65 und <75 Jahre			
Sulfonylharnstoff +Dapagliflozin 5 mg	29	2 (6,9)	1 (3,4)
Sulfonylharnstoff +Dapagliflozin 10 mg	23	2 (8,7)	2 (8,7)
Sulfonylharnstoff +Placebo	32	1 (3,1)	1 (3,1)
5 mg vs.Placebo OR [95%-KI]		5,91 [0,27;128,38]	1,11 [0,07;18,55]
10 mg vs. Placebo OR [95%-KI]		2,95 [0,25;34,68]	2,95 [0,25;34,68]
Interaktionstest			
5 mg vs.Placebo		I ² =0%; p=0,93	I ² =0%; p=0,74
10 mg vs. Placebo		I ² =0%;p=0,76	I ² =17,7%; p=0,27

Auch in der Anzahl der Ereignisse, die auf eine Genital- oder Harnwegsinfektion hinweisen können, zeigte sich keine Effektmodifikation bezüglich des Alters.

4.3.1.3.2.18 Unerwünschte Ereignisse stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT

Tabelle 4-77: Ergebnisse für Unerwünschte Ereignisse stratifiziert nach nach asiatischer und nicht-asiatischer Abstammung - RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate unerwünschter Ereignisse, n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
D1690C00005				
24 Wochen				
Asiaten				
Sulfonylharnstoff+Dapagliflozin 5 mg	43	18 (41,9)	5 (11,6)	0
Sulfonylharnstoff +Dapagliflozin 10 mg	40	24 (60,0)	1 (2,5)	0
Sulfonylharnstoff +Placebo	39	23 (59,0)	4 (10,3)	1 (2,6)
5 mg vs.Placebo OR [95%-KI]		0,50 [0,21;1,21]	1,15 [0,29;4,63]	0,30 [0,01;7,46]
10 mg vs. Placebo OR [95%-KI]		1,04 [0,42;2,56]	0,22 [0,02;2,10]	0,32 [0,01;8,02]
24 Wochen				
Nicht-Asiaten				
Sulfonylharnstoff +Dapagliflozin 5 mg	73	33 (45,2)	3 (4,1)	3 (4,1)
Sulfonylharnstoff +Dapagliflozin 10 mg	89	39 (43,8)	6 (6,7)	1 (1,1)
Sulfonylharnstoff +Placebo	75	30 (40,0)	1 (1,3)	1 (1,3)
5 mg vs.Placebo OR [95%-KI]		1,24 [0,64;2,38]	3,17 [0,32;31,21]	3,17 [0,32;31,21]
10 mg vs. Placebo OR [95%-KI]		1,17 [0,63;2,18]	5,35 [0,63;45,47]	0,84 [0,05;13,68]
Interaktionstest				
5 mg vs.Placebo		I ² =61,8%; p=0,11	I ² =0%; p=0,46	I ² =27,7%; p=0,24
10 mg vs. Placebo		I ² =0%; p=0,84	I ² =75,2%; p=0,04	I ² =0%; p=0,65
48 Wochen				
Asiaten				
Sulfonylharnstoff +Dapagliflozin 5 mg	43	25 (58,1)	7 (16,3)	0
Sulfonylharnstoff +Dapagliflozin 10 mg	40	28 (70,0)	3 (7,5)	0
Sulfonylharnstoff +Placebo	39	27 (69,2)	6 (15,4)	2 (5,1)
5 mg vs.Placebo OR [95%-KI]		0,62 [0,25;1,53]	1,07 [0,33;3,51]	0,17 [0,01;3,71]
10 mg vs. Placebo OR [95%-KI]		1,04 [0,40;2,71]	0,45 [0,10;1,93]	0,19 [0,01;3,98]
48 Wochen				
Nicht-Asiaten				
Sulfonylharnstoff +Dapagliflozin 5 mg	73	42 (57,5)	6 (8,2)	3 (4,1)
Sulfonylharnstoff +Dapagliflozin 10 mg	89	46 (51,7)	8 (9,0)	1 (1,1)
Sulfonylharnstoff +Placebo	75	36 (48,0)	2 (2,7)	1 (1,3)
5 mg vs.Placebo OR [95%-KI]		1,47 [0,77;2,81]	3,27 [0,64;16,75]	3,17 [0,32;31,21]
10 mg vs. Placebo OR [95%-KI]		1,16 [0,63;2,14]	3,60 [0,74;17,53]	0,84 [0,05;13,68]

Interaktionstest			
5 mg vs. Placebo		$I^2=56,6\%$; $p=0,13$	$I^2=14,8\%$; $p=0,28$
10 mg vs. Placebo		$I^2=0\%$; $p=0,85$	$I^2=72,3\%$; $p=0,06$

Bei der Gesamtrate unerwünschter Ereignisse zeigte sich eine positive Interaktion im Vergleich Sulfonylharnstoff+Dapagliflozin 5 mg gegenüber Sulfonylharnstoff+Placebo ($I^2=61,8\%$; $p=0,11$) nach Woche 24 und nach 48 Wochen ($I^2=56,6\%$; $p=0,13$), der Behandlungseffekt änderte sich jedoch nicht (alle Effektschätzer nicht signifikant). Zudem ließ sich keine Interaktion für die 10 mg Gruppe nachweisen. Bezüglich der schwerwiegenden unerwünschten Ereignisse gab es ebenfalls signifikante Interaktionstest ($I^2=75,2\%$; $p=0,04$ und $I^2=72,3\%$; $p=0,06$), hier allerdings nur für die 10 mg Gruppe und nicht für die 5 mg Gruppe. Erneut war kein Effektschätzer statistisch signifikant. Insgesamt ist daher bzgl. unerwünschter Ereignisse nicht von einer fazitrelevanten Modifikation des Behandlungseffektes durch die Abstammung auszugehen.

Tabelle 4-78: Ergebnisse für Ergebnisse für Anzahl der Patienten mit Anzeichen einer Genital- oder Harnwegsinfektion stratifiziert nach asiatischer und nicht-asiatischer Abstammung - RCT mit dem zu bewertenden Arzneimittel

Studie	N	Anzahl der Patienten mit Anzeichen auf Genitalinfektion, n (%)	Anzahl der Patienten mit Anzeichen auf Harnwegsinfektion, n (%)
D1690C00005			
24 Wochen			
Asiaten			
Sulfonylharnstoff +Dapagliflozin 5 mg	43	2 (4,7)	0
Sulfonylharnstoff +Dapagliflozin 10 mg	40	1 (2,5)	3 (2,5)
Sulfonylharnstoff +Placebo	39	1 (2,6)	1 (2,6)
5 mg vs. Placebo OR [95%-KI]		1,85 [0,16;21,28]	0,30 [0,01;7,46]
10 mg vs. Placebo OR [95%-KI]		0,97 [0,06;16,14]	3,08 [0,31;30,98]
24 Wochen			
Nicht-Asiaten			
Sulfonylharnstoff +Dapagliflozin 5 mg	73	5 (6,8)	5 (6,7)
Sulfonylharnstoff +Dapagliflozin 10 mg	89	6 (6,7)	2 (2,2)
Sulfonylharnstoff +Placebo	75	0	5 (6,7)
5 mg vs. Placebo OR [95%-KI]		12,12 [0,66;223,33]	1,03 [0,29;3,72]
10 mg vs. Placebo OR [95%-KI]		11,75 [0,65;212,19]	0,32 [0,06;1,71]
Interaktionstest			
5 mg vs. Placebo		I ² =0%; p=0,33	I ² =0%; p=0,48
10 mg vs. Placebo		I ² =31,8%; p=0,23	I ² =58,6%; p=0,12
48 Wochen			
Asiaten			
Sulfonylharnstoff +Dapagliflozin 5 mg	43	2 (4,7)	1 (2,3)
Sulfonylharnstoff +Dapagliflozin 10 mg	40	2 (5,0)	5 (12,5)
Sulfonylharnstoff +Placebo	39	2 (5,1)	1 (2,6)
5 mg vs. Placebo OR [95%-KI]		0,30 [0,12;6,73]	0,90 [0,05;14,97]
10 mg vs. Placebo OR [95%-KI]		0,97 [0,13;7,28]	5,43 [0,60;48,78]
48 Wochen			
Nicht-Asiaten			
Sulfonylharnstoff +Dapagliflozin 5 mg	73	5 (6,8)	5 (6,8)
Sulfonylharnstoff +Dapagliflozin 10 mg	89	6 (6,7)	3 (3,4)
Sulfonylharnstoff +Placebo	75	0	7 (9,3)
5 mg vs. Placebo OR [95%-KI]		12,12 [0,66;223,33]	0,71 [0,22;2,36]
10 mg vs. Placebo OR [95%-KI]		11,75 [0,65;212,19]	0,34 [0,08;1,36]
Interaktionstest			
5 mg vs. Placebo		I ² =51,7%; p=0,15	I ² =0%; p=0,88
10 mg vs. Placebo		I ² =47,9%; p=0,17	I ² =77,2%; p=0,04

Bezüglich der Abstammung ergaben sich signifikante Interaktionstests, deren Aussagekraft aber durch die geringe Anzahl an Ereignissen limitiert wird. Auch waren die Effekte in allen Subgruppen statistisch nicht signifikant. Es kann daher aus diesen Befunden nicht auf eine fazitrelevante Modifikation des Behandlungseffektes geschlossen werden.

4.3.1.3.2.19 Kardiale Ereignisse stratifiziert nach dem Geschlecht – RCT

Tabelle 4-79: Ergebnisse für Kardiale Ereignisse stratifiziert nach Geschlecht aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
D1690C00005		
24 Wochen		
Männlich		
Sulfonylharnstoff +Dapagliflozin 5 mg	58	4 (6,9)
Sulfonylharnstoff +Dapagliflozin 10 mg	56	2 (3,6)
Sulfonylharnstoff +Placebo	58	2 (3,4)
5 mg vs.Placebo OR [95%-KI]		2,07 [0,36;11,79]
10 mg vs. Placebo OR [95%-KI]		1,01 [0,14;7,63]
24 Wochen		
Weiblich		
Sulfonylharnstoff +Dapagliflozin 5 mg	58	0
Sulfonylharnstoff +Dapagliflozin 10 mg	73	1 (1,4)
Sulfonylharnstoff +Placebo	56	1 (1,8)
5 mg vs.Placebo OR [95%-KI]		0,32 [0,01;7,93]
10 mg vs. Placebo OR [95%-KI]		0,76 [0,05;12,49]
Interaktionstest		
5 mg vs.Placebo		I ² =1,4%; p=0,31
10 mg vs. Placebo		I ² =0%; p=0,86
48 Wochen		
Männlich		
Sulfonylharnstoff +Dapagliflozin 5 mg	58	6 (10,3)
Sulfonylharnstoff +Dapagliflozin 10 mg	56	2 (3,6)
Sulfonylharnstoff +Placebo	58	3 (5,2)
5 mg vs.Placebo OR [95%-KI]		2,12 [0,50;8,90]
10 mg vs. Placebo OR [95%-KI]		0,68 [0,11;4,23]
48 Wochen		
Weiblich		
Sulfonylharnstoff +Dapagliflozin 5 mg	58	1 (1,7)
Sulfonylharnstoff +Dapagliflozin 10 mg	73	2 (2,7)
Sulfonylharnstoff +Placebo	56	3 (5,4)
5 mg vs.Placebo OR [95%-KI]		0,31 [0,03;3,07]
10 mg vs. Placebo OR [95%-KI]		0,50 [0,08;3,08]
Interaktionstest		
5 mg vs.Placebo		I ² =48,3%; p=0,16
10 mg vs. Placebo		I ² =0%; p=0,81

Bezüglich des Geschlechts ergab sich ein signifikanter Interaktionstest (I²=48,3%; p=0,16 nach 48 Wochen), der aber für den Zeitpunkt 24 Wochen nicht zu beobachten war. Die Anzahl der Ereignisse war sehr gering und alle Effektschätzer waren statistisch nicht signifikant. Auf eine fazitrelevante Modifikation des Behandlungseffektes kann daher nicht geschlossen werden.

4.3.1.3.2.20 Kardiale Ereignisse stratifiziert nach dem Alter – RCT

Tabelle 4-80: Ergebnisse für Kardiale Ereignisse stratifiziert nach Alter aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
D1690C00005		
24 Wochen		
Alter <65 Jahre		
Sulfonylharnstoff +Dapagliflozin 5 mg	87	4 (4,6)
Sulfonylharnstoff +Dapagliflozin 10 mg	106	2 (1,9)
Sulfonylharnstoff +Placebo	82	2 (2,4)
5 mg vs.Placebo OR [95%-KI]		1,93 [0,34;10,82]
10 mg vs. Placebo OR [95%-KI]		0,77 [0,11;5,58]
24 Wochen		
Alter ≥65 und <75 Jahre		
Sulfonylharnstoff +Dapagliflozin 5 mg	29	0
Sulfonylharnstoff +Dapagliflozin 10 mg	23	1 (4,3)
Sulfonylharnstoff +Placebo	32	1 (3,1)
5 mg vs.Placebo OR [95%-KI]		0,36 [0,01;9,09]
10 mg vs. Placebo OR [95%-KI]		1,41 [0,08;23,76]
Interaktionstest		
5 mg vs.Placebo		I ² =0%; p=0,37
10 mg vs. Placebo		I ² =0%; p=0,73
48 Wochen		
Alter <65 Jahre		
Sulfonylharnstoff +Dapagliflozin 5 mg	87	7 (8,0)
Sulfonylharnstoff +Dapagliflozin 10 mg	106	2 (1,9)
Sulfonylharnstoff +Placebo	82	4 (4,9)
5 mg vs.Placebo OR [95%-KI]		1,71 [0,48;6,06]
10 mg vs. Placebo OR [95%-KI]		0,38 [0,07;2,10]
48 Wochen		
Alter ≥65 und <75 Jahre		
Sulfonylharnstoff +Dapagliflozin 5 mg	29	0
Sulfonylharnstoff +Dapagliflozin 10 mg	23	2 (8,7)
Sulfonylharnstoff +Placebo	32	2 (6,3)
5 mg vs.Placebo OR [95%-KI]		0,21 [0,01;4,49]
10 mg vs. Placebo OR [95%-KI]		1,43 [0,19;10,96]
Interaktionstest		
5 mg vs.Placebo		I ² =35,2%; p=0,21
10 mg vs. Placebo		I ² =0%; p=0,33

Es ergaben sich keine signifikanten Interaktionen bezüglich des Alters.

4.3.1.3.2.21 Kardiale Ereignisse stratifiziert nach asiatischer und nicht-asiatischer Abstammung – RCT

Tabelle 4-81: Ergebnisse für Kardiale Ereignisse stratifiziert nach asiatischer und nicht-asiatischer Abstammung aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
D1690C00005		
24 Wochen		
Asiaten		
Sulfonylharnstoff +Dapagliflozin 5 mg	43	1 (2,3)
Sulfonylharnstoff +Dapagliflozin 10 mg	40	0
Sulfonylharnstoff +Placebo	39	0
5 mg vs. Placebo OR [95%-KI]		2,79 [0,11;70,47]
10 mg vs. Placebo OR [95%-KI]		Nicht berechenbar
24 Wochen		
Nicht-Asiaten		
Sulfonylharnstoff +Dapagliflozin 5 mg	73	3 (4,1)
Sulfonylharnstoff +Dapagliflozin 10 mg	89	3 (3,4)
Sulfonylharnstoff +Placebo	75	3 (4,0)
5 mg vs. Placebo OR [95%-KI]		1,03 [0,20;5,27]
10 mg vs. Placebo OR [95%-KI]		0,84 [0,16;4,28]
Interaktionstest		
5 mg vs. Placebo		I ² =0%;p=0,59
10 mg vs. Placebo		Nicht berechenbar
48 Wochen		
Asiaten		
Sulfonylharnstoff +Dapagliflozin 5 mg	43	2 (4,7)
Sulfonylharnstoff +Dapagliflozin 10 mg	40	0
Sulfonylharnstoff +Placebo	39	0
5 mg vs. Placebo OR [95%-KI]		4,76 [0,22;102,26]
10 mg vs. Placebo OR [95%-KI]		Nicht berechenbar
48 Wochen		
Nicht-Asiaten		
Sulfonylharnstoff +Dapagliflozin 5 mg	73	5 (6,8)
Sulfonylharnstoff +Dapagliflozin 10 mg	89	4 (4,5)
Sulfonylharnstoff +Placebo	75	6 (8,0)
5 mg vs. Placebo OR [95%-KI]		0,85 [0,25;2,90]
10 mg vs. Placebo OR [95%-KI]		0,54 [0,15;1,99]
Interaktionstest		
5 mg vs. Placebo		I ² =0%;p=0,31
10 mg vs. Placebo		Nicht berechenbar

Es wurden keine signifikanten Interaktionen bezüglich der Abstammung der Patienten gefunden.

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 medizinische Nutzenbewertung eingeschlossene placebokontrollierte RCT D1690C00005 war eine Zulassungsstudie für Dapagliflozin. Im Rahmen des Zulassungsprozesses wurde die Zielpopulation für Dapagliflozin eingeschränkt auf Patienten unter 75 Jahren mit einer eGFR ≥ 60 mL/min/1,73 m² oder einem Kreatininclearance ≥ 60 mL/min sowie Patienten ohne Vorbehandlung mit Schleifendiuretika. Die Population der Patienten, für die Dapagliflozin zugelassen ist, ist die für die Nutzenbewertung primäre Population und wird im Folgenden jeweils für die beiden zugelassenen Dosierungen 5 mg und 10 mg betrachtet. Die Resultate der gesamten Studienpopulation wurden nur dann beschrieben, wenn Sie von denen der Zielpopulation abweichen.

Bezüglich des **Ausmaßes der Blutzuckerkontrolle** (HbA1c) ergaben sich nach 24 Wochen (LOCF) statistisch signifikante Unterschiede sowohl zwischen Sulfonylharnstoff+Dapagliflozin 5 mg und Sulfonylharnstoff+Placebo (MWD [95%-KI]: -0,50% [-0,70;-0,31]). als auch zwischen Sulfonylharnstoff+Dapagliflozin 10 mg und Sulfonylharnstoff+Placebo (MWD [95%-KI]: -0,74% [-0,92;-0,55]). Dieser Effekt blieb auch nach 48 Wochen (*Repeated-measurements* Analyse) bestehen. Hier ergab die MWD [95%-KI] zwischen der Sulfonylharnstoff+Dapagliflozin 5 mg und der Sulfonylharnstoff+Placebo-Gruppe -0,56% [-0,83;-0,30] und -0,77% [-1,02;-0,52] zwischen der Sulfonylharnstoff+Dapagliflozin 10 mg und Sulfonylharnstoff+Placebo-Gruppe.

Die unterstützenden Wirksamkeitsmessungen (Anteil von Patienten, die die Studie aufgrund von unzureichender Blutzuckerkontrolle abgebrochen haben oder die Notfallmedikation, die aufgrund unzureichender Blutzuckerkontrolle verabreicht wurde, bekommen haben) ergaben in den Sulfonylharnstoff+Dapagliflozin 5 mg/10 mg-Gruppen gegenüber der Sulfonylharnstoff+Placebo-Gruppe nach 24 und 48 Wochen signifikant geringere Raten (OR [95%-KI] nach 24 Wochen 0,25 [0,10;0,63] / 0,07 [0,02;0,30] und nach 48 Wochen 0,35 [0,20;0,61] / 0,18 [0,10;0,33]).

Hinsichtlich der in Subgruppenanalysen untersuchten Faktoren Alter und Geschlecht ergaben sich keine Hinweise auf Effektmofifikation, alle Interaktionstest zeigten entweder p-Werte $> 0,2$ oder konnten aufgrund der Konstellation des Befundes (unter Berücksichtigung von Konsistenz; Robustheit der beobachteten Interaktion; Plausibilität; und zugrunde liegende Gruppengröße) als nicht fazitrelevant identifiziert werden. Bei der Abstammung (asiatisch, nicht-asiatisch) ergaben sich signifikante Interaktionstests ($I^2=53,3\%$; $p=0,14$ und $I^2=63,5\%$; $p=0,09$), jedoch war der Effekt gleichgerichtet zugunsten der Sulfonylharnstoff+Dapagliflozin

5 mg/10 mg-Gruppen. Betrachtet man die absoluten Veränderungen unter Dapagliflozin 5 mg bzw. 10 mg sieht man, dass die HbA1c-Senkung in beiden Abstammungsgruppen ähnlich war (Dapagliflozin 5 mg: Patienten asiatischer Abstammung -0,63%, Patienten nicht-asiatischer Abstammung: -0,63; Dapagliflozin 10 mg: Patienten asiatischer Abstammung: -0,87, Patienten nicht-asiatischer Abstammung -0,85). In der Placebogruppe kam es auch bei den Patienten nicht-asiatischer Abstammung zu einer HbA1c-Senkung während es bei in der Gruppe mit den asiatischen Patienten zu einer Erhöhung des HbA1c-Wertes kam (-0,24% versus 0,07%). Dadurch kommt es in dieser Subgruppe zu den beobachtenden Behandlungsunterschieden und damit auch zur signifikanten Interaktion. Zu beobachten bleibt aber in allen betrachteten Subgruppen die Dosisabhängigkeit des Behandlungseffektes, sowie die signifikante Überlegenheit von Dapagliflozin gegenüber Placebo. Bei der Interpretation muss zudem die geringe Anzahl von asiatischen Patienten in der Placebogruppe berücksichtigt werden, die es problematisch erscheinen lässt, hieraus eindeutige Schlussfolgerungen abzuleiten. Insgesamt wurde aus den aufgezählten Gründen die beobachtete Interaktion als nicht fazitrelevant angesehen.

Hinsichtlich der Gesamtrate an **Hypoglykämien** ergaben sich nach 24 Wochen keine signifikanten Unterschiede (Sulfonylharnstoff+Dapagliflozin 5 mg / 10 mg versus Sulfonylharnstoff+Placebo OR [95%-KI]: 1,61 [0,51;5,09] / 2,24 [0,76;6,55]). Auch nach 48 Wochen ergaben sich keine signifikanten Unterschiede (Sulfonylharnstoff+Dapagliflozin 5 mg / 10 mg versus Sulfonylharnstoff+Placebo OR [95%-KI]: 1,53 [0,60;3,89] / 2,01 [0,83;4,86]). Schwerwiegende Hypoglykämien oder Studienabbrüche aufgrund von Hypoglykämien traten während der gesamten Behandlungsphase (48 Wochen) nicht auf.

Der Interaktionstest nach dem Geschlecht war nach Woche 48 im Vergleich von Sulfonylharnstoff+Dapagliflozin 5 mg versus Sulfonylharnstoff+Placebo signifikant ($I^2=42,4\%$, $p=0,19$), aber aufgrund des gleichgerichteten Effekts (kein Unterschied) und der geringen Ereignisrate in den Subgruppen ergab sich daraus keine Änderung der aus der Bewertung der Gesamtpopulation abgeleiteten Beurteilung. Im Vergleich von Sulfonylharnstoff+Dapagliflozin 10 mg (Dosis) versus Sulfonylharnstoff+Placebo war der Interaktionstest nicht signifikant. ($I^2=14,1\%$; $p=0,28$). Ansonsten ergab sich für Hypoglykämien kein Hinweis auf eine Effektmodifikation durch das Alter oder die Abstammung.

Bezüglich der **Gewichtsreduktion** ergab sich nach 24 Wochen (LOCF) ein signifikanter Vorteil zugunsten von Sulfonylharnstoff+Dapagliflozin (beide Dosisgruppen) versus Sulfonylharnstoff+Placebo (MWD [95%-KI]: -0,89 kg [-1,62;-0,15] und -1,72 kg [-2,43;-1,01]). Nach 48 Wochen (*Repeated measurements* Analyse) war der Unterschied in der Sulfonylharnstoff+Dapagliflozin 5 mg-Gruppe nicht mehr signifikant (MWD [95%-KI]: 0,77 kg [-1,81;0,27]). In der Behandlungsgruppe Sulfonylharnstoff+Dapagliflozin 10 mg (Hauptdosis) blieb der Effekt im Vergleich zum Sulfonylharnstoff+Placebo bestehen (MWD [95%-KI]: -1,80 kg [-2,80;-0,80]).

Bis zu 25,5% der Patienten, welche 10 mg Dapagliflozin in Kombination mit einem Sulfonylharnstoff erhielten hatten nach 24 Wochen eine **Gewichtsreduktion von $\geq 5\%$** . Nach

48 Wochen waren es noch 16,2%. Diese Unterschiede waren signifikant sowohl für die 5 mg wie auch für die 10 mg Gruppe (OR [95%-KI]: 3,56 [1,26;10,09] / 7,42 [2,79;19,78] nach 24 Wochen und 3,16 [1,20;8,34] / 3,47 [1,35;8,93] nach 48 Wochen).

Für die Gewichtsveränderung ergab sich kein Hinweis auf eine Effektmodifikation durch das Alter. Der Interaktionstest nach Geschlecht zeigte zum Zeitpunkt 48 Wochen bei der Gewichtsreduktion von $\geq 5\%$ ein signifikantes Ergebnis sowohl in der 5 mg als auch in der 10 mg Dapagliflozin-Gruppe ($I^2=56,1\%$; $p=0,13$ und $I^2=49,9\%$; $p=0,16$). Aufgrund der geringen Patientenzahlen in den Subgruppen ergibt sich daraus keine Änderung des aus der Bewertung der Gesamtpopulation abgeleiteten Fazits. Auch ergab sich ein Hinweis auf eine Effektmodifikation durch die Abstammung. (Interaktionstest: $I^2=61,3\%$; $p=0,11$ zwischen Sulfonylharnstoff+Dapagliflozin 10 mg versus Sulfonylharnstoff+Placebo). Auch für die Operationalisierung Gewichtsverlust ($\geq 5\%$) ergab sich ein Hinweis auf eine Effektmodifikation durch die Abstammung der Patienten ($I^2=66,8\%$; $p=0,08$), jedoch nur nach 48 Wochen und zwischen Sulfonylharnstoff+Dapagliflozin 10 mg versus Sulfonylharnstoff+Placebo. Aufgrund der geringen Ereigniszahlen kann die Interaktion als nicht valide angesehen werden.

Nach 24 Wochen Behandlung mit Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Placebo ließen sich für den Endpunkt **Therapiezufriedenheit** (die Analyse wurde anhand berichteter Werte durchgeführt) keine signifikanten Unterschiede feststellen (MWD [95%-KI]: 1,20 [-0,89;3,29] und 1,40 [-0,43;3,23]). Auch nach 48 Wochen ließen sich keine signifikanten Unterschiede feststellen (MWD [95%-KI]: 2,40 [-0,27;5,07] und 2,10 [-0,48;4,68]). In der Studienpopulation war der Unterschied zwischen den Behandlungsgruppen nach 48 Wochen etwas größer, wodurch er hier signifikant wurde (MWD [95%-KI]: 2,40 [0,26;4,54] und 2,20 [0,07;4,33]) zu Gunsten der Kombination Sulfonylharnstoff+Dapagliflozin.

Für die Therapiezufriedenheit ergab sich kein Hinweis auf eine Effektmodifikation durch das Geschlecht und der Abstammung. Der Interaktionstest im Vergleich Sulfonylharnstoff+Dapagliflozin 10 mg und Sulfonylharnstoff+Placebo-Gruppe nach dem Alter ist signifikant ($I^2=53,8\%$; $p=0,14$). Der Behandlungseffekt zeigte jedoch in beiden Fällen keinen Unterschied.

Die **Blutdrucksenkung** war in der Sulfonylharnstoff+Dapagliflozin-Gruppe höher als in der Sulfonylharnstoff+Placebo-Gruppe. Der systolische Blutdruck konnte nach 24 Wochen als auch nach 48 Wochen signifikant stärker in der Sulfonylharnstoff+Dapagliflozin 10 mg-Gruppe als in der Sulfonylharnstoff+Placebo-Gruppe gesenkt werden (MWD [95%-KI]: -4,4 mmHg [-7,2;-1,6] nach 24 Wochen / -4,46 mmHg [-8,62;-0,31] nach 48 Wochen). Der Unterschied war für die Sulfonylharnstoff+Dapagliflozin 5 mg-Gruppe nach 48 Wochen nicht mehr signifikant. Auch der diastolische Blutdruck konnte sowohl nach 24 als auch nach 48 Wochen in der 10 mg-Gruppe signifikant stärker gesenkt werden (MWD [95%-KI]: -2,0 mmHg [-3,8;-0,2] nach 24 Wochen / -3,12 mmHg [-5,68;-0,56] nach 48 Wochen). Der Unterschied war für die Sulfonylharnstoff+Dapagliflozin 5 mg-Gruppe nach 24 Wochen nicht signifikant.

Für die Veränderung des systolischen Blutdrucks ergab sich kein Hinweis auf eine Effektmodifikation durch das Alter, Abstammung und das Geschlecht. Für die Veränderung des diastolischen Blutdrucks gab es signifikante Interaktionstests bezüglich des Alters (Sulfonylharnstoff+Dapagliflozin 5 mg vs. Sulfonylharnstoff+Placebo: $I^2=70,5\%$; $p=0,07$ / Sulfonylharnstoff+Dapagliflozin 10 mg vs. Sulfonylharnstoff+Placebo: $I^2=37,9\%$; $p=0,20$) und bezüglich der Abstammung ($I^2=75,7\%$; $p=0,04$) zu Gunsten der Nicht-Asiaten.

Zusammenfassend zeigte sich im Vergleich Sulfonylharnstoff+Dapagliflozin 5 mg/10 mg versus Sulfonylharnstoff+Placebo keine statistisch signifikant erhöhte **Gesamtrate unerwünschter Ereignisse** (OR [95%-KI]: 0,90 [0,54;1,52]/ 1,10 [0,66;1,82] nach 24 Wochen und 1,11 [0,66;1,86] /1,09 [0,66;1,81] nach 48 Wochen); keine signifikant erhöhte **Gesamtrate der schwerwiegenden unerwünschten Ereignisse** (OR [95%-KI]: 1,61 [0,51;5,09] / 1,25 [0,39;4,06] nach 24 Wochen und 1,67 [0,67;4,20] /1,24 [0,48;3,19] nach 48 Wochen) und keine statistisch erhöhte **Abbruchrate aufgrund unerwünschter Ereignisse** (OR [95%-KI]: 1,49 [0,24;9,07] /0,44 [0,04;4,89] nach 24 Wochen und 0,98 [0,19;4,97]/0,29 [0,03;2,82] nach 48 Wochen).

Bei der Anzahl der Patienten mit Anzeichen einer **Genitalinfektion** und Anzahl der Patienten mit Anzeichen einer **Harnwegsinfektion** ergaben sich keine statistisch signifikanten Unterschiede, jedoch sind die Anzeichen einer Genitalinfektion in der Sulfonylharnstoff+Dapagliflozin numerisch höher. Die OR [95%-KI] für die Anzeichen einer Genitalinfektion betrug 7,26 [0,88;59,96] / 6,48 [0,79;53,52] nach 24 Wochen und 3,60 [0,73;17,70] / 3,70 [0,77;17,81] nach 48 Wochen. Die OR [95%-KI] für die Anzeichen einer Harnwegsinfektion betrug 0,81 [0,24;2,74] / 0,73 [0,22;2,45] nach 24 Wochen und 0,72 [0,24;2,15] / 0,88 [0,32;2,42] nach 48 Wochen. Auch bei der Betrachtung der gesamten Studienpopulation gab es bei den Anzeichen auf eine Harnwegsinfektion sowohl nach 24 als auch nach 48 Wochen keine signifikanten Unterschiede, Hinweise auf Genitalinfektionen traten allerdings hier signifikant häufiger unter Sulfonylharnstoff+Dapagliflozin-Behandlung auf. Dabei gilt jedoch zu berücksichtigen, dass die Anzahl dieser Ereignisse primär nicht gesicherte Diagnosen, sondern hauptsächlich auf den Angaben der Patienten beruht, die in der Studie aktiv gebeten wurden Anzeichen und Symptome einer möglichen Genital- bzw. Harnwegsinfektion (wie z.B. Pruritus und Dysurie) zu melden. Diese Ereignisse waren mild bis moderat in ihrer Intensität, traten in den meisten Fällen nur einmal auf und waren in Falle einer wahren Infektion mit den gängigen Therapieoptionen, wie Antibiose einfach zu behandeln.

Hinsichtlich der **unerwünschten Ereignisse** ergaben sich keine Belege oder Hinweise auf eine Interaktion durch das Alter. Hinsichtlich des Geschlechts zeigte sich bei der Gesamtrate unerwünschter Ereignisse eine positive Interaktion im Vergleich Sulfonylharnstoff+Dapagliflozin 10 mg gegenüber Sulfonylharnstoff+Placebo ($I^2=41,3\%$; $p=0,19$) nach Woche 24. Jedoch ändert sich der Behandlungseffekt nicht und nach Woche 48 gab es keine Interaktion mehr ($I^2=0\%$; $p=0,84$). Bezüglich der Abstammung ergaben sich bei der Gesamtrate unerwünschter Ereignisse positive Interaktionen im Vergleich Sulfonylharnstoff+Dapagliflozin 5 mg gegenüber Sulfonylharnstoff+Placebo ($I^2=61,8\%$;

p=0,11) nach Woche 24 und nach 48 Wochen ($I^2=56,6\%$; p=0,13). Jedoch änderte sich der Behandlungseffekt nicht. Bezüglich der schwerwiegenden unerwünschten Ereignisse gab es ebenfalls signifikante Interaktionstests bei der 10 mg-Dosisgruppe ($I^2=75,2\%$; p=0,04 und $I^2=72,3\%$; p=0,06) aber aufgrund der geringen Ereigniszahlen kann dies nicht als valide angesehen werden.

Bei den Anzeichen auf Harnwegsinfektion gab es hinsichtlich des Geschlechts eine signifikante Interaktion nach Woche 48 ($I^2=48,1\%$; p=0,17) auch hinsichtlich der Abstammung ergaben sich signifikante Interaktionen, aber bei der geringen Anzahl der Ereignisse in den Subgruppen, lässt sich keine Effektmodifikation zeigen. Ansonsten gab es bezüglich der Hinweise auf Harnwegs- und Genitalinfektionen keine weiteren signifikanten Interaktionstests.

Bezüglich **kardialer Ereignisse** ergaben sich keine signifikanten Unterschiede zwischen den Behandlungsoptionen (Sulfonylharnstoff+Dapagliflozin 5 mg vs. Sulfonylharnstoff+Placebo / Sulfonylharnstoff+Dapagliflozin 10 mg vs. Sulfonylharnstoff+Placebo). Das OR [95%-KI] nach 24 Wochen betrug 1,32 [0,29;6,04] / 0,88 [0,17;4,45] und 1,16 [0,38;3,55] / 0,58 [0,16;2,09] nach 48 Wochen.

Hinsichtlich der in Subgruppenanalysen untersuchten Faktoren Alter und Abstammung ergaben sich keine Hinweise auf Effektmodifikation, alle Interaktionstests zeigten p-Werte $>0,2$. Bezüglich des Geschlechts fand sich ein signifikanter Interaktionstest in der Gruppe Sulfonylharnstoff+Dapagliflozin 5 mg vs. Sulfonylharnstoff+Placebo, aber bei der geringen Anzahl an Ereignissen ist dieser nicht interpretierbar.

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.

Es wurden zwei indirekte Vergleiche durchgeführt:

- Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin
- Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor

Beide indirekte Vergleiche wurden über den Brückenkomparator Sulfonylharnstoff+Placebo durchgeführt. Die Ergebnisse der Informationsbeschaffung, wie auch die Ergebnisse der indirekten Vergleiche selbst sind gemeinsam in den nachfolgenden Abschnitten dargestellt.

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*

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 4-82: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
MB102029	ja	abgeschlossen	104 Wochen	Dapagliflozin 5 mg+Vorbehandlung Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
D1690C00005	ja	abgeschlossen	48 Wochen	Dapagliflozin 2,5 mg+Glimepirid Dapagliflozin 5 mg+Glimepirid Dapagliflozin 10 mg+Glimepirid Placebo+ Glimepirid
MB102020	nein	abgebrochen	4 Wochen	Dapagliflozin 5 mg Dapagliflozin 20 mg
MB102035	nein	abgeschlossen	12 Wochen	Dapagliflozin 10 mg+Vorbehandlung (Metformin und/oder Sulfonylharnstoff) Placebo+Vorbehandlung (Metformin und/oder Sulfonylharnstoff)
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
MB102054	nein	abgeschlossen*	24 Wochen	Dapagliflozin 5 mg+Vorbehandlung Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
MB102073	nein	laufend	12 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung
MB102077	nein	laufend	12 Wochen	Dapagliflozin 10 mg+Vorbehandlung Placebo+Vorbehandlung

* Es liegt noch kein Studienbericht vor.

Geben Sie an, welchen Stand die Information in Tabelle 4-82 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 Studienstatus in Tabelle 4-82 ist zum 29.10.2012 abgebildet.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-82 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-83: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
MB102029	Die Studie untersucht nur Patienten mit Nierenfunktionsstörung (eGFR zwischen 30 mL/min/1,73 m ² und 59 mL/min/1,73 m ²). Für diese Patienten wird eine Behandlung mit Dapagliflozin nicht empfohlen.
MB102020	Phase-1-Studie. Die Studiendauer beträgt nur 4 Wochen.
MB102035	Die Vergleichstherapie ist Hydrochlorothiazid. Die Studiendauer ist zu gering.
D1690C00018 (inkl. LT I und II)	Die Vorbehandlung der Patienten ist heterogen. Eine Subgruppenanalyse nach Vorbehandlung wird für die von uns betrachteten Sulfonylharnstoffe (Glimepirid, Glibenclamid und Glipizid) nicht berichtet.
D1690C00019 (inkl. LT I und II)	Die Vorbehandlung der Patienten ist heterogen. Eine Subgruppenanalyse nach Vorbehandlung wird für die von uns betrachteten Sulfonylharnstoffe (Glimepirid, Glibenclamid und Glipizid) nicht berichtet.
MB102054	Die Studie wurde kürzlich abgeschlossen. Es liegen noch keine Ergebnisse vor. Die Studie wurde daher in Anhang 4-E aufgenommen.
MB102073	Die Studie läuft noch. Es liegen keine Ergebnisse vor. Die Studiendauer ist zu gering.
MB102077	Die Studie läuft noch. Es liegen keine Ergebnisse vor. Die Studiendauer ist zu gering.

eGFR: geschätzte glomeruläre Filtrationsrate

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

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.

Indirekter Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin

Zur Bewertung des medizinischen Zusatznutzens von Dapagliflozin wurde ein indirekter Vergleich herangezogen, da für das Anwendungsgebiet *Add-on* zu Sulfonylharnstoff keine Studien vorliegen, die einen direkten Vergleich zwischen dem zu bewertenden Arzneimittel und der zweckmäßigen Vergleichstherapie ermöglichen.

Wie aus Abschnitt 4.3.2.1.1.1 ersichtlich, ist die einzige identifizierte relevante RCT mit dem zu bewertenden Arzneimittel die Studie D1690C00005. In dieser Studie von BMS / AZ wurde das zu bewertende Arzneimittel Dapagliflozin in Kombination mit dem Sulfonylharnstoff Glimepirid verabreicht und mit der Kombinationsbehandlung Glimepirid und Placebo verglichen.

Der Vergleich mit der zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Metformin) ist daher durch einen indirekten Vergleich möglich, welcher Sulfonylharnstoff+Placebo als Brückenkomparator benötigt.

Da die Studie D1690C00005 den Sulfonylharnstoff Glimepirid verwendete, wurde zur Identifikation von relevanten Studien mit der zweckmäßigen Vergleichstherapie zunächst eine umfassende bibliographische Literaturrecherche nach randomisierten, kontrollierten Studien mit Metformin in Kombination mit Glimepirid durchgeführt.

Diese Literaturrecherche vom 23.10.2012 erzielte insgesamt 547 Treffer. Nach automatisiertem Ausschluss der Duplikate (n=127) wurden die verbleibenden 420 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den in Abschnitt 4.2.2 definierten Kriterien selektiert. Die Studien wurden von zwei Personen unabhängig voneinander nach Relevanz selektiert. Basierend auf den Informationen des Titels und Abstracts wurden 419 Publikationen begründet durch Ausschlusskriterien als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Die verbleibende Publikation, deren Titel und Abstract die Einschlusskriterien erfüllte, wurde im Anschluss der Volltext hinsichtlich dieser Kriterien überprüft und aus der Bewertung ausgeschlossen (Ausschlussgrund: nicht passende Patientenpopulation, Anhang 4-C, Tabelle 4-147)

Somit ergab die bibliographische Literaturrecherche keine relevante Publikation.

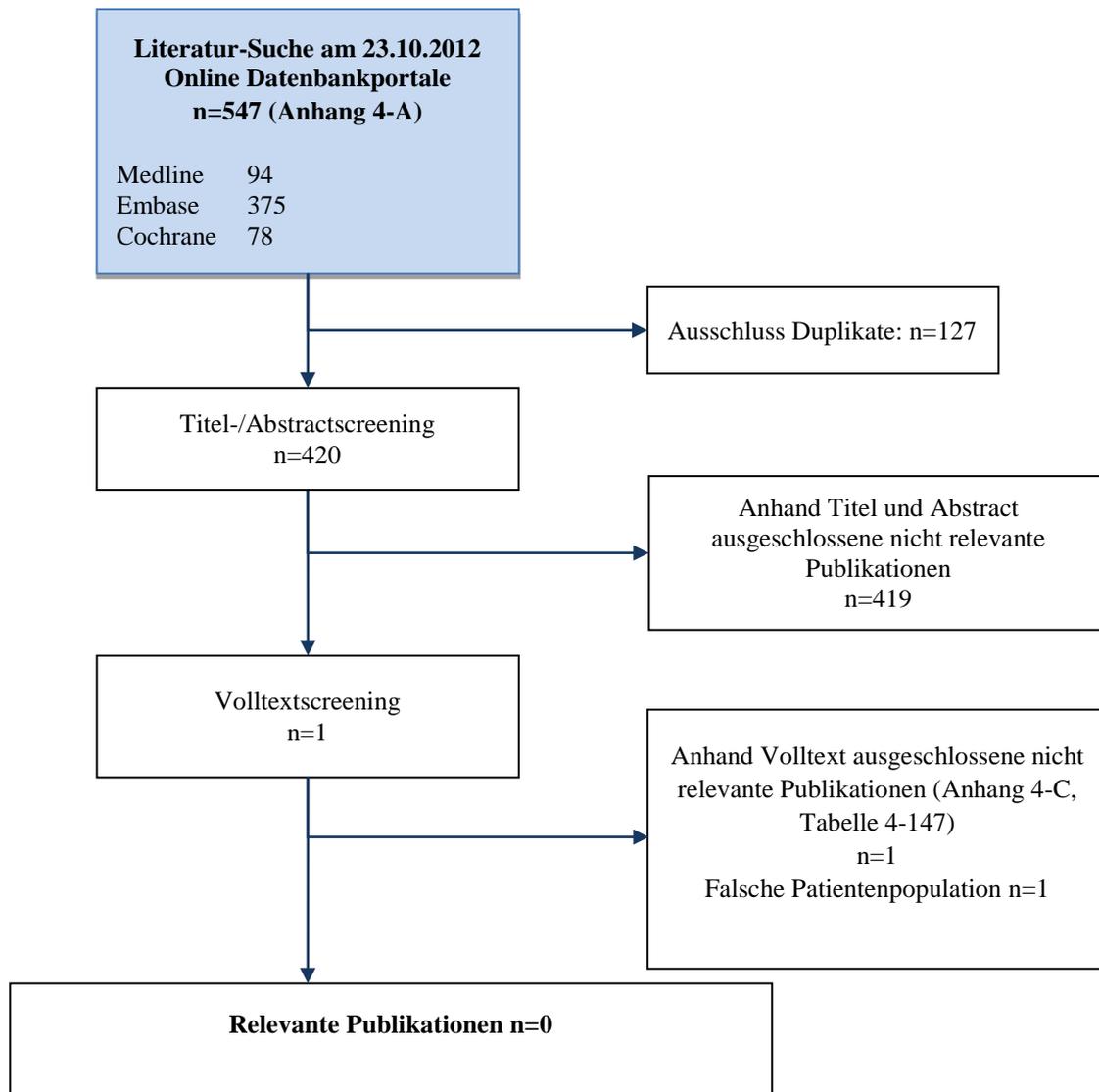


Abbildung 9: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie Metformin in Kombination mit Glimperid

Um die Evidenzlage zu erhöhen, wurde eine zweite Literaturrecherche durchgeführt. Hier wurde nach randomisierten, kontrollierten Studien mit Metformin in Kombination mit den Sulfonylharnstoffen Glibenclamid oder Glipizid gesucht.

Die umfassende bibliographische Literaturrecherche vom 23.10.2012 nach randomisierten, kontrollierten Studien mit den zweckmäßigen Vergleichstherapien Metformin in Kombination mit Glibenclamid oder Glipizid erzielte insgesamt 1012 Treffer. Nach automatisiertem Ausschluss der Duplikate (n=270) wurden die verbleibenden 742 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den in Abschnitt 4.2.2 definierten Kriterien eingegrenzt. Die Studien wurden von zwei Personen unabhängig

voneinander nach Relevanz selektiert. Basierend auf den Informationen des Titels und Abstracts wurden 729 Publikationen begründet durch Ausschlusskriterien als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Für die verbleibenden 13 Publikationen, deren Titel oder Abstract die Einschlusskriterien erfüllen, wurden im Anschluss die Volltexte hinsichtlich dieser Kriterien überprüft. Im Ergebnis entsprachen 11 nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C, Tabelle 4-147. Ausschlussgründe: Patientenpopulation, Studiendauer, Endpunkte und/oder Studiendesign).

Somit ergab die bibliographische Literaturrecherche zwei relevante Publikationen. Bei den identifizierten Studien handelt es sich um die Publikationen von DeFronzo et al., 1995 und Goldstein et al., 2003. Für die von DeFronzo et al. 1995 publizierte Studie lag zudem in den Archiven von BMS der vollständige Studienbericht vor (Lipha Pharmaceutical Inc., 1992).

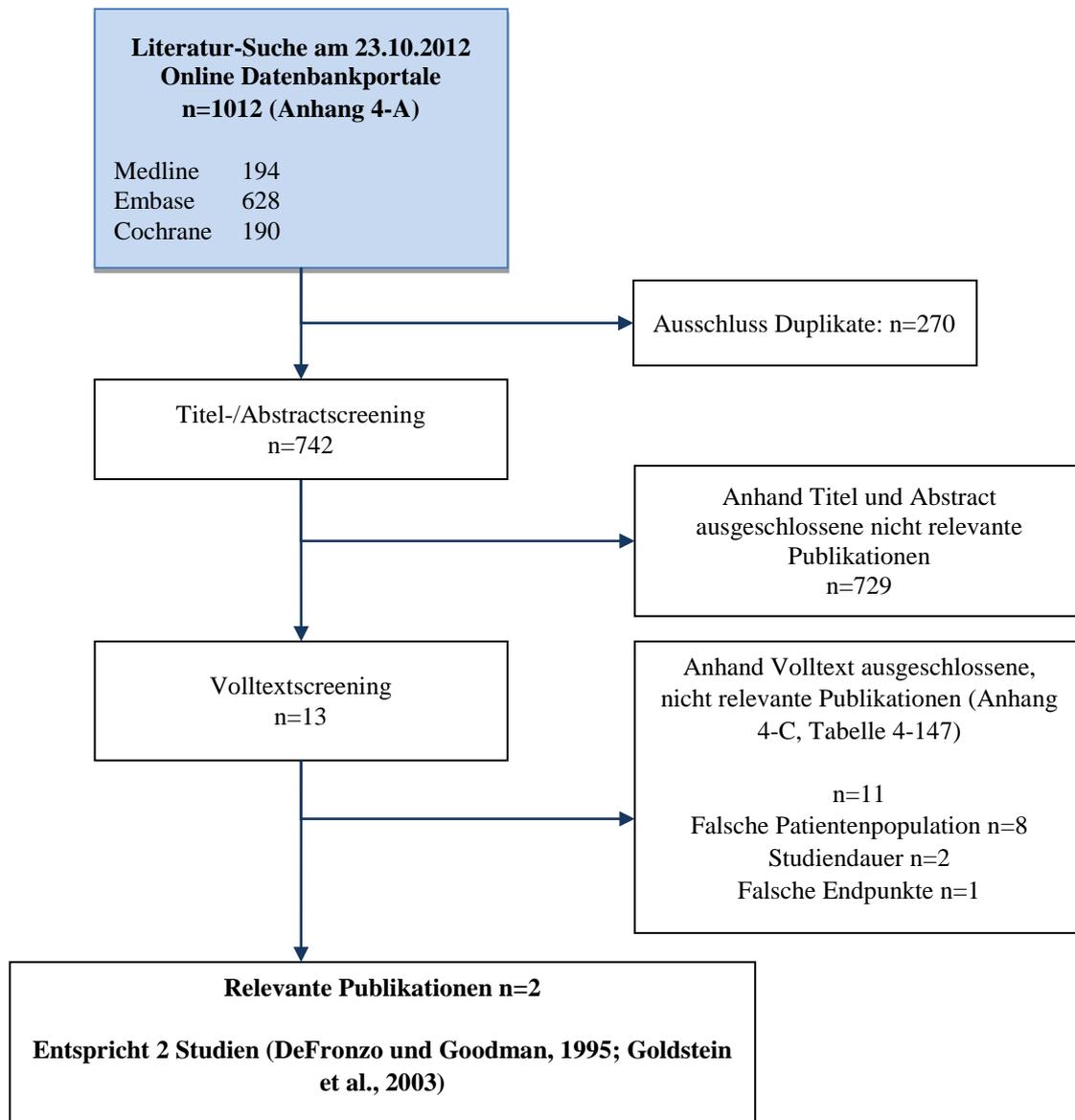


Abbildung 10: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie Metformin in Kombination mit Glibenclamid oder Glipizid

Indirekter Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor

Die Literaturrecherche nach Studien zu Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo vom 17.10.2012 erzielte insgesamt 418 Treffer. Nach Ausschluss der Duplikate (n=33) wurden die verbleibenden 385 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den in Abschnitt 4.2.2 definierten Kriterien selektiert. Die Studien wurden von zwei Personen unabhängig voneinander nach Relevanz selektiert. Basierend auf den Informationen des Titels und Abstracts wurden 342 Publikationen

begründet durch Ausschlusskriterien als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Die verbleibenden 43 Publikationen wurden im Anschluss im Volltext gesichtet und hinsichtlich der Ein- und Ausschlusskriterien überprüft. Insgesamt 37 wurden begründet ausgeschlossen (Ausschlussgründe: Review/Meta-Analysen, unpassende Behandlungsgruppen, kein RCT, falsche Population, oder zu kurze Beobachtungsdauer. Siehe Anhang 4-C, Tabelle 4-147 und Abbildung 11). Als relevant identifiziert wurden sechs Publikationen (Hermansen et al., 2007; Garber et al., 2008; Chacra et al., 2009; Pratley et al., 2009; Chacra et al., 2011; Lewin et al., 2012). Die Studie von Chacra 2011 konnte allerdings für einen indirekten Vergleich nicht verwendet werden, da sie nur Daten zu Wirksamkeit und Sicherheit nach Woche 76 berichtet. Da keine Analysen anderer Studien zu diesem Zeitpunkt vorlagen, wurde die Publikation von Chacra 2011 ebenfalls begründet ausgeschlossen (Anhang 4-C).

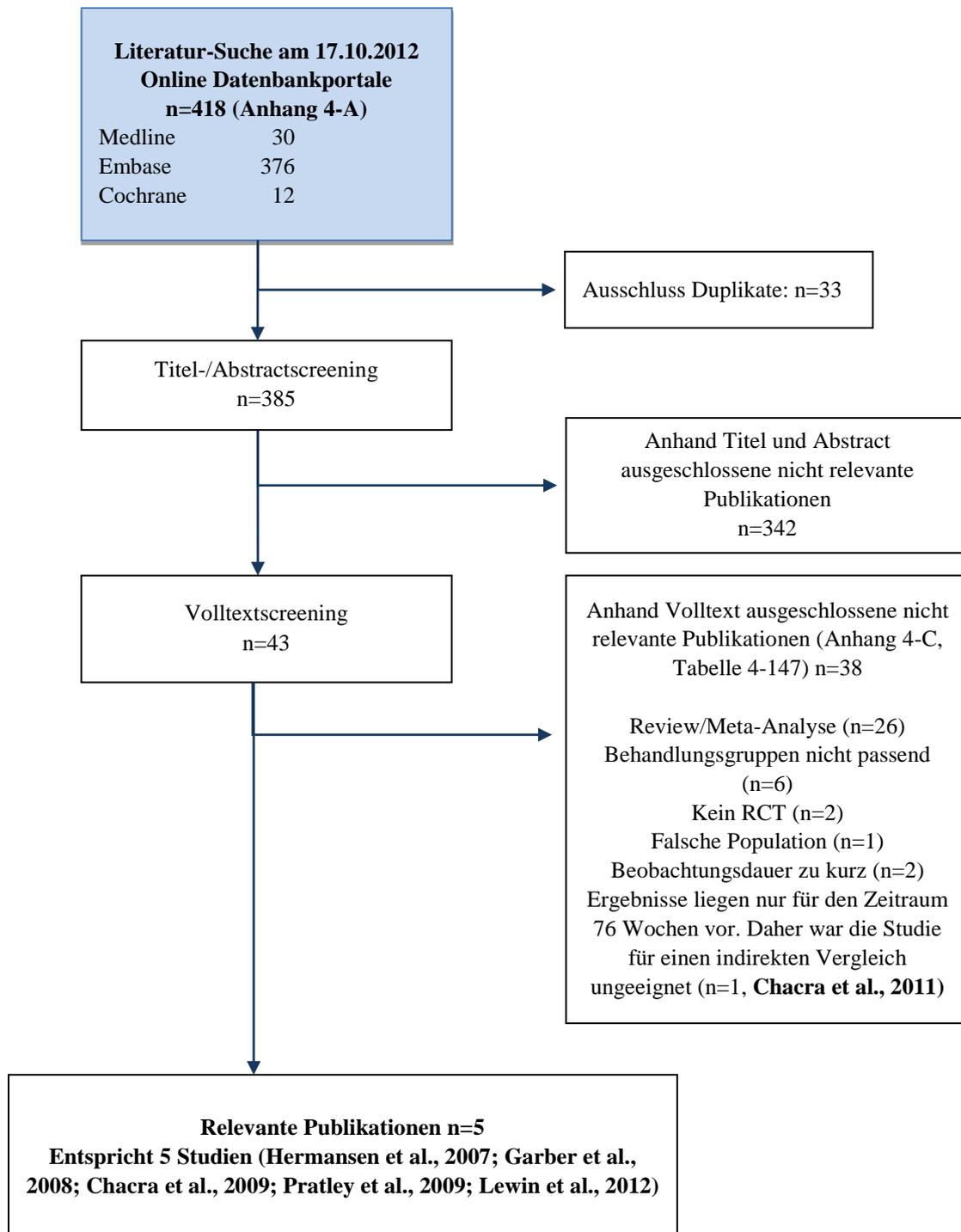


Abbildung 11: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der alternativen Vergleichstherapie DPP-4-Inhibitor in Kombination mit Sulfonylharnstoff

4.3.2.1.1.3 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 4-5) 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.

Die für indirekte Vergleiche verwendeten Studien aus der Studienregistersuche nach RCT mit dem zu bewertenden Arzneimittel sind in Tabelle 4-7 aufgeführt.

Indirekter Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin

Von den Studien aus der Studienregistersuche nach RCT mit der zweckmäßigen Vergleichstherapie erfüllte keine die in Abschnitt 4.2.2 genannten Einschlusskriterien.

Tabelle 4-84: Relevante Studien aus der Suche in Studienregistern – RCT mit der zweckmäßigen Vergleichstherapie

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)
-	-	-	-

a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.
b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.

Indirekter Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor

Von den Studien aus der Studienregistersuche nach RCT mit der alternativen Vergleichstherapie erfüllten fünf die in Abschnitt 4.2.2 genannten Einschlusskriterien. Alle Studien sind bereits in der Literaturrecherche identifiziert worden.

Tabelle 4-85: Relevante Studien aus der Suche in Studienregistern – RCT mit der alternativen Vergleichstherapie

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)
Pratley et al. 2009	clinicaltrials.gov [NCT00286468] (Clinicaltrials.gov, 2012b)	nein	ja
Garber et al. 2008	clinicaltrials.gov [NCT00099944] (Clinicaltrials.gov, 2012c)	nein	ja
Chacra et al. 2009	clinicaltrials.gov [NCT00313313] (Clinicaltrials.gov, 2012d)	nein	ja
Hermansen et al. 2007	clinicaltrials.gov [NCT00106704] (Clinicaltrials.gov, 2012e)	nein	ja
Lewin et al. 2012	clinicaltrials.gov [NCT00819091] (Clinicaltrials.gov, 2012a)	nein	ja
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.			

4.3.2.1.1.4 Resultierender Studienpool: RCT für indirekte Vergleiche

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.2.1.1.1, 4.3.2.1.1.2 und 4.3.2.1.1.3) 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 4-86: Studienpool – RCT für indirekte Vergleiche

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponsernte 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, zu bewertendes Arzneimittel						
D1690C00005	ja	ja	nein	ja (AstraZeneca, 2010a) (AstraZeneca, 2010b)	ja [NCT00680745] (Clinicaltrials.gov, 2012l) [EUCTR2007-005931-27-HU] (ICTRP Search Portal, 2012d) [NCT00680745] (ICTRP Search Portal, 2012c)	ja (Strojek et al., 2011)
placebokontrolliert, zweckmäßige Vergleichstherapie(n) Sulfonylharnstoff+Metformin						
DeFronzo 1995	nein	nein	ja	ja (Lipha Pharmaceutical Inc., 1992)	nein	ja (DeFronzo und Goodman, 1995)
Goldstein 2003	nein	nein	ja	nein	nein	ja (Goldstein et al., 2003)
placebokontrolliert, alternative Vergleichstherapie(n) Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	nein	nein	ja	nein	clinicaltrials.gov [NCT00106704] (Clinicaltrials.gov, 2012e)	ja (Hermansen et al., 2007)
Chacra et al. 2009	nein	ja	nein	nein	clinicaltrials.gov [NCT00313313] (Clinicaltrials.gov, 2012d)	ja (Chacra et al., 2009)

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gespon-serte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
Garber et al. 2008	nein	nein	ja	nein	clinicaltrials.gov [NCT00099944] (Clinicaltrials.gov, 2012c)	ja (Garber et al., 2008)
Pratley et al. 2009	nein	nein	ja	nein	clinicaltrials.gov [NCT00286468] (Clinicaltrials.gov, 2012b)	ja (Pratley et al., 2009)
Lewin et al. 2012	nein	nein	ja	nein	ja [NCT00819091] (Clinicaltrials.gov, 2012a)	ja (Lewin et al., 2012)

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.

4.3.2.1.2 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-A zu hinterlegen.

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

Tabelle 4-87: Charakterisierung der eingeschlossenen Studien – Studien für indirekte Vergleiche

Studie	Studiendesign <RCT, doppelblind/ein- fach 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
Studie D1690C00005 siehe Abschnitt 4.3.1.2						
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin						
DeFronzo 1995 (Protokoll 2)	RCT, doppelblind, parallel, aktivkontrolliert, multizentrisch (25 Zentren in den USA)	Erwachsene Patienten im Alter von 40 bis 70 mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit einem Sulfonylharnstoff nicht ausreichend kontrolliert werden kann (FBG≥140 mg/dl)	Glibenclamid +Metformin (n=213) Glibenclamid +Placebo (n=209) Metformin+Placebo (n= 210)	Prärandomisation- Phase: 5Wochen (Glibenclamid) Metformin-Titration Phase: 5 Wochen Behandlungsphase: 24 Wochen	USA	Relevante primäre Endpunkte lt. Studienprotokoll: HbA1c-Veränderung im Vergleich zum Ausgangswert Relevante sekundäre Endpunkte lt. Studienprotokoll: Gewichtsveränderung im Vergleich zum Ausgangswert Blutdruckveränderungen im Vergleich zum Ausgangswert Anzahl Hypoglykämien Unerwünschte Ereignisse
Goldstein 2003	RCT, doppelblind, parallel, aktivkontrolliert, multizentrisch (108 Zentren in den USA)	Erwachsene Patienten mit Diabetes mellitus Typ 2, deren Blutzuckerwert alleine mit einem Sulfonylharnstoff	Glipizid+Metformin (n=87) Glipizid+Placebo (n=84) Metformin+Placebo (n=76)	Glipizid (einfachblinde)- Lead-in Phase: 2 Wochen Behandlung: 18 Wochen	USA Zeitraum: n.b.	Primärer Endpunkt: HbA1c-Änderung in Woche 18 im Vergleich zum Ausgangswert

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

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
		nicht ausreichend kontrolliert werden kann (HbA1c \geq 7,5% und \leq 12%)				Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtveränderung in Woche 18 im Vergleich zum Ausgangswert Anzahl Hypoglykämien Unerwünschte Ereignisse
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	RCT, doppelblind, parallel, aktivkontrolliert, multizentrisch, international	Patienten im Alter von 18 bis 75 mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit Glimepirid oder Glimepirid+Metformin nicht ausreichend kontrolliert werden kann (HbA1c \geq 7,5% und \leq 10,5%)	Stratum 1: Glimepirid+Placebo (n=106) Glimepirid+Sitagliptin (n=106) Stratum 2: Glimepirid+Metformin+Placebo (n=113) Glimepirid+Metformin+Sitagliptin (n=116)	Umstellungsphase (wenn erforderlich): bis zu 14 Wochen <i>Placebo-run-in</i> Phase: 2 Wochen Behandlungsphase: 24 Wochen	Weltweit; 2005-2007	Primärer Endpunkt: HbA1c-Änderung in Woche 24 im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtveränderung in Woche 24 im Vergleich zum Ausgangswert Anzahl Hypoglykämien Unerwünschte Ereignisse

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

Studie	Studiendesign <RCT, doppelblind/ein- fach verblindet/offen, paralle/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
Chacra et al. 2009	RCT, doppelblind, parallel, aktivkontrolliert, multizentrisch, international	Erwachsene Patienten im Alter von 18 bis 77 mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit einer submaximalen Sulfonylharnstoff- Dosis nicht ausreichend kontrolliert werden kann (HbA1c ≥7,5% und ≤10,5%)	Glibenclamid +Saxagliptin 2.5 mg/Tag (n=248) Glibenclamid +Saxagliptin 5 mg/Tag (n=253) Glibenclamid +Placebo (n=267)	Placebo-lead-in Phase: 4 Wochen Behandlungsphase: 24 Wochen	Weltweit; 2006- 2008	Primärer Endpunkt: HbA1c-Änderung in Woche 24 im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtveränderung in Woche 24 im Vergleich zum Ausgangswert Anzahl Hypoglykämien Unerwünschte Ereignisse
Garber et al. 2008	RCT, doppelblind, parallel, aktivkontrolliert, multizentrisch, international	Patienten im Alter von 18 bis 80 mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit Sulfonylharnstoff- Monotherapie nicht ausreichend kontrolliert werden kann (HbA1c 7,5% bis 11,0%)	Glimepirid+Vildagli ptin 50 mg/Tag (n=132) Glimepirid+Vildagli ptin 100 mg/Tag (n=132) Glimepirid+Placebo (n=144)	Umstellungsphase: 4 Wochen Behandlungsphase: 24 Wochen	Weltweit; 2004- 2005	Primärer Endpunkt: HbA1c-Änderung in Woche 24 im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtveränderung in Woche 24 im Vergleich zum Ausgangswert Anzahl bestätigter Hypoglykämien Unerwünschte Ereignisse

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

Studie	Studiendesign <RCT, doppelblind/ein- fach 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
Pratley et al. 2009	RCT, doppelblind, parallel, aktivkontrolliert, multizentrisch, international	Patienten im Alter von 18 bis 80 mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit Sulfonylharnstoff- Monotherapie nicht ausreichend kontrolliert werden kann (HbA1c 7,0% bis 10,0%)	Glibenclamid+Alo- gliptin 12,5mg (n=203) Glibenclamid+Alo- gliptin 25mg (n=198) Glibenclamid+ Placebo (n=99)	<i>Run-in Phase:</i> 4 Wochen <i>Behandlungsphase:</i> 26 Wochen	Weltweit; 2006- 2007	Primärer Endpunkt: HbA1c-Änderung in Woche 24 im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Gewichtveränderung in Woche 24 im Vergleich zum Ausgangswert Anzahl bestätigter Hypoglykämien Unerwünschte Ereignisse
Lewin et al. 2012	RCT, doppelblind, parallel, aktivkontrolliert, multizentrisch, international	Patienten im Alter von 18 bis 80 mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit Sulfonylharnstoff- Monotherapie nicht ausreichend kontrolliert werden kann (HbA1c 7,0% bis 10,0%)	Sulfonylharnstoff+ Linagliptin (n=161) Sulfonylharnstoff+ Placebo (n=84)	<i>Wash-out Phase:</i> 4 Wochen (Absetzen anderer oraler Antidiabetika) <i>Placebo-run-in Phase:</i> 2 Wochen <i>Behandlungsphase:</i> 18 Wochen	Weltweit; 2008- 2010	Primärer Endpunkt: HbA1c-Änderung in Woche 18 im Vergleich zum Ausgangswert Patientenrelevante sekundäre Ziele sowie weitere Endpunkte: Anzahl Hypoglykämien Unerwünschte Ereignisse

Tabelle 4-88: Charakterisierung der Interventionen – Studien für indirekte Vergleiche

Studie	Sulfonylharnstoff + Metformin bzw. DPP-Inhibitor	Sulfonylharnstoff + Placebo	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
Studie D1690C00005 siehe Abschnitt 4.3.1.2			
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin			
DeFronzo 1995 (Protokoll 2)*	Glibenclamid 20 mg/Tag + Metformin 2500 mg/Tag	Glibenclamid 20 mg/Tag + Placebo	Vorrandomisation-Phase: 5 Wochen Glibenclamid (1. Woche: 2 mal 5 mg/Tag; danach 2 mal 10 mg/Tag) Metformin-Titration Phase: 5 Wochen (Metformin Auftitration auf max 2500 mg/Tag, aufgeteilt auf 3 Tagesdosen) Behandlungsphase: 24 Wochen Glibenclamid 20 mg/Tag und Metformin 2500 mg/Tag (Dosisreduktion möglich bei Auftreten von UAW)
Goldstein 2003*	Glipizid max. 20 mg/Tag. + Metformin max. 2000 mg/Tag	Glipizid max. 30 mg/Tag + Placebo	Vorbehandlung :Sulfonylharnstoff (mind. mit der Hälfte der Höchstdosis) Glipizid 15 mg zweimal täglich (einfachblind)-Lead-in Phase: 2 Wochen Behandlung: 18 Wochen Entweder Glipizid 30mg, Metformin 500 mg oder Glipizid 5 mg +Metformin 500 mg Metformin und Glipizid+Metformin –Dosierung konnten gesteigert werden, um Blutzuckerkontrolle zu erreichen. Maximaldosis: Glipizid 30mg, Metformin 2000 mg , bzw. Metformin+Glipizid 2000 mg+30 mg Dosisreduktion war erlaubt, falls der Blutzuckerwert mehr als einmal unter 60 mg/dL war und Hypoglykämien auftraten
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor			
Hermansen 2007*	Glimepirid 4 bis 8 mg/Tag + Sitagliptin 100 mg/Tag	Glimepirid 4 bis 8 mg/Tag + Placebo	Vorbehandlung: Patienten im für den indirekten Vergleich relevanten Stratum 1 wurden auf eine stabile Dosis von Glimepirid eingestellt (Titrationphase bis zu 4 Wochen; Stabilisierungsphase von bis zu 10 Wochen) Danach erfolgte eine 2-wöchige Placebo-run-in Phase. Behandlung (Stratum 1): 24 Wochen entweder Glimepirid oder

Studie	Sulfonylharnstoff + Metformin bzw. DPP-Inhibitor	Sulfonylharnstoff + Placebo	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
			<p>Glimepirid+Sitagliptin. Eine Dosisveränderung der Hintergrundtherapie (Glimepirid) war nicht vorgesehen; <i>Rescue</i>-Therapie mit Pioglitazon war unter vordefinierten Bedingungen möglich.</p>
Chacra 2009*	<p>Glibenclamid 7,5 mg/Tag + Saxagliptin 5 mg/Tag</p>	<p>Glibenclamid 10 mg/Tag + Placebo</p>	<p>Vorbehandlung: Patienten wurden in einer 4-wöchigen <i>Placebo-lead-in</i> Phase auf Glibenclamid 7,5 mg/Tag eingestellt. Behandlung: Alle Patienten erhielten weiterhin <i>open-label</i> Glibenclamid 7,5 mg/Tag. Zusätzlich erhielten die Patienten randomisiert Saxagliptin (2,5mg oder 5mg) oder Glibenclamid (2,5mg). Eine Dosisreduktion des <i>open-label</i> Glibenclamids von 7,5 mg auf 5 mg/Tag war einmalig möglich (bei Hypoglykämien). Eine Dosiserhöhung des Glibenclamids auf max. 15 mg/Tag war unter vordefinierten Bedingungen ebenfalls möglich.</p>
Garber 2008*	<p>Glimepirid 4 mg/Tag + Vildagliptin 50 mg/Tag</p>	<p>Glimepirid 4 mg/Tag + Placebo</p>	<p>Vorbehandlung: Alle Patienten wurden auf Glimepirid 4 mg/Tag eingestellt. Nach 4 Wochen stabiler Therapie und Erfüllen der Ein- und Ausschlusskriterien erfolgte die Randomisierung in die Behandlungsgruppen (Vildagliptin 50 mg oder 100mg, oder Placebo) Behandlung: 24 Wochen Vildagliptin 50 mg oder 100mg, oder Placebo zusätzlich zu Glimepirid 4 mg/Tag. Die Glimepirid-Dosis konnte nach vordefinierten Kriterien bei Auftreten von Hypoglykämien auf 2 mg/Tag reduziert werden.</p>
Pratley 2009*	<p>Glibenclamid 10 mg/Tag + Alogliptin 25 mg/Tag</p>	<p>Glibenclamid 10 mg/Tag + Placebo</p>	<p>Vorbehandlung: Alle Patienten wurden auf Glibenclamid 10 mg/Tag eingestellt (5 mg/Tag war möglich, wenn die höhere Dosis nicht vertragen werden konnte). Behandlung: Zusätzlich zur stabilen Glibenclamid-Dosis erhielten die Patienten randomisiert Alogliptin 12.5 mg, 25 mg, oder Placebo. Bei Auftreten von Hypoglykämien konnte die Glybruid-Dosis reduziert</p>

Studie	Sulfonylharnstoff + Metformin bzw. DPP-Inhibitor	Sulfonylharnstoff + Placebo	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
			werden (um 2.5 mg / Woche). Rescue-Therapie war möglich (predefinierte Kriterien).
Lewin 2012*	Sulfonylharnstoff in stabiler Tagesdosis + Linagliptin 5 mg/Tag	Sulfonylharnstoff in stabiler Tagesdosis + Placebo	<p>Vorbehandlung: In einer <i>wash.out</i> Phase von 4 Wochen wurde der Patient nur mit Sulfonylharnstoffen behandelt, danach für 2 Wochen zusätzlich mit Placebo (<i>Run-In</i> Phase).</p> <p>Behandlung: Zusätzlich zur stabilen Hintergrundtherapie mit einem Sulfonylharnstoff wurden die Patienten randomisiert auf Linagliptin 5 mg/Tag oder Placebo (Behandlung für 18 Wochen).</p> <p>Eine <i>Rescue</i>-Therapie bei Auftreten von Hyperglykämie war möglich, eine Dosisänderung des Sulfonylharnstoffs war nicht vorgesehen.</p>

*Eventuelle für den indirekten Vergleich nicht relevante Studiengruppen (z.B. Metformin+Placebo oder nicht relevante Dosisgruppen) sind in der vorliegenden Tabelle nicht mit dargestellt.

FBG: Nüchternblutzucker

Tabelle 4-89: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche

Studie Gruppe	N	Alter (Jahre) Mittelwert [Range] / Mittelwert (SD) / (SE)	Geschlecht w / m (%)	BMI (kg/m ²) Mittelwert [Range] / Mittelwert (SD) / (SE)	Gewicht (kg) Mittelwert [Range] / Mittelwert (SD)/ (SE)
Studie D1690C00005 siehe Abschnitt 4.3.1.2					
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin					
DeFronzo 1995 (Protokoll 2)					
Glibenclamid+Metformin	213	55 (1) ^a	54,0 / 46,0	29,0 (0,3) ^a	92,1 (1,1) ^a
Glibenclamid+Placebo	209	56 (1) ^a	50,7 / 49,3	29,1 (0,3) ^a	92,6 (1,0) ^a
Goldstein 2003					
Glipizid+Metformin	87	54,6 (11,3) ^b	41,4 / 58,6	31,7 (4,9) ^b	94,0 (18,0) ^b
Glipizid+Placebo	84	57,4 (9,2) ^b	35,7 / 64,3	30,6 (4,8) ^b	89,9 (17,3) ^b
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor					
Hermansen et al. 2007 (Stratum 1)					
Glimepirid+Sitagliptin	106	54,4 (10,3) ^b	47,2 / 52,8	31,0 (6,7) ^b	85,8 (22,5) ^b
Glimepirid+Placebo	106	55,2 (10,2) ^b	45,3 / 54,7	30,7 (6,4) ^b	85,1 (22,6) ^b
Chacra et al. 2009					
Glibenclamid+Saxagliptin 5mg	253	54,9 (10,0) ^b	56,5 / 43,5	29,2 (4,6) ^b	76,2 (17,6) ^b
Glibenclamid+Placebo	267	55,1 (10,7) ^b	53,9 / 46,1	28,8 (4,7) ^b	75,6 (17,4) ^b
Garber et al. 2008					
Glimepirid+Vildagliptin 50 mg	132	58,6 (10,6) ^b	40,9 / 59,1	32,2 (4,9) ^b	91,5 (1,6) ^a
Glimepirid+Placebo	144	57,9 (10,5) ^b	41,7 / 58,3	31,0 (5,5) ^b	89,4 (1,6) ^a
Pratley et al. 2009					
Glibenclamid+Alogliptin 25 mg	198	56,5 (11,7) ^b	50,0 / 50,0	30,0 (5,3) ^b	Nicht berichtet
Glibenclamid+Placebo	99	57,1 (10,0) ^b	48,5 / 51,5	30,0 (4,8) ^b	Nicht berichtet
Lewin et al. 2012					
Sulfonylharnstoff+ Linagliptin 5 mg	161	57,2 (9,8) ^b	52,2 / 47,8	28,4 (5,0) ^b	74,5 (17,0) ^b
Sulfonylharnstoff+Placebo	84	56,2 (10,2) ^b	38,1 / 61,9	28,2 (5,1) ^b	76,1 (17,0) ^b

^aStandardfehler (SE); ^bStandardabweichung (SD)

Tabelle 4-90: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche

Studie Gruppe	N	HbA1c (%) Ausgangswert Mittelwert (SD) / (SE)	Diabetesdauer (Jahre) Median [Range] / Mittelwert (SD) / (SE)
Studie D1690C00005 siehe Abschnitt 4.3.1.2			
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin			
DeFronzo 1995 (Protokoll 2)			
Glibenclamid+Metformin	213	8,8 (0,1) ^a	7,8 (0,4) ^a
Glibenclamid+Placebo	209	8,5 (0,1) ^a	8,7 (0,4) ^a
Goldstein 2003			
Glipizid+Metformin	87	8,7 (1,2) ^b	5,9 (5,3) ^b
Glipizid+Placebo	84	8,9 (1,1) ^b	6,5 (4,4) ^b
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor			
Hermansen K et al. 2007 (Stratum 1)			
Glimepirid+Sitagliptin	106	8,42 (0,79) ^b	7,2 (5,0) ^b
Glimepirid+Placebo	106	8,43 (0,80) ^b	8,0 (6,5) ^b
Chacra AR et al. 2009			
Glibenclamid+Saxagliptin 5 mg	253	8,5 (0,9) ^b	6,8 (5,8) ^b
Glibenclamid+Placebo	267	8,4 (0,9) ^b	6,8 (5,7) ^b
Garber AJ et al. 2008			
Glimepirid+Vildagliptin 50 mg	132	8,5 (0,9) ^b	6,9 (5,2) ^b
Glimepirid+Placebo	144	8,5 (1,0) ^b	7,8 (5,8) ^b
Pratley RE et al. 2009			
Glibenclamid+Alogliptin 25 mg	198	8,09 (0,90) ^b	7,6 (6,0) ^b
Glibenclamid+Placebo	99	8,15 (0,85) ^b	7,7 (5,3) ^b
Lewin et al. 2012			
Sulfonylharnstoff+Linagliptin 5 mg	161	8,61 (0,07) ^a	Nicht berichtet
Sulfonylharnstoff+Placebo	84	8,60 (0,08) ^a	Nicht berichtet

^aStandardfehler (SE); ^bStandardabweichung (SD)

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

Studie D1690C00005: siehe Abschnitt 4.3.1.2.1

Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin

DeFronzo (1995): zwei randomisierte, placebokontrollierte, multizentrische Studienprotokolle. In beiden Studienprotokollen wird die Wirksamkeit von Metformin zur Behandlung von Patienten zwischen 40 und 70 Jahren mit Typ-2-Diabetes, deren Blutzucker ungenügend kontrolliert war (FBG ≥ 140 mg/dL), untersucht. Studienprotokoll 1 beinhaltete eine achtwöchige Vorrandomisierungsphase, in der die Patienten eine Diät hielten. Anschließend wurde die Monotherapie mit Metformin mit der Gabe von Placebo verglichen.

Für den hier durchzuführenden indirekten Vergleich war Studienprotokoll 2 relevant. Dieses beinhaltete eine fünfwöchige Vorrandomisierungsphase, in der die Patienten angeleitet wurden eine Diät zu halten und mit zunächst 10 mg Glibenclamid pro Tag (1. Woche), und danach mit 20 mg Glibenclamid pro Tag (2. bis 5. Woche) vorbehandelt wurden. Danach erfolgte (bei Erfüllen der Ein- und Ausschlusskriterien, inkl. eines Nüchternblutzuckerwertes von >140 mg/dl) die Randomisierung in die Gruppen Glibenclamid+Placebo (N=209 Patienten), Metformin+Placebo (N=210 Patienten) und Glibenclamid+Metformin (N=213 Patienten). Glibenclamid wurde in einer Dosierung von 20 mg/Tag verabreicht, während Metformin während der ersten 5 Wochen nach Randomisierung auf maximal 2500 mg/Tag auftitriert wurde (Steigerung um 500 mg Tagesdosis je Woche). Die Metformin-Dosis konnte im Laufe der Studie bei Auftreten von unerwünschten Wirkungen reduziert werden.

Goldstein (2003): randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte, 18-wöchige Studie, die die Wirksamkeit und Verträglichkeit der Kombinationstherapie Glipizid plus Metformin bei Patienten mit einem Durchschnittsalter von 56,2 Jahren mit Typ-2-Diabetes, welche durch die Behandlung mit einem Sulfonylharnstoff ungenügend kontrolliert waren ($HbA1c \geq 7,5\%$ und $\leq 12\%$, FBG <300 mg/dL), untersuchte. In einer zweiwöchigen *Run-In*-Phase wurden die Patienten zunächst auf 30 mg Glipizid pro Tag eingestellt. Bei Erfüllen der Ein- und Ausschlusskriterien erfolgte dann die Randomisierung in die Gruppen Glipizid (N=84; max. 30 mg pro Tag); Metformin (N=76; max. 2000 mg pro Tag) und Glipizid+Metformin (N=87; max. 2000 mg+20 mg pro Tag). Dosisreduzierungen waren (z.B. bei Auftreten von Hypoglykämien) erlaubt. Die Behandlung im Rahmen der Studie wurde abgebrochen, wenn der Blutzuckerspiegel (durchschnittlicher täglicher Blutglukosespiegel) vordefinierte Werte überschritt (>280 mg/dl bei Woche 5; >240 mg/dl während Wochen 5 bis 12; und >200 mg/dl ab Woche 12).

Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor

Hermansen K et al. (2007): Es handelt sich um eine randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte, 24-wöchige Studie, welche die Wirksamkeit und Sicherheit von Sitagliptin als *Add-on* Therapie zu Glimепirid (alleine oder in Kombination mit Metformin) untersucht hat. Eingeschlossen wurden Patienten im Alter von 18 bis 75 mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit Glimепirid oder Glimепirid+Metformin nicht ausreichend kontrolliert werden konnte ($HbA1c \geq 7,5\%$ und $\leq 10,5\%$). Nach einer Einstellungsphase auf die Hintergrundmedikation (Glimепirid Monotherapie 4-8 mg/Tag oder Glimепirid 4-8 mg/Tag +Metformin 1500-3000 mg/Tag) von insgesamt bis zu 14 Wochen, erfolgte eine zweiwöchige *Placebo-run-in* Phase. Wenn danach die Ein- und Ausschlusskriterien der Studie erfüllt waren, erfolgte die Randomisierung in die Behandlungsgruppe (Sitagliptin 100 mg/Tag) oder Placebo zusätzlich zu der weiterhin stabilen Hintergrundbehandlung mit Glimепirid (mit oder ohne Metformin). Die Studiendauer nach Randomisierung betrug 24 Wochen. Die Auswertung der Studie erfolgte in zwei Strata (Vorbehandlung mit Glimепirid Monotherapie oder Glimепirid+Metformin). Zur Möglichkeit einer Dosisveränderung der Hintergrundtherapie (Glimепirid) z.B. bei Auftreten von Hypoglykämien wurde in der Publikation keine Stellung genommen; *Rescue*-Therapie mit

Pioglitazon war hingegen unter vordefinierten Bedingungen möglich. Die in dieser Studie eingesetzte Dosierung von Sitagliptin (100 mg/Tag) wurde als die relevante Dosierung für den indirekten Vergleich angesehen, da dies die für diese Indikation zugelassene Dosierung in Deutschland ist.

Chacra AR et al. (2009): Es handelt sich um eine randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte, 24-wöchige Studie, welche die Wirksamkeit und Sicherheit von Saxagliptin (2,5 mg oder 5 mg) als *Add-on* Therapie zu einer submaximalen Dosis von Glibenclamid mit einer Glibenclamid-Dosissteigerung verglichen hat. Eingeschlossen wurden Patienten im Alter von 18 bis 77 Jahren mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit einer submaximalen Sulfonylharnstoff-Dosis nicht ausreichend kontrolliert werden konnte ($\text{HbA1c} \geq 7,5\%$ und $\leq 10,5\%$). Alle Patienten wurden zunächst (wenn nicht schon zutreffend) auf eine Tagesdosis von 7,5mg Glibenclamid pro Tag umgestellt (4-wöchige *Placebo-lead-in* Phase). Wenn die Ein- und Ausschlusskriterien nach Beendigung der Lead-In-Phase erfüllt waren, wurden die Patienten im Verhältnis 1:1:1 auf die drei Arme Saxagliptin 2,5mg; Saxagliptin 5mg; oder Glibenclamid 2.5mg randomisiert (der Glibenclamid-Monotherapie-Arm entsprach somit einer Glibenclamid-Dosissteigerung von 7.5mg auf 10 mg pro Tag). Die Studiendauer nach Randomisierung betrug 24 Wochen. Eine Dosisreduktion des *open-label* Glibenclamids von 7,5 mg auf 5 mg/Tag war einmalig möglich (bei Hypoglykämien). Eine Dosiserhöhung des Glibenclamids auf max. 15 mg/Tag war unter vordefinierten Bedingungen ebenfalls möglich. Bezüglich der Tagesdosis von Saxagliptin wurde der Studienarm 5 mg/Tag als relevante Dosierung für den indirekten Vergleich angesehen, da dies die für diese Indikation zugelassene Dosierung in Deutschland ist.

Garber AJ et al. (2008): Es handelt sich um eine randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte, 24-wöchige Studie, welche die Wirksamkeit und Sicherheit von Vildagliptin (in zwei Dosierungen) als *Add-on* Therapie zu Glimepirid 4mg/Tag mit Placebo verglichen hat. Eingeschlossen wurden Patienten im Alter von 18 bis 80 mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit Sulfonylharnstoff-Monotherapie nicht ausreichend kontrolliert werden konnte (HbA1c 7,5% bis 11,0%). Alle Patienten wurden auf Glimepirid 4 mg/Tag eingestellt. Nach 4 Wochen stabiler Therapie und Erfüllen der Ein- und Ausschlusskriterien erfolgte die Randomisierung in die Behandlungsgruppen (Vildagliptin 50 mg oder 100mg, oder Placebo). Die Glimepirid-Dosis konnte nach vordefinierten Kriterien bei Auftreten von Hypoglykämien auf 2 mg/Tag reduziert werden. Bezüglich der Tagesdosis von Vildagliptin wurde der Studienarm 50 mg/Tag als relevante Dosierung für den indirekten Vergleich angesehen, da dies die für diese Indikation zugelassene Dosierung in Deutschland ist.

Pratley RE et al. (2009): Es handelt sich um eine randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte, 26-wöchige Studie, welche die Wirksamkeit und Sicherheit von Alogliptin (in zwei Dosierungen) als *Add-on* Therapie zu Glibenclamid mit Placebo verglichen hat. Nach Screening wurden für die Studienteilnahme infrage kommende Patienten auf Glibenclamid 10 mg/Tag eingestellt (5 mg/Tag war möglich, wenn 10 mg/Tag

nicht vertragen werden konnte). Gleichzeitig wurde für 4 Wochen (single-blind) Placebo verabreicht. Patienten, die nach der 4-wöchigen *Placebo-run-in* Phase die Ein- und Ausschlusskriterien erfüllten, wurden dann im Verhältnis 2:2:1 auf Alogliptin (12,5 mg/Tag); Alogliptin 25 mg/Tag; oder Placebo randomisiert. Die Randomisierung erfolgte dabei stratifiziert für den HbA1C Wert bei Woche -1 (HbA1c<8.0% vs. ≥8.0%) und für Region. Die Behandlungsphase dauerte insgesamt 26 Wochen; bei Auftreten von Hypoglykämien während dieser Zeit konnte die Glibenclamid-Dosis reduziert werden (um 2.5 mg/Woche). Rescue-Therapie bei Hyperglykämie war ebenfalls möglich (nach predefinierte Kriterien). Bezüglich der Frage, welche Tagesdosis von Alogliptin relevant für den indirekten Vergleich ist, konnte nicht auf den Zulassungsstatus in Deutschland Bezug genommen, da Alogliptin zum Zeitpunkt der Dossiererstellung noch keine Zulassung in Deutschland hatte. Es wurde daher die insgesamt wirksamere Tagesdosis (25 mg/Tag) als relevante Dosierung für den indirekten Vergleich angesehen.

Lewin AJ et al. 2012: Es handelt sich um eine randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte, 18-wöchige Studie, welche die Wirksamkeit und Sicherheit von Linagliptin 5mg/Tag als *Add-on* Therapie zu einer Sulfonylharnstoff-Therapie mit Placebo verglichen hat. Es gab bei dieser Studie keine Vorschriften hinsichtlich der Sulfonylharnstoff-Einzelsubstanz. In einer *wash-out* Phase von 4 Wochen wurden Patienten, die zusätzlichen mit anderen oralen Antidiabetika behandelt wurden, auf eine Sulfonylharnstoff-Monotherapie umgestellt; danach erfolgte für alle Patienten eine zweiwöchige *Placebo-run-in* Phase. Das jeweilige Sulfonylharnstoff-Präparat wurde für die gesamte Dauer der Studie (inklusive der 4-wöchigen *wash-out* Phase und der zweiwöchigen *Placebo-run-in* Phase in derselben konstanten Dosierung verabreicht). Bei Erfüllen der Ein- und Ausschlusskriterien erfolgte nach der *Placebo-run-in* Phase eine Randomisierung in die Gruppen Linagliptin oder Placebo zusätzlich zur Sulfonylharnstoff-Hintergrundtherapie. Die Behandlungsphase nach Randomisierung betrug 18 Wochen; eine *Rescue*-Therapie bei Auftreten von Hyperglykämie war möglich. Die in dieser Studie eingesetzte Dosierung von Linagliptin (5 mg/Tag) wurde als die relevante Dosierung für den indirekten Vergleich angesehen, da dies die für nahe verwandte Indikationen (z.B. *Add-on* zur Metformin-Monotherapie oder zur Sulfonylharnstoff-Metformin-Kombinationstherapie) zugelassene Dosierung in Deutschland ist.

Um eine möglichst hohe Vergleichbarkeit der Studien im indirekten Vergleich zu gewährleisten, wurden wenn möglich Wirksamkeits- und Verträglichkeitsdaten zum Zeitpunkt 24 Wochen (oder mit möglichst geringem Abstand zum Zeitpunkt 24 Wochen) verwendet.

4.3.2.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-91: Verzerrungspotenzial auf Studienebene – Studien für indirekte Vergleiche

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
Studie D1690C00005 siehe Abschnitt 4.3.1.2.2							
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin							
DeFronzo 1995 (Protokoll 2)	ja	ja	ja	ja	nein	nein	niedrig
Goldstein 2003	ja	ja	ja	ja	unklar*	nein	niedrig
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor							
Hermansen K et al. 2007	ja	ja	ja	ja	unklar*	nein	niedrig
Chacra AR et al. 2009	ja	ja	ja	ja	unklar*	nein	niedrig
Garber AJ et al. 2008	ja	ja	ja	ja	unklar*	nein	niedrig
Pratley RE et al. 2009	ja	ja	ja	ja	unklar*	nein	niedrig
Lewin AJ et al. 2012	ja	ja	ja	ja	unklar*	nein	niedrig

*es liegt nur die Publikation der Studie vor (kein Studienbericht)

Begründen Sie für jede Studie die abschließende Einschätzung.

Für alle zu beurteilenden Studien gilt, dass das Verzerrungspotenzial auf Studienebene als „niedrig“ eingestuft werden kann, da die Studien mit der zweckmäßigen bzw. alternativen Vergleichstherapie randomisiert durchgeführt wurden. Außerdem erfolgte die Gruppenzuteilung jeweils verdeckt, sowohl Patient als auch Behandler waren verblindet (doppelblind) und es gab keine sonstigen das Verzerrungspotenzial beeinflussenden Faktoren. Für Goldstein et al., sowie für die DPP-4-Studien war eine ergebnisgesteuerte Berichterstattung nicht abschließend zu beurteilen (auszuschließen), da keine Studienberichte bzw. Studienprotokolle vorlagen. Hinweise auf eine ergebnisgesteuerte Berichterstattung ergaben sich aber nicht.

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 4-92: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	Hypo- glykämien	Gewichts- veränderung	Un- erwünschte Ereignisse	Folge- komplikationen	Gesamt- mortalität	Systolischer und diastolischer Blutdruck
D1690C00005	ja	ja	ja	ja	nein	ja
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin						
DeFronzo 1995 (Protokoll 2)	ja	ja	ja	nein	nein	ja
Goldstein 2003	ja	ja	ja	nein	nein	nein
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen K et al. 2007	ja	ja	ja	nein	nein	nein
Chacra AR et al. 2009	ja	ja	ja	nein	nein	nein
Garber AJ et al. 2008	ja	ja	ja	nein	nein	nein
Pratley RE et al. 2009	ja	ja	ja	nein	nein	nein
Lewin AJ et al. 2012	ja	nein	ja	nein	nein	nein

Tabelle 4-93: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	Hyperosmolare Komata	Stationäre Behandlungen	Gesundheitsbezogene Lebensqualität	Therapiezufriedenheit	Ausmaß der Blutzuckersenkung (HbA1c)
D1690C00005	nein	nein	nein	ja	ja
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin					
DeFronzo 1995 (Protokoll 2)	nein	nein	nein	nein	ja
Goldstein 2003	nein	nein	nein	nein	ja
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor					
Hermansen K et al. 2007	nein	nein	nein	nein	ja
Chacra AR et al. 2009	nein	nein	nein	nein	ja
Garber AJ et al. 2008	nein	nein	nein	nein	ja
Pratley RE et al. 2009	nein	nein	nein	nein	ja
Lewin AJ et al. 2012	nein	nein	nein	nein	ja

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 4-94: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Dapagliflozin	Sulfonylharnstoff + Placebo	Sulfonylharnstoff + Metformin	Metformin+ Placebo*
1	D1690C00005	●	●		
2	DeFronzo 1995 (Protokol 2) Goldstein 2003		●	●	●
			●	●	●

*dieser Behandlungsarm wird nicht betrachtet

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

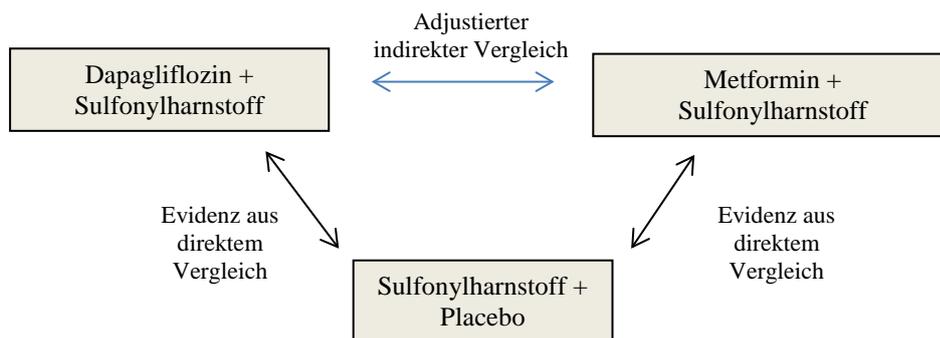
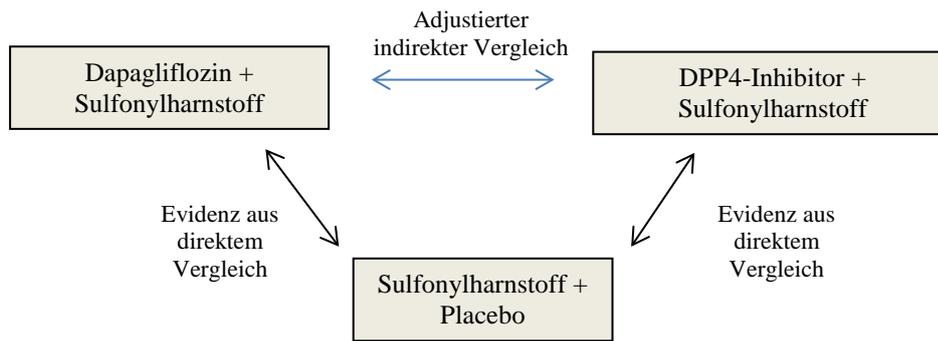


Tabelle 4-95: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Dapagliflozin	Sulfonylharnstoff + Placebo	Sulfonylharnstoff + DPP-4-Inhibitor
1	D1690C00005	●	●	
5*	Hermansen et al. 2007		●	●
	Chacra et al. 2009		●	●
	Garber et al. 2008		●	●
	Pratley et al. 2009		●	●
	Lewin et al. 2012		●	●

Für den indirekten Vergleich irrelevante Studienarme dieser Studien sind in der Tabelle nicht berücksichtigt..

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.



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 Ausmaß der Blutzuckerkontrolle (HbA1c)

Studie	Operationalisierung
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1	
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin	
DeFronzo 1995 (Protokoll 2)	anhand der Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 29
Goldstein 2003	anhand der Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 18
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor	
Hermansen et al. 2007	anhand der Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 24
Chacra et al. 2009	anhand der Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 24
Garber et al. 2008	anhand der Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 24
Pratley et al. 2009	anhand der Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 26
Lewin et al. 2012	anhand der Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 18

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-97: 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
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1						
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin						
DeFronzo 1995 (Protokoll 2)	niedrig	ja	ja	nein	nein	niedrig
Goldstein 2003	niedrig	ja	ja	unklar*	nein	niedrig
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	niedrig	ja	ja	unklar*	nein	niedrig
Chacra et al. 2009	niedrig	ja	ja	unklar*	nein	niedrig
Garber et al. 2008	niedrig	ja	ja	unklar*	nein	niedrig
Pratley et al. 2009	niedrig	ja	ja	unklar*	nein	niedrig
Lewin et al. 2012	niedrig	ja	ja	unklar*	nein	niedrig

*es liegt nur die Publikation der Studie vor (kein Studienbericht)

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für den zu bewertenden Endpunkt „Ausmaß der Blutzuckerkontrolle (HbA1c)“ für alle Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es keine sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

Eine ergebnisunabhängige Berichterstattung kann mit hoher Sicherheit nur ausgeschlossen werden, wenn sowohl Publikation wie auch der vollständige Studienbericht vorliegen. Für die RCT, die einen DPP-4-Inhibitor eingeschlossen haben, sowie für die Studie von Goldstein et al. (2003) lagen keine Studienberichte vor. Explizite Hinweise auf eine ergebnisabhängige Berichterstattung ergaben sich aber für keine der vorliegenden Studien.

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 4-98: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT für indirekte Vergleiche

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff +Dapagliflozin (B)				Komparator: Sulfonylharnstoff +Placebo (A)				Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4- Inhibitor (C)				Gruppen- unterschied % Mittelwert (SE) [95%-KI]
	N	Ausgangs- wert % Mittel- wert (SD)	End- wert % Mittel- wert (SD)	Verände- rung % Mittel- wert (SE)	N	Ausgangs- wert % Mittel- wert (SE/SD)	Endwert % Mittel- wert (SE/SD)	Veränder- ung % Mittelwert (SE/SD)	N	Ausgang swert % Mittel- wert (SE/SD)	Endwert % Mittel- wert (SE/SD)	Veränder- ung % Mittelwert (SE/SD)	
Studie D1690C00005													
24 Wochen Studien- population	150	8,07 (0,792)	7,27 (0,843)	-0,82 (0,0610)	143	8,15 (0,741)	8,00 (0,928)	-0,13 (0,0625)					-0,68 (0,0873) [-0,86;-0,51]
Zielpopulation	129	8,09 (0,800)	7,25 (0,829)	-0,86 (0,0647)	111	8,17 (0,744)	8,03 (0,951)	-0,13 (0,0698)					-0,74 (0,0952) [-0,92;-0,55]
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin													
DeFronzo 1995 (Protokoll 2) 25 Wochen					191	8,52 (0,11) ^a	8,43 (0,11) ^a	-0,08 (0,10) ^a	200	8,80 (0,11) ^a	6,82 (0,10) ^a	-1,96 (0,11) ^a	-1,88 (0,15) ^{a,*} [-2,17;-1,59] [*]
Goldstein 2003 18 Wochen					79	8,9 (1,1) ^b	8,5 (0,1) ^a	-0,4 (0,11) ^{a,*}	80	8,7 (1,2) ^b	7,4 (0,1) ^a	-1,3 (0,11) ^{a,*}	-1,06 (0,15) ^a [-1,35;-0,77] [*]
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor													
Hermansen et al. 2007					106	8,42 (0,79) ^b		0,27 [0,09;0,45]	106	8,43 (0,80) ^b		-0,30 [-0,48;- 0,12]	-0,57 [-0,82;-0,32]

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff +Dapagliflozin (B)		Komparator: Sulfonylharnstoff +Placebo (A)			Zweckmäßige Vergleichstherapie:			Gruppen- unterschied % Mittelwert (SE) [95%-KI]
						Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4- Inhibitor (C)			
Chacra et al. 2009			253	8,5 (0,9) ^b	0,08 (0,03) ^a	267	8,4 (0,9) ^b	-0,64 (0,03) ^a	-0,72 (0,0413) ^{a,*}
Garber et al. 2008			144	8,5 (1,0) ^b	0,07 (0,09) ^a	132	8,5 (0,9) ^b	-0,58 (0,10) ^a	-0,64 (0,13) ^a
Pratley et al. 2009			99	8,15 (0,85) ^b	0,01	198	8,09 (0,90) ^b	-0,53	-0,53 [-0,73;-0,33]
Lewin et al. 2012			84	8,60 (0,08) ^a	-0,07 (0,10) ^a	161	8,61 (0,07) ^a	-0,54 (0,07) ^a	-0,47 (0,12) ^a

^aStandardfehler (SE); ^bStandardabweichung (SD); KI: Konfidenzintervall

*eigene Berechnung / Schätzung

Ergebnisse der Studien mit der zweckmäßigen Vergleichstherapie

Für den Endpunkt „Ausmaß der Blutzuckerkontrolle“ stehen beide Studien (DeFronzo und Goldstein) für den Vergleich von Sulfonylharnstoff+Placebo versus Sulfonylharnstoff+Metformin zur Verfügung. Bevor eine Meta-Analyse der zwei Studien durchgeführt werden konnte, musste zunächst überprüft werden, ob die Zusammenfassung der betrachteten Studien überhaupt sinnvoll ist, da die Studien bezüglich der Fragestellung vergleichbar sein mussten. Beide Studien verglichen Sulfonylharnstoff+Placebo versus Sulfonylharnstoff+Metformin und in jeder einzelnen Studie war die Behandlung von Sulfonylharnstoff+Metformin der Behandlung von Sulfonylharnstoff+Placebo in Bezug auf die Senkung des HbA1c nachweislich überlegen. Die Studien unterschieden sich jedoch nicht nur in Bezug auf die Zeitpunkte der HbA1c-Auswertung (18 Wochen vs. 25 Wochen). Untersuchte man die Heterogenität der Studien bezüglich der Veränderung des HbA1c, so zeigte sich ein hohes Ausmaß an Heterogenität ($I^2=93\%$; $p<0,0001$). Von der Berechnung eines Gesamtschätzers wurde daher Abstand genommen, stattdessen wurde der indirekte Vergleich für beide Studien einzeln durchgeführt.

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 4-99: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) - indirekte Vergleiche

	Mittelwertdifferenz des indirekten Vergleiches Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin HbA1c% [95%-KI]
D1690C00005 (Studienpopulation) und DeFronzo	1,20 [0,86;1,54]
D1690C00005 (Zielpopulation) und DeFronzo	1,14 [0,79;1,49]
D1690C00005 (Studienpopulation) und Goldstein	0,38 [0,04;0,72]
D1690C00005 (Zielpopulation) und Goldstein	0,32 [-0,03;0,67]

KI: Konfidenzintervall

Der indirekte Vergleich unter Berücksichtigung von DeFronzo et al. zeigte einen signifikanten Unterschied zugunsten von Sulfonylharnstoff+Metformin (MWD [95%-KI]: 1.20 [0.86;1.54] bzw. 1.14 [0.79;1.49]). Die Vergleichbarkeit der Studie von DeFronzo und der Studie D1690C00005 war allerdings in Bezug auf verschiedene Aspekte eingeschränkt. So war der

Ausgangswert für HbA1c in der Studie von DeFronzo höher als in der Studie D1690C00005 (siehe Tabelle 4-90); zudem hatten in der DeFronzo Studie die Patienten in der Sulfonylharnstoff+Metformin Gruppe einen höheren HbA1c-Ausgangswert (8,80%) als die Patienten in der Sulfonylharnstoff+Placebo Gruppe (8,52%). Dieses kann zu einem scheinbar besonders ausgeprägten Behandlungseffekt durch Metformin *Add-on* beigetragen haben.

Der indirekte Vergleich unter Berücksichtigung von Goldstein et al. zeigte bei Berücksichtigung der Dapagliflozin-Zielpopulation keinen signifikanten Unterschied (MWD [95%-KI]: 0.32 [-0.03;0.67]), sondern nur bei Berücksichtigung der Studienpopulation (MWD [95%-KI]: 0.38 [0.04;0.72]). Patienten in der Studie von Goldstein und in der Studie D1690C00005 wurden mit einem Sulfonylharnstoff vorbehandelt. Die Studiendauer war unterschiedlich (18 / 24 Wochen), die HbA1c-Ausgangswerte waren bei der Publikation von Goldstein (8,7% / 8,9%) höher als in Studie D1690C00005. Im Vergleich zur Dapagliflozin-Studienpopulation kann dieses zu einem scheinbar besonders ausgeprägten Behandlungseffekt durch Metformin *Add-on* beigetragen haben.

Ergebnisse der Studien mit der alternativen Vergleichstherapie

Für den Endpunkt „Ausmaß der Blutzuckerkontrolle“ stehen alle Studien des Vergleiches Sulfonylharnstoff+Placebo versus Sulfonylharnstoff+DPP-4-Inhibitor zur Verfügung. Alle Studien zeigten einen statistisch signifikanten Vorteil des Sulfonylharnstoff+DPP-4-Inhibitor Arms gegenüber der Kontrollgruppe. Die Größenordnungen dieses Effektes waren in ähnlichem Bereich. Die Meta-Analyse wies auf eine moderate Heterogenität der Studienergebnisse hin ($I^2=39\%$; $p=0.16$), welche vor allem durch den relativ starken Effekt (mit engen Konfidenzintervallen) der Studie von Chacra 2009 begründet war.

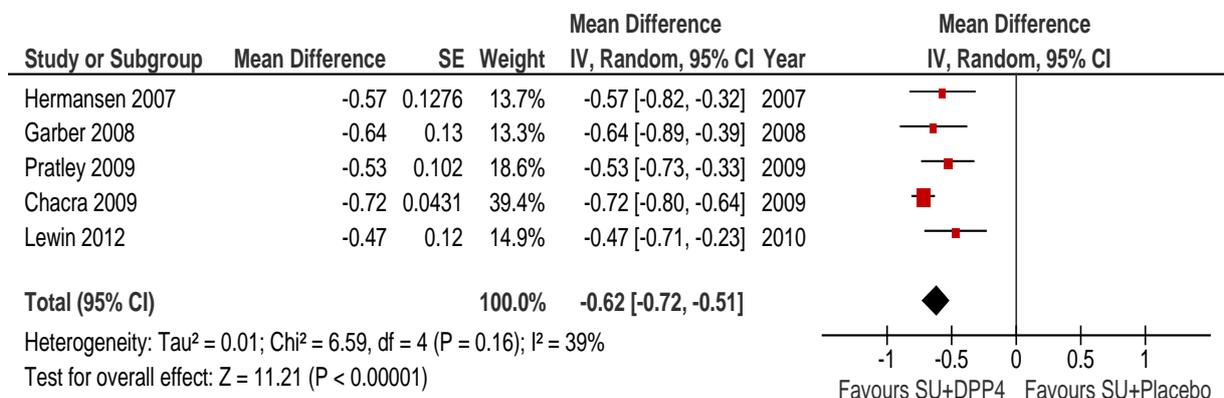


Abbildung 12: Meta-Analyse der eingeschlossenen Studie zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Ausmaß der Blutzuckerkontrolle“

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 4-100: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) - indirekte Vergleiche

	Mittelwertsdifferenz des indirekten Vergleiches Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor HbA1c% [95%-KI]
D1690C00005 (Studienpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	-0,06 [-0,26;0,14]
D1690C00005 (Zielpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	-0,12 [-0,33;0,09]
D1690C00005 (Studienpopulation) und Hermansen 2007 (Sitagliptin)	-0,11 [-0,41;0,19]
D1690C00005 (Zielpopulation) und Hermansen 2007 (Sitagliptin)	-0,17 [-0,48;0,14]

KI: Konfidenzintervall

Der indirekte Vergleich ergab in Bezug auf den Endpunkt „Ausmaß der Blutzuckerkontrolle (HbA1c)“ keine statistisch signifikanten Unterschiede zwischen der zu bewertenden Therapie und der alternativen Vergleichstherapie (DPP-4-Inhibitoren als Gruppe bzw. Sitagliptin).

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 4-101: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Dapagliflozin	Sulfonylharnstoff +Placebo	Sulfonylharnstoff + Metformin	Metformin+ Placebo*
1	D1690C00005	•	•		
2	DeFronzo 1995 (Protokoll 2)		•	•	•
	Goldstein 2003		•	•	•

*dieser Behandlungsarm wurde nicht betrachtet

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

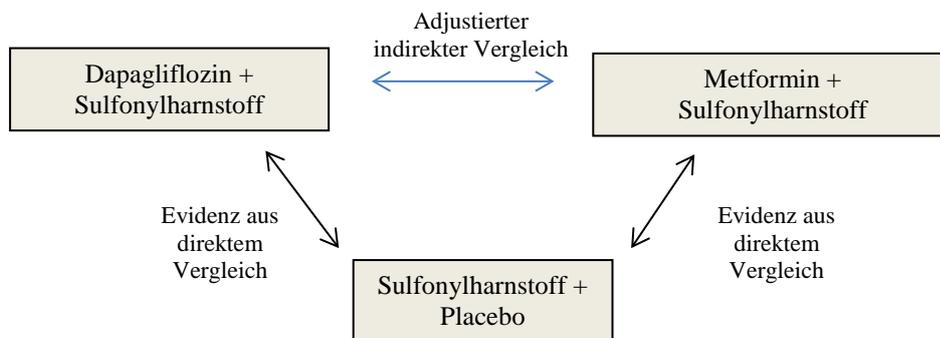
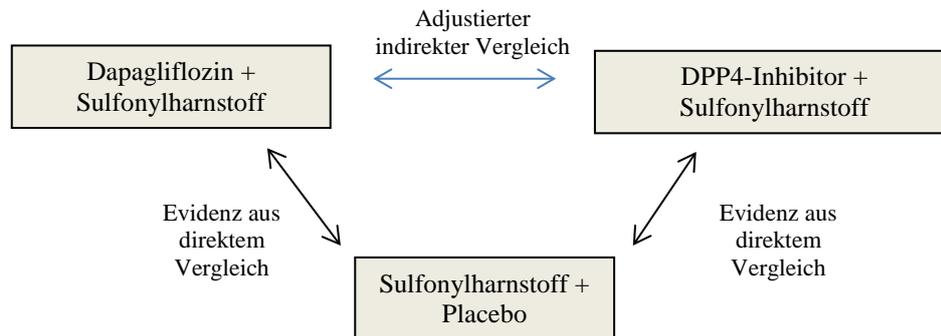


Tabelle 4-102: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Dapagliflozin	Sulfonylharnstoff + Placebo	Sulfonylharnstoff + DPP-4-Inhibitor
1	D1690C00005	•	•	
5*	Hermansen et al. 2007		•	•
	Chacra et al. 2009		•	•
	Garber et al. 2008		•	•
	Pratley et al. 2009		•	•
	Lewin et al. 2012		•	•

Für den indirekten Vergleich irrelevante Studienarme dieser Studien sind in der Tabelle nicht berücksichtigt..

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.



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

Tabelle 4-103: Operationalisierung von Hypoglykämien

Studie	Operationalisierung
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1.1	
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin	
DeFronzo 1995 (Protokoll 2)	<u>Gesamt:</u> Durch den Patienten berichtete, subjektive Hypoglykämie-Symptome innerhalb der 29-wöchigen Behandlungsphase <u>Bestätigt:</u> dokumentierter, biochemischer Nachweis von Hypoglykämien
Goldstein 2003	<u>Bestätigt:</u> Blutzuckerwert ≤ 50 mg/dL (Patienten führten ein Tagebuch und führten mittels Fingerpunktion eine Blutzuckermessung durch, falls Symptome einer Hypoglykämie auftraten) innerhalb der 18-wöchigen Behandlungsphase <u>Schwer:</u> Benötigung medizinischer Hilfe
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor	
Hermansen 2007	<u>Gesamt:</u> Durch den Patienten berichtete Ereignisse von Hypoglykämien <u>Schwer:</u> Vom Studienarzt als schwer eingeschätzt oder Ereignisse, die medizinische Hilfe erforderten
Chacra 2009	<u>Gesamt:</u> Ereignisse vereinbar mit den Zeichen / Symptomen einer Hypoglykämie (unabhängig davon, ob Blutglukosespiegel dokumentiert wurden oder nicht) <u>Bestätigt:</u> Hypoglykämie-Symptome und Blutzuckerwert ≤ 50 mg/dL (Blutzuckermessung aus der Fingerkuppe) <u>Schwer:</u> Vom Studienarzt als schwer eingeschätzt
Garber 2008	<u>Bestätigt:</u> Hypoglykämie-Symptome und Blutzuckerwert (aus Selbstmessung) $\leq 3,1$ mmol/l
Pratley 2009	<u>Bestätigt:</u> Hypoglykämie-Symptome und Blutzuckerwert ≤ 50 mg/dL (Blutzuckermessung aus der Fingerkuppe) <u>Schwer:</u> Ereignisse, die medizinische Hilfe erforderten mit Blutzuckerwert ≤ 60 mg/dL
Lewin 2012	<u>Gesamt:</u> Durch den Studienarzt als Hypoglykämie berichtet <u>Schwer:</u> Ereignisse, die medizinische Hilfe Dritter erforderten (zur Verabreichung von Glukose oder Medikation oder zur Reanimation)

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-104: 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
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1.1						
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin						
DeFronzo 1995 (Protokoll 2)	niedrig	ja	ja	nein	nein	niedrig
Goldstein 2003	niedrig	ja	ja	unklar*	nein	niedrig
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen 2007	niedrig	ja	ja	unklar*	nein	niedrig
Chacra 2009	niedrig	ja	ja	unklar*	nein	niedrig
Garber 2008	niedrig	ja	ja	unklar*	nein	niedrig
Pratley 2009	niedrig	ja	ja	unklar*	nein	niedrig
Lewin 2012	niedrig	ja	ja	unklar*	nein	niedrig

*es lag nur die Publikation der Studie vor (kein Studienbericht)

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wurde das Verzerrungspotenzial für den zu bewertenden Endpunkt „Hypoglykämien“ für alle Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es keine sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

Da für Goldstein et al. 2003, sowie für die RCT zu DPP-4-Inhibitoren nur eine Publikation der Studie und kein Studienbericht vorlag, konnte die ergebnisunabhängige Berichterstattung nicht mit Sicherheit überprüft werden. Explizite Hinweise auf eine ergebnisabhängige Berichterstattung fanden sich aber nicht.

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 4-105: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff +Dapagliflozin (B)		Komparator: Sulfonylharnstoff +Placebo (A)		Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4-Inhibitor (C)	Odds Ratio [95%-KI]	
	Sulfonylharnstoff+Dapagliflozin 10 mg		Sulfonylharnstoff+Placebo				
Studie D1690C00005	N	n (%)	N	n (%)	N	n (%)	
Gesamt (Studienpopulation)	151	12 (7,9)	146	7 (4,8)		1,71 [0,66;4,48]	
schwer	151	0 (0,0)	146	0 (0,0)		Nicht berechenbar	
leicht	151	10 (6,6)	146	3 (2,1)		3,38 [0,91;12,54]	
andere	151	3 (2,0)	146	5 (3,4)		0,57 [0,13;2,44]	
Gesamt (Zielpopulation)	129	12 (9,3)	114	5 (4,4)		2,24 [0,76;6,55]	
schwer	129	0	114	0		Nicht berechenbar	
leicht	129	10 (7,8)	114	2 (1,8)		4,71 [1,01;21,95]	
andere	129	3 (2,3)	114	4 (3,5)		0,65 [0,14;2,99]	
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin							
DeFronzo 1995 (Protokoll 2)							
Gesamt			209	7 (3,3)	213	38 (17,8)	6,26 [2,73;14,39]*
bestätigte			209	0 (0)	213	0 (0)	Nicht berechenbar

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff +Dapagliflozin (B)	Komparator: Sulfonylharnstoff +Placebo (A)	Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4-Inhibitor (C)	Odds Ratio [95%-KI]
Goldstein 2003				
bestätigte		79 0 (0,0)	80 11(12,6)	26,33 [1,52;455,15]*,a
schwer		79 0 (0,0)	80 0 (0,0)	Nicht berechenbar
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor				
Hermansen et al. 2007				
Gesamt		106 3 (2,8)	106 8 (7,5)	2,80 [0,72;10,87]*
schwer		106 0	106 0	Nicht berechenbar
Chacra et al. 2009				
Gesamt		267 27 (10,1)	253 37 (14,6)	1,52 [0,84;2,42]*
bestätigt		267 2 (0,7)	253 3 (0,8)	1,51 [0,25;9,08]*
schwer		267 0	253 0	Nicht berechenbar
Garber et al. 2008				
bestätigt		176 1 (0,6)	170 2 (1,2)	2,08 [0,19;23,19]*
Pratley et al. 2009				
Gesamt		99 11 (11,1)	198 19 (9,6)	0,85 [0,39;1,86]*
Bestätigt		99 6 (6,1)	198 8 (4,0)	0,65 [0,22;1,94]*
Schwer		99 1 (1,0)	198 0	1,51 [0,06;37,44]*,a

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff +Dapagliflozin (B)	Komparator: Sulfonylharnstoff +Placebo (A)	Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4-Inhibitor (C)	Odds Ratio [95%-KI]
Lewin et al. 2012				
Gesamt		84 4 (4,8)	161 9 (5,6)	1,18 [0,35;3,97]
Schwer		84 0	161 0	Nicht berechenbar

*eigene Berechnung

^a Odds-Ratio zur Veranschaulichung des Gruppenunterschieds berechnet nach Woolf-Haldane Korrektur (Addition von 0.5 zu jeder Zelle der Vierfeldertafel). Odds-Ratio wurde nicht zur Durchführung eines indirekten Vergleiches verwendet.

KI: Konfidenzintervall

Ergebnisse der Studien mit der zweckmäßigen Vergleichstherapie

Signifikante Unterschiede bei den Hypoglykämien gab es in der gesamten Anzahl der Hypoglykämien in der Studie von DeFronzo. Hier zeigte die Sulfonylharnstoff+Metformin-Gruppe ein höheres Risiko für Hypoglykämien als in der Sulfonylharnstoff+Placebo-Gruppe (OR [95%-KI]: 6,26 [2.73;14.39]). In der Studie von Goldstein zeigte sich ebenfalls ein höheres Risiko von (in diesem Fall bestätigten) Hypoglykämien in der Glipizid+Metformin Gruppe im Vergleich zur Glipizid+Placebo-Gruppe (geschätztes OR nach Addition von 0.5 zu jeder Zelle aufgrund von 0 Ereignissen in der Glipizid+Placebo-Gruppe: 26,33 [1.52;455.15]).

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.*

Gesamtrate der Patienten mit mindestens einer Hypoglykämie:

Diese Gesamtrate wurde in Studie D1690C00005 und in der Publikation DeFronzo beschrieben. In der Studie von Goldstein wurde diese Rate nicht berichtet, eine Meta-Analyse kam hier daher nicht infrage. Im indirekten Vergleich zeigte sich für Sulfonylharnstoff +Dapagliflozin ein Vorteil gegenüber Sulfonylharnstoff+Metformin bei Berücksichtigung der Studienpopulation, der statistisch signifikant war. Keine statistische Signifikanz fand sich hingegen für den indirekten Vergleich unter Berücksichtigung der Zielpopulation aus der D1690C00005-Studie (Tabelle 4-106), vermutlich aufgrund der geringeren statistischen Power in dieser Analyse.

Tabelle 4-106: Ergebnisse für Hypoglykämien aus indirekte Vergleiche

	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff +Dapagliflozin versus Sulfonylharnstoff+Metformin [95%-KI]
D1690C00005 (Studienpopulation) und DeFronzo	0,27 [0,08;0,97]
D1690C00005 (Zielpopulation) und DeFronzo	0,36 [0,09;1,39]

KI: Konfidenzintervall

Gesamtrate der Patienten mit mindestens einer leichten Hypoglykämie:

Kein Vergleich möglich, da nur die Studie D1690C00005 dies berichtete.

Gesamtrate der Patienten mit mindestens einer schwerwiegenden Hypoglykämie:

Kein Vergleich möglich. Die Gesamtrate wurde in D1690C00005 und in der Studie von Goldstein berichtet, aber da keine Ereignisse aufgetreten sind ließen sich keine Effektschätzer für einen indirekten Vergleich berechnen.

Ergebnisse der Studien mit der alternativen Vergleichstherapie

Signifikante Unterschiede bei den Hypoglykämien zwischen den betrachteten Studienarmen fanden sich in den Studien zu DPP-4 Inhibitoren nicht.

Gesamtrate der Patienten mit mindestens einer Hypoglykämie:

Diese Gesamtrate wurde in Studie D1690C00005 und allen fünf in den indirekten Vergleich eingeschlossenen Studien berichtet. Es zeigten sich keine Hinweise auf bedeutsame Heterogenität der Ergebnisse ($I^2=0\%$; $p=0.59$).

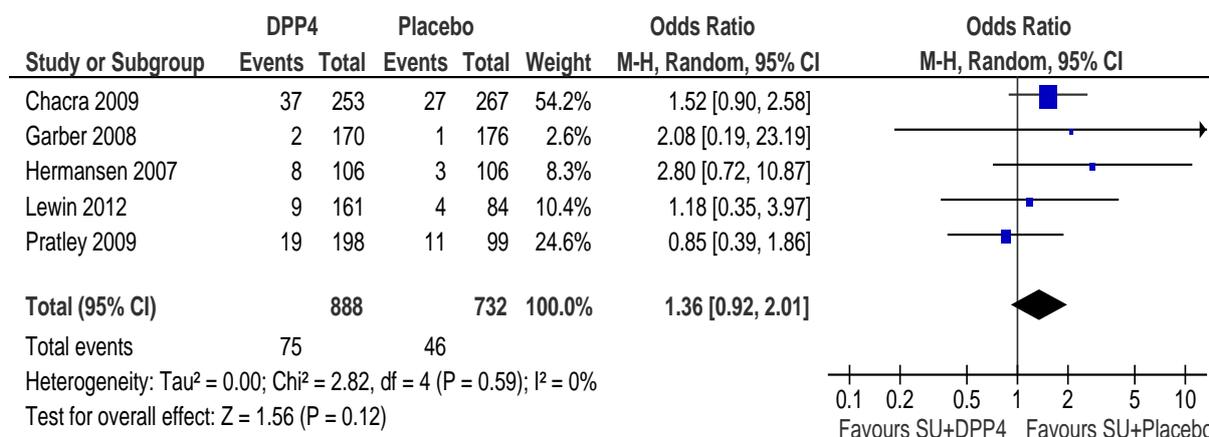


Abbildung 13: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate der Patienten mit mindestens einer Hypoglykämie“

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 4-107: Ergebnisse für Gesamtrate der Patienten mit mindestens einer Hypoglykämie - indirekte Vergleiche

	Odds ratio des indirekten Vergleiches Sulfonylharnstoff +Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor [95%-KI]
D1690C00005 (Studienpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	1,26 [0,45;3,54]
D1690C00005 (Zielpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	1,65 [0,52;5,18]
D1690C00005 (Studienpopulation) und Hermansen 2007 (Sitagliptin)	0,61 [0,12;3,22]
D1690C00005 (Zielpopulation) und Hermansen 2007 (Sitagliptin)	0,80 [0,14;4,52]

KI: Konfidenzintervall

Der indirekte Vergleich ergab in Bezug auf den Endpunkt "Gesamtrate der Patienten mit mindestens einer Hypoglykämie" keine statistisch signifikanten Unterschiede zwischen der zu bewertenden Therapie und der alternative Vergleichstherapie (DPP-4-Inhibitoren als Gruppe bzw. Sitagliptin).

Gesamtrate der Patienten mit mindestens einer leichten Hypoglykämie:

Kein Vergleich möglich, da nur die Studie D1690C00005 dies berichtete.

Gesamtrate der Patienten mit mindestens einer schwerwiegenden Hypoglykämie:

Kein Vergleich möglich. Die Gesamtrate wurde in D1690C00005 und in den Studien von Hermansen et al. 2007; Chacra et al. 2009; Pratley et al. 2009; und Lewin et al. 2012 berichtet. Lediglich in der Studie von Pratley et al. ist dabei eine schwerwiegende Hypoglykämie berichtet worden, diese trat in der Gruppe Sulfonylharnstoff+Placebo auf. Einen Effektschätzer für einen indirekten Vergleich ließ sich aus diesen Informationen nicht berechnen.

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

Im Folgenden wird zunächst die in den vorherigen Abschnitten 4.3.2.1.3.1 und 4.3.2.1.3.1.1 der Endpunkte „Ausmaß der Blutzuckerkontrolle“ und „Hypoglykämien“ gemeinsam

betrachtet, da das Auftreten von Hypoglykämien nur unter Berücksichtigung der Blutzuckerkontrolle sinnvoll beurteilt werden kann (siehe Abschnitt 4.3.1.3).

Tabelle 4-108: Ergebnisse für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel – indirekter Vergleich

Zeitpunkte	Studie(n)	Blutzuckerkontrolle	Hypoglykämien	Gesamt- Beurteilung
Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin				
25 bzw. 29 Wochen (DeFronzo) und 24 Wochen (D1690C00005)	D1690C00005; DeFronzo et al. 1995	Nachteil	Kein Unterschied	kein Zusatznutzen
Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor				
24 Wochen (D1690C00005; Hermansen et al.; Chacra et al. 2009); 26 Wochen (Pratley et al.); 18 Wochen (Lewin et al. 2012)	D1690C00005; Hermansen et al. 2007; Chacra et al. 2009; Garber et al. 2008; Pratley et al. 2009; Lewin et al. 2012	Kein Unterschied	Kein Unterschied	Kein Zusatznutzen

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 FDA, auf die sich auch das IQWiG im Abschlussbericht Blutglukoseseibstmessung bezieht. 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.

Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin:

Hinsichtlich der Änderung des HbA1c-Wertes ergab sich im indirekten Vergleich von DeFronzo und D1690C00005 ein Nachteil in der Blutzuckersenkung für Sulfonylharnstoff+Dapagliflozin. Bezüglich der Gesamtrate an Hypoglykämien ergab sich für die Zielpopulation kein Unterschied zwischen den Behandlungsoptionen, für die Studienpopulation zeigte sich ein Vorteil für Sulfonylharnstoff+Dapagliflozin. Insgesamt ergab sich aus dieser Datenlage kein Vorteil für eine der beiden Behandlungsoptionen.

Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor:

Der indirekte Vergleich ergab für keinen der relevanten Endpunkte einen statistisch signifikanten Unterschied zwischen den betrachteten Gruppen. In der Gesamtbeurteilung ergab sich somit aus dieser Datenlage ebenfalls kein Vorteil für eine der beiden 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 4-109: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Dapagliflozin	Sulfonylharnstoff + Placebo	Sulfonylharnstoff + Metformin	Metformin+ Placebo*
1	D1690C00005	•	•		
2	DeFronzo 1995 (Protokoll 2) Goldstein 2003		•	•	•

*dieser Behandlungsarm wurde nicht betrachtet

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

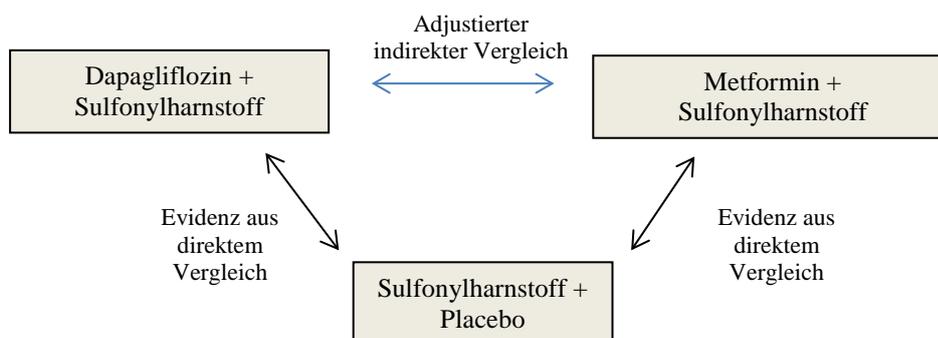
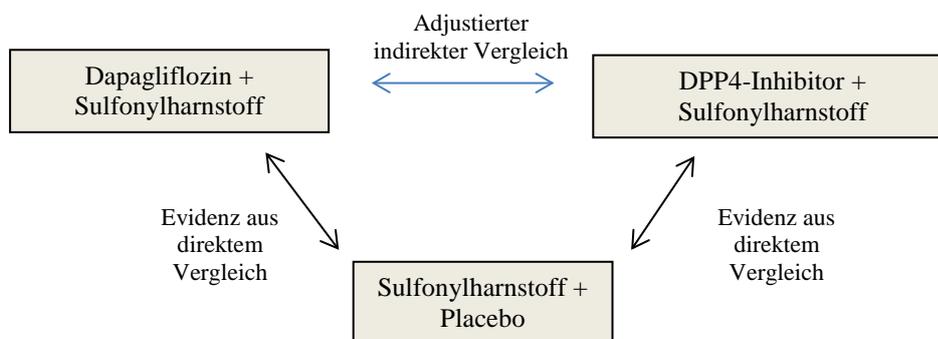


Tabelle 4-110: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Dapagliflozin	Sulfonylharnstoff + Placebo	Sulfonylharnstoff + DPP-4-Inhibitor
1	D1690C00005	•	•	
5*	Hermansen et al. 2007		•	•
	Chacra et al. 2009		•	•
	Garber et al. 2008		•	•
	Pratley et al. 2009		•	•
	Lewin et al. 2012		•	•

*Für den indirekten Vergleich irrelevante Studienarme dieser Studien sind in der Tabelle nicht berücksichtigt

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.



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

Tabelle 4-111: Operationalisierung von Gewichtsveränderung

Studie	Operationalisierung
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1.1	
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin	
DeFronzo 1995 (Protokoll 2)	anhand der Änderung des Gewichts in kg von Studienbeginn bis Woche 29
Goldstein 2003	anhand der Änderung des Gewichts in kg von Studienbeginn bis Woche 18
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor	
Hermansen et al. 2007	anhand der Änderung des Gewichts in kg von Studienbeginn bis Woche 24
Chacra et al. 2009	anhand der Änderung des Gewichts in kg von Studienbeginn bis Woche 24
Garber et al. 2008	anhand der Änderung des Gewichts in kg von Studienbeginn bis Woche 24
Pratley et al. 2009	anhand der Änderung des Gewichts in kg von Studienbeginn bis Woche 26
Lewin et al. 2012	anhand der Änderung des Gewichts in kg von Studienbeginn bis Woche 26

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-112: 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
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1.1						
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin						
DeFronzo 1995 (Protokoll 2)	niedrig	ja	ja	nein	nein	niedrig
Goldstein 2003	niedrig	ja	ja	unklar*	nein	niedrig
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	niedrig	ja	ja	unklar*	nein	niedrig
Chacra et al. 2009	niedrig	ja	ja	unklar*	nein	niedrig
Garber et al. 2008	niedrig	ja	ja	unklar*	nein	niedrig
Pratley et al. 2009	niedrig	ja	ja	unklar*	nein	niedrig
Lewin et al. 2012	niedrig	ja	ja	unklar*	nein	niedrig

*es liegt nur die Publikation der Studie vor (kein Studienbericht)

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wurde das Verzerrungspotenzial für den zu bewertenden Endpunkt „Gewichtsveränderung“ für alle Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es keine sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

Eine ergebnisunabhängige Berichterstattung kann mit hoher Exaktheit nur kontrolliert werden, wenn sowohl Publikation wie auch der vollständige Studienbericht vorliegen. Für die RCT, die einen DPP-4-Inhibitor eingeschlossen haben, sowie für die Studie von Goldstein et al. (2003) lagen keine Studienberichte vor. Explizite Hinweise auf eine ergebnisabhängige Berichterstattung ergaben sich aber für keine der vorliegenden Studien.

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 4-113: Ergebnisse für Gewichtsveränderung aus RCT für indirekte Vergleiche

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Dapagliflozin (B)				Komparator: Sulfonylharnstoff +Placebo (A)				Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4- Inhibitor (C)				Gruppen- unterschied kg Mittelwert (SE) [95%-KI]
	N	Ausgangs- wert kg Mittelwert (SD)	End- wert kg Mittel- wert (SD)	Veränder- ung kg Mittelwert (SE)	N	Ausgangs- wert kg Mittel- wert (SE/SD)	Endwert kg Mittel- wert (SE/SD)	Veränder- ung kg Mittelwert (SE/SD)	N	Ausgangs- wert kg Mittel- wert (SE/SD)	End- wert kg Mittel- wert (SE/SD)	Verände- rung kg Mittel- wert (SE/SD)	
D1690C00005													
24 Wochen Studien- population	151	80,56 (17,869)	78,31 (17,804)	-2,26 (0,2217)	145	80,94 (15,773)	80,22 (15,516)	-0,72 (0,2263)					-1,54 (0,3168) [-2,17;-0,92]
Zielpopulation	129	80,76 (18,300)	78,42 (18,273)	-2,34 (0,2468)	113	80,72 (16,178)	80,11 (16,031)	-0,62 (0,2637)					-1,72 (0,3611) [-2,43;-1,01]
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin													
DeFronzo 1995					206	92,09 (1,01) ^a	91,82 (1,02) ^a	-0,30 (0,15) ^a	212	91,71 (1,07) ^a	92,30 (1,08) ^a	0,53 (0,20) ^a	0,83 (0,25) [0,34;1,32]*
Goldstein 2003					79	89,9 (17,3) ^b		-0,4	80	94,0 (18,0) ^b		-0,3	nicht berichtet ⁺
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor													
Hermansen et al. 2007					104	85,1 (22,6) ^b		0,0 (0,33 ^{a,*}) [-0,6;0,7]	104	85,8 (22,5) ^b		1,1 (0,33 ^{a,*}) [0,5;1,8]	1,10 (0,4667) [0,19;2,01]*
Chacra et al. 2009					267	75,6 (17,4) ^b		0,3	253	76,2 (17,6) ^b		0,8	0,50 (0,3) [-0,09;1,09]**

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Dapagliflozin (B)			Komparator: Sulfonylharnstoff +Placebo (A)			Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4- Inhibitor (C)			Gruppen- unterschied kg Mittelwert (SE) [95%-KI]			
	N	Ausgangs- wert kg Mittelwert (SD)	End- wert kg Mittel- wert (SD)	Veränder- ung kg Mittelwer- t/(SE)	N	Ausgang swert kg Mittel- wert (SE/SD)	Endwert kg Mittel- wert (SE/SD)	Verände- rung kg Mittel- wert (SE/SD)	N		Ausgang swert kg Mittel- wert (SE/SD)	End- wert kg Mittel- wert (SE/SD)	Verände- rung kg Mittel- wert (SE/SD)
Garber et al. 2008					144	89,4 (1,6) ^a		-0,4 (0,3) ^a	132	91,5 (1,6) ^a		-0,1 (0,3) ^a	0,30 (0,4243) [-0,53;1,13]*
Pratley et al. 2009					99			-0,2 (0,28) ^a	198			0,68 (0,19) ^a	0,88 (0,3384) [0,22;1,54]*
Lewin et al. 2012					82	76,1 (17,0) ^b		-0,01	158	74,5 (17,0) ^b		0,43	0,44 (0,4628) [-0,47;1,35]**

^a: Standardfehler (SE); ^b: Standardabweichung (SD); KI: Konfidenzintervall

*eigene Berechnung / Schätzung

**Streuungsmaß geschätzt aus dem Median der Standardabweichungen der anderen DPP-4-Studien

[†]Gruppenunterschied nicht berechenbar, da bei der Gewichtsveränderung die Standardabweichung oder Standardfehler nicht berichtet wird.

Ergebnisse der Studien mit der zweckmäßigen Vergleichstherapie

In der Publikation DeFronzo zeigte sich eine Gewichtszunahme unter Sulfonylharnstoff+Metformin, während sich unter Sulfonylharnstoff+Placebo eine Gewichtsabnahme zeigte. Der Unterschied zwischen beiden Gruppen war signifikant. In der Studie von Goldstein konnte ein Effektschätzer aufgrund des fehlenden Streuungsmaßes nicht berechnet werden. Die absoluten Zahlen zeigten aber keinen Unterschied in beiden Behandlungsarmen.

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.*

Der indirekte Vergleich wurde zwischen der Studie von DeFronzo und der Studie D1690C00005 durchgeführt, da in der Publikation von Goldstein keine verwendbaren Daten berichtet wurden.

Tabelle 4-114: Ergebnisse für Gewichtsveränderung aus indirekte Vergleiche

	Mittelwertsdifferenzen des indirekten Vergleiches Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin kg [95%-KI]
D1690C00005 (Studienpopulation) DeFronzo	-2,37 [-3,16;-1,58]
D1690C00005 (Zielpopulation) und DeFronzo	-2,55 [-3,41;-1,69]

KI: Konfidenzintervall

Der indirekte Vergleich von Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin zeigte einen signifikanten Unterschied zugunsten von Sulfonylharnstoff+Dapagliflozin. Dieses Ergebnis lässt sich nicht durch ein höheres Ausgangsgewicht erklären, da die Patienten in der Studie von DeFronzo im Schnitt deutlich schwerer waren als die Patienten in der Studie D1690C00005 und daher eher ein Potential für eine Gewichtsabnahme gehabt hätten. Selbst wenn also eine Studienverzerrung durch das Ausgangsgewicht besteht, so würde die Verzerrung eher zum Nachteil von Sulfonylharnstoff+Dapagliflozin wirksam werden.

Ergebnisse der Studien mit der alternativen Vergleichstherapie

Für den Endpunkt „Gewichtsveränderung“ standen alle identifizierten DPP-4-Studien des Vergleiches Sulfonylharnstoff+Placebo versus Sulfonylharnstoff+DPP-4-Inhibitor zur Verfügung. Zwei der Studien (Chacra 2009; Lewin 2012) berichteten keine Streuungsmaße (Standardabweichung oder Standardfehler), als Standardabweichung wurde für diese Studien daher der Median der Standardabweichungen der anderen DPP-4-Inhibitor-Studien angenommen (Median der Standardabweichungen für Mittelwertsdifferenz für DPP-4-Studienarm: SD=3,4; für Placebo-Studienarm: SD=3,4). In einer Sensitivitätsanalyse wurde überprüft, ob sich die Ergebnisse des indirekten Vergleiches durch den Ausschluss dieser zwei Studien in ihrer Aussage veränderten.

Alle Studien wiesen in Bezug auf Veränderungen des Körpergewichts auf einen Nachteil von DPP-4-Inhibitoren hin; die beobachteten Gruppenunterschiede waren aber nur in den Studien von Hermansen et al. 2007 und Pratley et al. 2009 statistisch signifikant. Die Meta-Analyse wies auf keine Heterogenität der Studienergebnisse hin ($I^2=0\%$; $p=0.64$). Die Zusammenfassung der Ergebnisse in einen gemeinsamen Effektschätzer erschien daher gerechtfertigt.

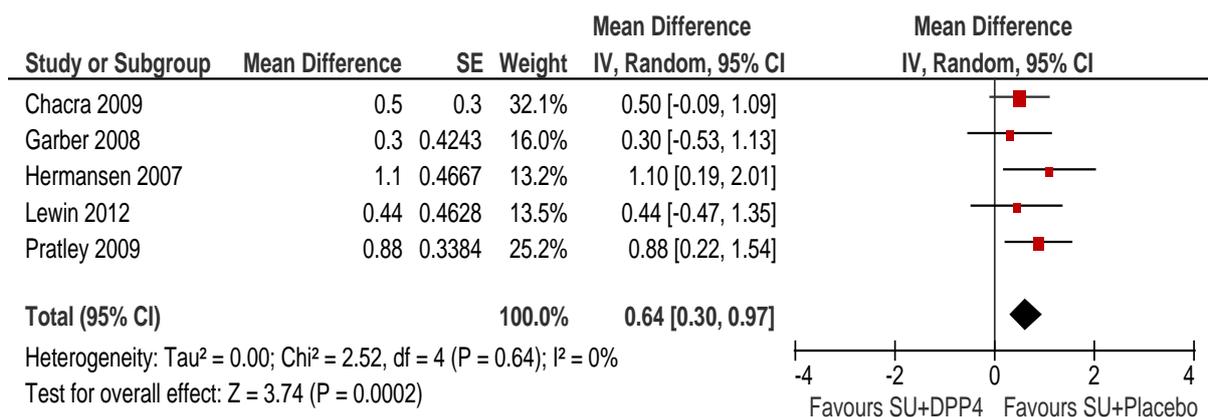


Abbildung 14: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gewichtsveränderung“

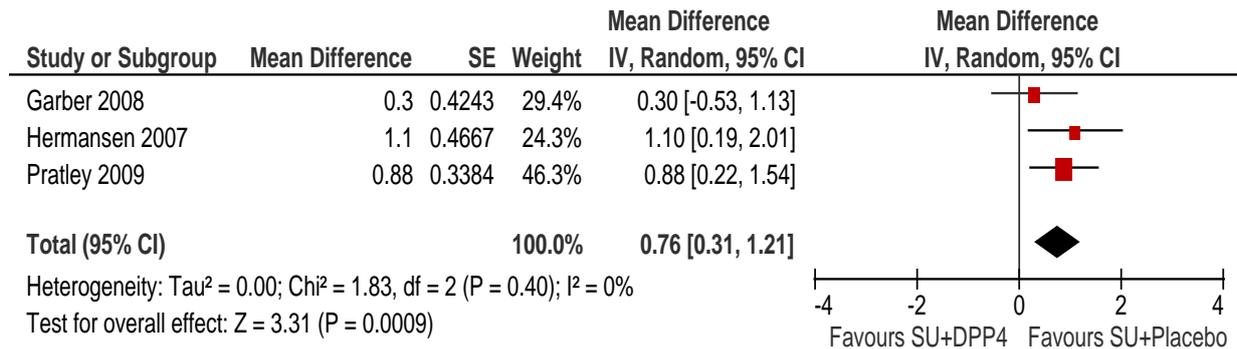


Abbildung 15: Sensitivitätsanalyse: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor (nur Studien mit Angabe von Streuungsmaßen) vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gewichtsveränderung“

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 4-115: Ergebnisse für Gewichtsveränderung - indirekte Vergleiche

	Mittelwertsdifferenz des indirekten Vergleiches Sulfonylharnstoff +Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor Gewicht (kg) [95%-KI]
D1690C00005 (Studienpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	-2,18 [-2,89;-1,47]
D1690C00005 (Zielpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	-2,36 [-3,14;-1,58]
D1690C00005 (Studienpopulation) und Hermansen 2007 (Sitagliptin)	-2,64 [-3,75;-1,53]
D1690C00005 (Zielpopulation) und Hermansen 2007 (Sitagliptin)	-2,82 [-3,98;-1,66]
Sensitivitätsanalyse: D1690C00005 (Studienpopulation) und alle DPP-4-Studien, die Streuungsmaße berichtet haben	-2,30 [-3,07;-1,53]
Sensitivitätsanalyse: D1690C00005 (Zielpopulation) und alle DPP-4-Studien, die Streuungsmaße berichtet haben	-2,48 [-3,32;-1,64]

KI: Konfidenzintervall

Der indirekte Vergleich von Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor zeigte einen signifikanten Unterschied zugunsten von Sulfonylharnstoff+Dapagliflozin, gleiches galt für den indirekten Vergleich gegen die Leitsubstanz Sitagliptin. Die Sensitivitätsanalyse ergab sehr ähnliche Ergebnisse wie die Hauptanalyse.

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:

Für den indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor wurden in den identifizierten RCT keine relevanten Daten berichtet. Die einzige Studie, die über das Ausmaß der Veränderungen von Blutdruckwerten berichtete, war Chacra

et al. 2009, es wurden allerdings keine Streuungsmaße (Standardfehler / Standardabweichung) berichtet. Im Folgenden werden daher ausschließlich die Ergebnisse des indirekten Vergleiches Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin berichtet.

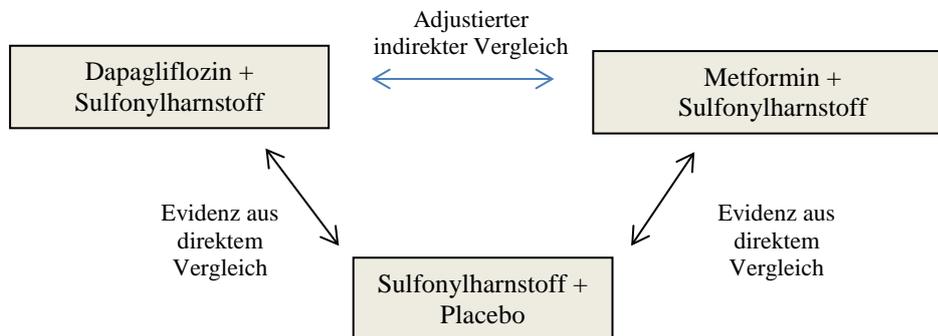
Tabelle 4-116: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Dapagliflozin	Sulfonylharnstoff +Placebo	Sulfonylharnstoff + Metformin	Metformin+ Placebo*
1	D1690C00005	•	•		
1	DeFronzo 1995 (Protokoll 2)		•	•	•

*dieser Behandlungsarm wird nicht beachtet

In der Studie von Goldstein et al. 2003 werden keine Daten zu den Effekten der Behandlung auf den Blutdruck berichtet, sie findet daher bzgl. dieses Endpunktes keine Berücksichtigung.

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.



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

Tabelle 4-117: Operationalisierung von Blutdrucksenkung

Studie	Operationalisierung
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1.1	
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin	
DeFronzo 1995 (Protokoll 2)	Anhand der Änderung des systolischen und diastolischen Blutdrucks (mmHg) von Studienbeginn bis Woche 25 (fehlende Werte nach der LOCF Methode ergänzt). Der systolische und diastolische Blutdruck wurde bei jedem Studienbesuch gemessen, sowohl in Rückenlage, wie auch im Sitzen. Es wurden keine speziellen Vorkehrungen getroffen, um die Blutdruckmessung zu standardisieren. Für den indirekten Vergleich fanden die Messungen in Rückenlage Verwendung, um eine möglichst hohe Vergleichbarkeit mit der Studie D1690C00005 zu gewährleisten.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-118: 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
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1.5						
DeFronzo 1995 (Protokoll 2)	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wurde das Verzerrungspotenzial für den zu bewertenden Endpunkt „Blutdrucksenkung“ für alle Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es keine sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

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 zusammenfassen

Tabelle 4-119: Ergebnisse für Blutdrucksenkung aus RCT für indirekte Vergleiche

Studie		Zu bewertende Arzneimittel: Sulfonylharnstoff +Dapagliflozin (B)			Komparator: Sulfonylharnstoff +Placebo (A)			Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin (C)			Gruppen- unter- schied % Mittel- wert mmHg (SE) [95%-KI]		
		N	Ausgangs- wert % Mittelwert mmHg (SD)	Endwert % Mittel- wert mmHg (SD)	Verän- derung % Mittel- wert° mmHg (SE)	N	Ausgangs- wert % Mittelwert mmHg (SE/SD)	Endwert % Mittel- wert mmHg (SE/SD)	Verän- derung % Mittel- wert mmHg (SE/ SD)	N		Aus- gangs- wert % Mittel- wert mmHg (SE/SD)	Endwert % Mittel- wert mmHg (SE/SD)
Studie D1690C 00005													
Studien- popu- lation	Systo- lisch	150	132,4 (13,71)	127,6 (14,20)	-5,0 (0,930)	145	133,3 (13,86)	131,9 (13,49)	-1,2 (0,946)				-3,8 (1,327) [-6,4;-1,2]
	Diasto- lisch	150	78,7 (8,24)	76,0 (8,05)	-2,8 (0,585)	145	79,8 (7,76)	77,9 (7,79)	-1,4 (0,596)				-1,4 (0,836) [-3,0;0,3]
Ziel- popu- lation	Systo- lisch	129	131,9 (13,52)	126,5 (13,14)	-5,5 (0,978)	113	134,0 (14,42)	132,1 (13,8)	-1,1 (1,047)				-4,4 (1,432) [-7,2;-1,6]
	Diasto- lisch	129	79,1 (8,46)	76,1 (8,17)	-3,0 (0,616)	113	80,6 (7,55)	78,8 (7,42)	-1,0 (0,661)				-2,0 (0,904) [-3,8;-0,2]

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff +Dapagliflozin (B)				Komparator: Sulfonylharnstoff +Placebo (A)			Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin (C)			Gruppen- unter- schied % Mittel- wert mmHg (SE) [95%-KI]		
	N	Ausgangs- wert % Mittelwert mmHg (SD)	Endwert % Mittel- wert mmHg (SD)	Verän- derung % Mittel- wert ^o mmHg (SE)	N	Ausgangs- wert % Mittelwert mmHg (SE/SD)	Endwert % Mittel- wert mmHg (SE/SD)	Verän- derung % Mittel- wert mmHg (SE/ SD)	N	Aus- gangs- wert % Mittel- wert mmHg (SE/SD)		Endwert % Mittel- wert mmHg (SE/SD)	Verän- der- ung % Mittel- wert mmHg (SE/ SD)
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin													
De Fronzo 1995	Systo- lisch				206	131,63 (1,14) ^a	134,39 (1,19) ^a	2,70 (1,06) ^a	212	130,75 (1,11) ^a	133,84 (1,16) ^a	3,14 (1,08) ^a	0,44 (0,15133) ^{a,*} [-2,53;- 3,41] [*]
	Diasto- lisch				206	80,40 (0,61) ^a	81,70 (0,68) ^a	1,32 (0,66) ^a	212	80,68 (0,59) ^a	82,07 (0,60) ^a	1,40 (0,62) ^a	0,08 (0,9055) ^{a,*} [-1,69;- 1,85] [*]

^a Standardfehler (SE); KI: Konfidenzintervall; ^{*}eigene Berechnung / Schätzung

Ergebnisse der Studien mit der zweckmäßigen Vergleichstherapie

In der Publikation von DeFronzo zeigten sich keine statistisch signifikanten Unterschiede in Bezug auf Veränderungen des systolischen oder diastolischen Blutdruckes zwischen den beiden Behandlungsarmen.

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 Blutdrucksenkung - indirekte Vergleiche

		Mittelwertsdifferenzen des indirekten Vergleiches Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin mm Hg [95%-KI]
Studienpopulation	Systolischer Blutdruck: D1690C00005 und DeFronzo	-4,24 [-6,86;-1,62]
	Diastolischer Blutdruck: D1690C00005 und DeFronzo	-1,48 [-3,90;0,94]
Zielpopulation	Systolischer Blutdruck: D1690C00005 und DeFronzo	-4,84 [-7,66;-2,02]
	Diastolischer Blutdruck: D1690C00005 und DeFronzo	-2,08 [-4,59;0,43]

KI: Konfidenzintervall

Der indirekte Vergleich von Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin zeigte in Bezug auf die Senkung des systolischen Blutdrucks einen signifikanten Unterschied zugunsten von Sulfonylharnstoff+Dapagliflozin (sowohl für die Studienpopulation, wie auch für die Zielpopulation). Bezüglich des diastolischen Blutdrucks fand sich ebenfalls ein Vorteil von Sulfonylharnstoff+Dapagliflozin, der aber für keine der untersuchten Populationen statistisch signifikant war.

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 (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff	Sulfonylharnstoff	Sulfonylharnstoff	Metformin
		+ Dapagliflozin	+ Placebo	+ Metformin	+ Placebo*
1	D1690C00005	•	•		
2	DeFronzo 1995 (Protokoll 2)		•	•	•
	Goldstein 2003		•	•	•

*dieser Behandlungsarm wird nicht beachtet

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

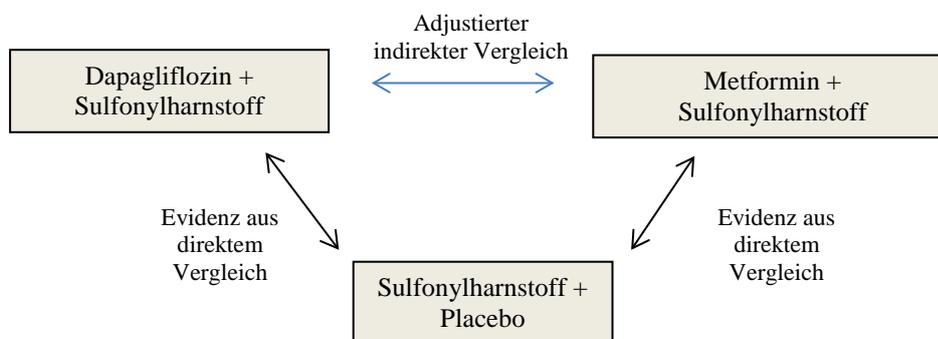
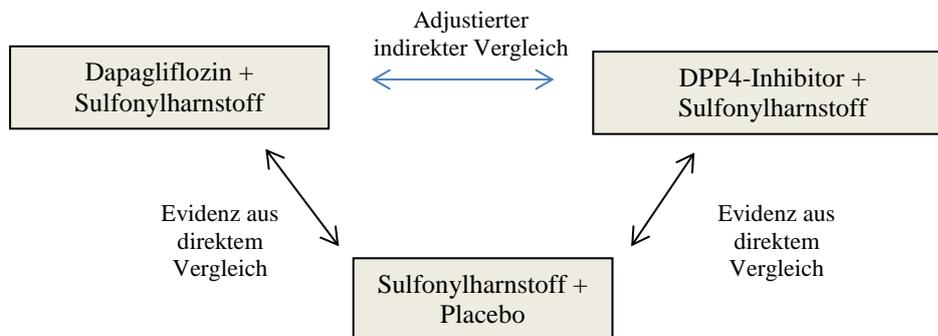


Tabelle 4-122 Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Dapagliflozin	Sulfonylharnstoff + Placebo	Sulfonylharnstoff + DPP-4-Inhibitor
1	D1690C00005	•	•	
5*	Hermansen et al. 2007		•	•
	Chacra et al. 2009		•	•
	Garber et al. 2008		•	•
	Pratley et al. 2009		•	•
	Lewin et al. 2012		•	•

*Für den indirekten Vergleich irrelevante Studienarme dieser Studien sind in der Tabelle nicht berücksichtigt.

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.



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
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1.1	
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin	
DeFronzo 1995 (Protokoll 2)	<p>Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrachen (innerhalb der Behandlungsphase)</p>
Goldstein 2003	<p>Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrachen (innerhalb der Behandlungsphase)</p>
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor	
Hermansen et al. 2007	<p>Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrachen (innerhalb der Behandlungsphase)</p>
Chacra et al. 2009	<p>Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrachen (innerhalb der Behandlungsphase)</p>
Garber et al. 2008	<p>Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrachen (innerhalb der Behandlungsphase)</p>
Pratley et al. 2009	<p>Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens</p>

Studie	Operationalisierung
	<p>einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrechen (innerhalb der Behandlungsphase)</p>
Lewin et al. 2012	<p>Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase)</p> <p>Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrechen (innerhalb der Behandlungsphase)</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. 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-F.

Tabelle 4-124: Bewertung des Verzerrungspotenzials für Unerwünschte 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
Studie D1690C00005 siehe Abschnitt 4.3.1.3.1.5						
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin						
DeFronzo 1995 (Protokoll 2)	niedrig	ja	ja	nein	nein	niedrig
Goldstein 2003	niedrig	ja	ja	unklar*	nein	niedrig
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	niedrig	ja	ja	unklar*	nein	niedrig
Chacra et al. 2009	niedrig	ja	ja	unklar*	nein	niedrig
Garber et al. 2008	niedrig	ja	ja	unklar*	nein	niedrig
Pratley et al. 2009	niedrig	ja	ja	unklar*	nein	niedrig
Lewin et al. 2012	niedrig	ja	ja	unklar*	nein	niedrig

*es lag nur die Publikation der Studie vor (kein Studienbericht)

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wurde das Verzerrungspotenzial für den zu bewertenden Endpunkt „unerwünschte Ereignisse“ für alle Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es keine sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

Eine ergebnisunabhängige Berichterstattung kann mit hoher Exaktheit nur kontrolliert werden, wenn sowohl Publikation wie auch der vollständige Studienbericht vorliegen. Für die RCTs, die einen DPP-4-Inhibitor eingeschlossen haben, sowie für die Studie von Goldstein et al. (2003) lagen keine Studienberichte vor. Explizite Hinweise auf eine ergebnisabhängige Berichterstattung ergaben sich aber für keine der vorliegenden Studien.

Bezüglich der Vergleichbarkeit der Studien für die Zwecke eines indirekten Vergleiches muss berücksichtigt werden, dass einige der Studien die „nicht-schweren Hypoglykämien“ nicht in die unerwünschten Ereignisse eingeschlossen haben (z.B. D1690C00005; Chacra, Lewin). Andere Studien schlossen wiederum die „nicht schweren Hypoglykämien“ mit ein (DeFronzo;

Garber) oder machten hierzu keine klaren Angaben (Goldstein; Hermansen; Pratley). Aufgrund der niedrigen Anzahl schwerwiegender Hypoglykämien in den vorliegenden Studien dürften diese methodischen Unterschiede aber keinen erheblichen Einfluss auf den Endpunkt „Gesamtrate schwerwiegender unerwünschter Ereignisse“ gehabt haben. Ein erheblicher Einfluss auf den Endpunkt „Gesamtrate unerwünschter Ereignisse“ erscheint ebenfalls eher unwahrscheinlich, da die Gesamtanzahl von unerwünschten Ereignissen gegenüber der Anzahl von Patienten mit mindestens einer Hypoglykämie groß war. Dennoch sollten die Ergebnisse des indirekten Vergleichs für diese Endpunkte vor dem Hintergrund methodischer Studienheterogenität mit besonderer Vorsicht interpretiert 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 zusammenfassen

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

Tabelle 4-125: Ergebnisse für Unerwünschte Ereignisse aus RCT für indirekte Vergleiche

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Dapagliflozin (B)		Komparator: Sulfonylharnstoff+Placebo (A)		Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4- Inhibitor (C)	Odds Ratio [95%-KI]
	N	n (%)	N	n (%)		
Studie D1690C00005						
Studienpopulation 24 Wochen						
Gesamtrate unerwünschter Ereignisse	151	76 (50,3)	146	69 (47,3)		1,13 [0,72;1,78]
Schwerwiegende, unerwünschte Ereignisse	151	9 (6,0)	146	7 (4,8)		1,26 [0,46;3,47]
Studienabbruch aufgrund un- erwünschter Ereignisse	151	4 (2,6)	146	3 (2,1)		1,30 [0,29;5,90]
Zielpopulation 24 Wochen						
Gesamtrate unerwünschter Ereignisse	129	63 (48,8)	114	53 (46,5)		1,09 [0,66;1,81]
Schwerwiegende, unerwünschte Ereignisse	129	7 (5,4)	114	5 (4,4)		1,25 [0,39;4,06]
Studienabbruch aufgrund un- erwünschter Ereignisse	129	1 (0,8)	114	2 (1,8)		0,44 [0,04;4,89]

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Dapagliflozin (B)		Komparator: Sulfonylharnstoff+Placebo (A)		Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4- Inhibitor (C)		Odds Ratio [95%-KI]
	N	n (%)	N	n (%)	N	n (%)	
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin							
DeFronzo 1995 (Protokoll 2)							
Gesamtrate unerwünschter Ereignisse			209	171 (81,8)	213	188 (88,3)	1,67 [0,97;2,88]
Schwerwiegende, unerwünschte Ereignisse			209	15 (7,2)	213	8 (3,8)	0,50 [0,21;1,22]
Studienabbruch aufgrund un- erwünschter Ereignisse			209	5 (2,4)	213	4 (1,9)	0,78 [0,21;2,95]
Goldstein 2003							
Gesamtrate unerwünschter Ereignisse			84	57 (67,9)	87	55 (63,2)	0,81 [0,43;1,53]
Schwerwiegende, unerwünschte Ereignisse			84	0 (0,0)	87	0 (0,0)	Nicht berechenbar
Studienabbruch aufgrund un- erwünschter Ereignisse			84	3 (3,6)	87	11 (12,6)	3,91 [1,05;14,55]
Studien für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor							
Hermansen 2007							
Gesamtrate unerwünschter Ereignisse			106	43 (40,6)	106	59 (55,7)	1,84 [1,07;3,17]
Schwerwiegende, unerwünschte Ereignisse			106	6 (5,7)	106	5 (4,7)	0,83 [0,24;2,79]
Studienabbruch aufgrund un-			106	1 (0,9)	106	3 (2,8)	3,06 [0,31;29,88]

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Dapagliflozin (B)		Komparator: Sulfonylharnstoff+Placebo (A)		Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4- Inhibitor (C)		Odds Ratio [95%-KI]
	N	n (%)	N	n (%)	N	n (%)	
erwünschter Ereignisse							
Chacra 2009							
Gesamtrate unerwünschter Ereignisse			267	205 (76,8)	253	183 (72,3)	0,79 [0,53;1,17]
Schwerwiegende, unerwünschte Ereignisse			267	6 (2,2)	253	6 (2,2)	1,06 [0,34;3,32]
Studienabbruch aufgrund unerwünschter Ereignisse			267	1 (0,4)	253	1 (0,4)	1,06 [0,07;16,97]
Garber 2008							
Gesamtrate unerwünschter Ereignisse			176	113 (64,2)	170	114 (67,1)	1,13 [0,73;1,77]
Schwerwiegende, unerwünschte Ereignisse			176	9 (5,1)	170	5 (2,9)	0,56 [0,18;1,71]
Studienabbruch aufgrund unerwünschter Ereignisse			176	3 (1,7)	170	4 (2,4)	1,39 [0,31;6,30]
Pratley 2009							
Gesamtrate unerwünschter Ereignisse			99	53 (53,5)	198	125 (63,1)	1,49 [0,91;2,42]
Schwerwiegende, unerwünschte Ereignisse			99	2 (2,0)	198	11 (5,6)	2,85 [0,62;13,13]
Studienabbruch aufgrund unerwünschter Ereignisse			99	2 (2,0)	198	4 (2,0)	1,00 [0,18;5,56]

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Dapagliflozin (B)		Komparator: Sulfonylharnstoff+Placebo (A)		Zweckmäßige Vergleichstherapie: Sulfonylharnstoff +Metformin bzw. alternative Vergleichstherapie DPP-4- Inhibitor (C)		Odds Ratio [95%-KI]
	N	n (%)	N	n (%)	N	n (%)	
Lewin 2012							
Gesamtrate unerwünschter Ereignisse			84	36 (42,9)	161	68 (42,2)	0,97 [0,57;1,66]
Schwerwiegende, unerwünschte Ereignisse			84	1 (1,19)	161	5 (3,11)	2,66[0,31;23,15]
Studienabbruch aufgrund un- erwünschter Ereignisse			84	3 (3,57)	161	5 (3,11)	0,87 [0,20; 3,71]

KI: Konfidenzintervall

Ergebnisse der Studien mit der zweckmäßigen Vergleichstherapie

In der von Goldstein et al. berichteten Studie waren Studienabbrüche aufgrund von unerwünschten Wirkungen unter Sulfonylharnstoff+Metformin signifikant häufiger als unter Sulfonylharnstoff+Placebo. Desweiteren fanden sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen.

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.*

Gesamtrate unerwünschter Ereignisse:

Bezüglich des Endpunktes „Gesamtrate unerwünschter Ereignisse“ fanden sich Hinweise auf eine moderate Studienheterogenität, die eine Berechnung eines gemeinsamen Risikoschätzers aus Goldstein et al. 2003 und DeFronzo et al. 1995 nicht sinnvoll erscheinen ließen ($I^2=65\%$; $p=0.09$). Der indirekte Vergleich wurde daher getrennt für beide Studien durchgeführt und im Folgenden dargestellt.

Tabelle 4-126: Ergebnisse für Gesamtrate unerwünschter Ereignisse – indirekte Vergleiche

	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff +Dapagliflozin versus Sulfonylharnstoff+Metformin [95%-KI]
D1690C00005 (Studienpopulation) und DeFronzo	0,68 [0,33;1,37]
D1690C00005 (Zielpopulation) und DeFronzo	0,65 [0,31;1,37]
D1690C00005 (Studienpopulation) und Goldstein	1,40 [0,64;3,04]
D1690C00005 (Zielpopulation) und Goldstein	1,35 [0,60;3,03]

KI: Konfidenzintervall

Es ließ sich kein Unterschied zwischen Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin bezüglich der Gesamtrate unerwünschter Ereignisse feststellen.

Gesamtrate schwerwiegender unerwünschter Ereignisse:

Für diesen Endpunkt konnte ein indirekter Vergleich unter Verwendung der Studie von DeFronzo durchgeführt werden. Bei Goldstein wurde für keinen Patienten ein

schwerwiegendes Ereigniss berichtet, so dass eine Meta-analytische Zusammenfassung beider Studien nicht infrage kam.

Tabelle 4-127: Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse – indirekte Vergleiche

	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff +Dapagliflozin versus Sulfonylharnstoff+Metformin [95%-KI]
D1690C00005 (Studienpopulation) und DeFronzo	2,52 [0,66;9,62]
D1690C00005 (Zielpopulation) und DeFronzo	2,50 [0,58;10,82]

KI: Konfidenzintervall

Es ließ sich kein signifikanter Unterschied zwischen Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin bezüglich der Gesamtrate schwerwiegender unerwünschter Ereignisse feststellen.

Studienabbrüche aufgrund unerwünschter Ereignisse:

Beide Studien für den direkten Vergleich Sulfonylharnstoff+Metformin versus Sulfonylharnstoff+Placebo berichteten die Anzahl der Studienabbrüche aufgrund unerwünschter Ereignisse. Die statistische Heterogenität der Studien ist moderat ($I^2=65%$, $p=0,09$) daher wurde der indirekte Vergleich mit den Einzelergebnissen der Studien durchgeführt.

Tabelle 4-128: Ergebnisse der Studienabbrüche aufgrund unerwünschter Ereignisse - indirekte Vergleiche

	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin [95%-KI]
D1690C00005 (Studienpopulation) und DeFronzo	1,67 [0,22;12,36]
D1690C00005 (Zielpopulation) und DeFronzo	0,56 [0,04;8,76]
D1690C00005 (Studienpopulation) und Goldstein	0,33 [0,05;2,45]
D1690C00005 (Zielpopulation) und Goldstein	0,11 [0,01;1,74]

KI: Konfidenzintervall

Es ließ sich kein signifikanter Unterschied bezüglich der Studienabbrüche aufgrund unerwünschter Ereignisse feststellen.

Ergebnisse der Studien mit der alternativen Vergleichstherapie

In der von Hermansen et al. berichteten Studie war die Gesamtrate unerwünschter Ereignisse unter Sulfonylharnstoff+DPP-4-Inhibitor signifikant häufiger als unter Sulfonylharnstoff+Placebo. Desweiteren fanden sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen.

Gesamtrate unerwünschter Ereignisse:

Bezüglich des Endpunktes „Gesamtrate unerwünschter Ereignisse“ fanden sich Hinweise auf eine moderate Studienheterogenität mit $I^2=48\%$; $p=0.10$, vermutlich aufgrund der bereits diskutierten Unterschiede in der Berücksichtigung von Hypoglykämien in der Gesamtrate unerwünschter Ereignisse.

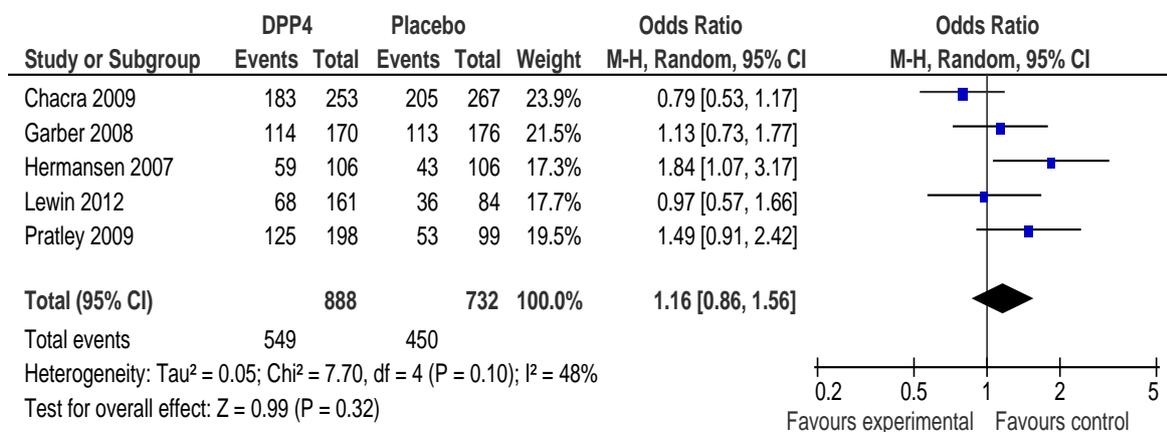


Abbildung 16: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate unerwünschter Ereignisse“

Gesamtrate schwerwiegender unerwünschter Ereignisse:

Bezüglich des Endpunktes „Gesamtrate schwerwiegender unerwünschter Ereignisse“ fanden sich keine Hinweise auf eine bedeutende Studienheterogenität ($I^2=0\%$; $p=0.44$). Die Zusammenfassung der Ergebnisse in einen gemeinsamen Effektschätzer erschien daher gerechtfertigt.

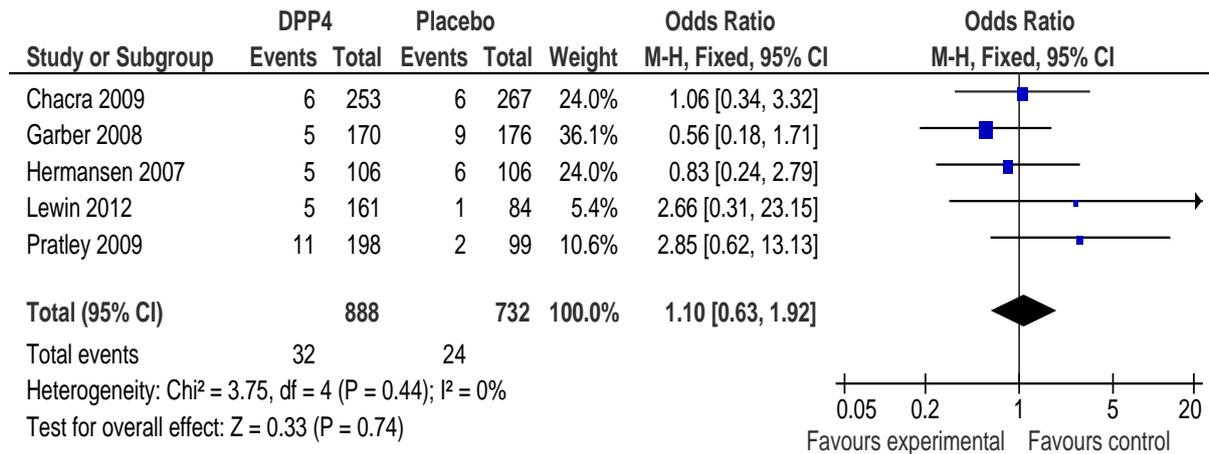


Abbildung 17: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate schwerwiegender unerwünschter Ereignisse“

Studienabbrüche aufgrund unerwünschter Ereignisse:

Bezüglich des Endpunktes „Studienabbrüche aufgrund unerwünschter Ereignisse“ fanden sich keine Hinweise auf eine bedeutende Studienheterogenität (I²=0%; p=0.92). Die Zusammenfassung der Ergebnisse in einen gemeinsamen Effektschätzer erschien daher gerechtfertigt.

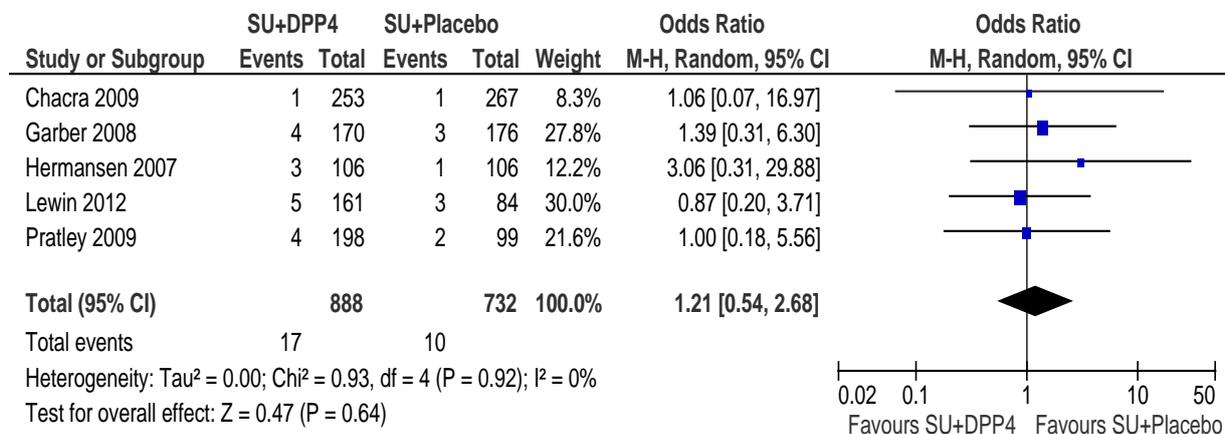


Abbildung 18: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Studienabbrüche aufgrund unerwünschter Ereignisse“

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 4-129: Ergebnisse für Gesamtrate unerwünschter Ereignisse - indirekte Vergleiche

	Odds ratio des indirekten Vergleiches Sulfonylharnstoff +Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor [95%-KI]
D1690C00005 (Studienpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	0,97 [0,57;1,67]
D1690C00005 (Zielpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	0,94 [0,52;1,69]
D1690C00005 (Studienpopulation) und Hermansen 2007 (Sitagliptin)	0,61 [0,30;1,25]
D1690C00005 (Zielpopulation) und Hermansen 2007 (Sitagliptin)	0,59 [0,28;1,24]

KI: Konfidenzintervall

Damit ergab der indirekte Vergleich in Bezug auf den Endpunkt „Gesamtrate unerwünschter Ereignisse“ keine statistisch signifikanten Unterschiede zwischen der zu bewertenden Therapie und der alternativen Vergleichstherapie (DPP-4-Inhibitoren als Gruppe bzw. Sitagliptin).

Tabelle 4-130 Ergebnisse für Gesamtrate schwerwiegender unerwünschter Ereignisse - indirekte Vergleiche

	Odds ratio des indirekten Vergleiches Sulfonylharnstoff +Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor [95%-KI]
D1690C00005 (Studienpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	1,15 [0,36;3,63]
D1690C00005 (Zielpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	1,14 [0,31;4,16]
D1690C00005 (Studienpopulation) und Hermansen 2007 (Sitagliptin)	1,52 [0,31;7,44]
D1690C00005 (Zielpopulation) und Hermansen 2007 (Sitagliptin)	1,51 [0,28;8,21]

KI: Konfidenzintervall

Damit ergab der indirekte Vergleich in Bezug auf den Endpunkt „Gesamtrate schwerwiegender unerwünschter Ereignisse“ keine statistisch signifikanten Unterschiede

zwischen der zu bewertenden Therapie und der alternativen Vergleichstherapie (DPP-4-Inhibitoren als Gruppe bzw. Sitagliptin).

Tabelle 4-131: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse - indirekte Vergleiche

	Odds ratio des indirekten Vergleiches Sulfonylharnstoff +Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor [95%-KI]
D1690C00005 (Studienpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	1,07 [0,20;5,92]
D1690C00005 (Zielpopulation) und DPP-4-Inhibitoren insgesamt (aus Meta-Analyse)	0,36 [0,03;4,58]
D1690C00005 (Studienpopulation) und Hermansen 2007 (Sitagliptin)	0,42 [0,03;6,55]
D1690C00005 (Zielpopulation) und Hermansen 2007 (Sitagliptin)	0,14 [0,01;3,96]

KI: Konfidenzintervall

Damit ergab der indirekte Vergleich in Bezug auf den Endpunkt „Studienabbrüche aufgrund unerwünschter Ereignisse“ keine statistisch signifikanten Unterschiede zwischen der zu bewertenden Therapie und der alternativen Vergleichstherapie (DPP-4-Inhibitoren als Gruppe bzw. Sitagliptin).

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 4.3.1.3.1.1.

Es konnten keine Subgruppenanalysen auf Basis indirekter Vergleiche vorgelegt werden, da für die im indirekten Vergleich verwendeten Studien nicht ausreichend Daten zur Durchführung von aussagekräftigen Subgruppenanalysen zur Verfügung standen.

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 wurden keine nicht randomisierten vergleichende Studien herangezogen.

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-F. 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-F.

Tabelle 4-132: 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 4-133: 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-F. 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-F.

Tabelle 4-134: 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 vergleichende Studien herangezogen.

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 4.3.1.3.1.1.

Es wurden keine nicht randomisierten vergleichende Studien herangezogen.

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*

Es wurden keine weiteren Untersuchungen zur Bewertung des Zusatznutzen eingesetzt.

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 4-135: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter 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 weiteren Untersuchungen zur Bewertung des Zusatznutzen eingesetzt.

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 weiteren Untersuchungen zur Bewertung des Zusatznutzen eingesetzt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.1.

Es wurden keine weiteren Untersuchungen zur Bewertung des Zusatznutzen eingesetzt.

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.

Zur Bewertung des medizinischen Zusatznutzens von Dapagliflozin wurden indirekte Vergleiche herangezogen, da keine direkten Studien weder mit der zweckmäßigen noch mit der alternativen Vergleichstherapie vorlagen.

Im Folgenden werden die Ergebnisse zu den einzelnen Endpunkten für beide indirekte Vergleiche kurz zusammengefasst. Wenn nicht weiter spezifiziert, beziehen sich alle Angaben auf den Vergleich mit der **Zielpopulation** für Dapagliflozin.

Ergebnisse des indirekten Vergleichs Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin

Ausmaß der Blutzuckerkontrolle:

Die Studien DeFronzo und Goldstein berichteten diesen Endpunkt mit dem Behandlungsarm Sulfonylharnstoff+Metformin, doch aufgrund der unterschiedlichen Behandlungsdauer und der klinischen Heterogenität lässt sich kein Gesamtschätzer berechnen. Der indirekte Vergleich der Studie D1690C00005 mit der Studie von Goldstein ergab keinen signifikanten Unterschied in der Blutzuckerkontrolle (MWD [95%-KI] des indirekten Vergleiches mit Goldstein 0,32% [-0,03;0,67]). Die MWD des indirekten Vergleiches mit DeFronzo zeigte einen signifikanten Unterscheid zugunsten der Vergleichstherapie Sulfonylharnstoff+Metformin (MWD [95%-KI]: 1,14% [0,79;1,49]). Die Vergleichbarkeit der Studie von DeFronzo und der Studie D1690C00005 war allerdings in Bezug auf verschiedene Aspekte eingeschränkt. So war der Ausgangswert für HbA1c in der Studie von DeFronzo höher als in der Studie D1690C00005; zudem hatten in der DeFronzo-Studie die Patienten in der Sulfonylharnstoff+Metformin Gruppe einen höheren HbA1c-Ausgangswert (8,80%) als die Patienten in der Sulfonylharnstoff+Placebo Gruppe (8,52%). Dieses kann zu einem scheinbar besonders ausgeprägten Behandlungseffekt durch Metformin beigetragen haben. Darüber hinaus wurden in der Studie von DeFronzo die Patienten über eine Zeit von vier Wochen mit einer Sulfonylharnstoff-Dosis behandelt, die fast doppelt so hoch war wie in Deutschland maximal zugelassen, bevor die Randomisierung erfolgte. Ein „Nachwirkeffekt“ von Sulfonylharnstoff könnte ebenfalls zu einer besonders ausgeprägten HbA1c-Senkung von Sulfonylharnstoff+Metformin beigetragen haben.

Hypoglykämien:

Die Publikation von DeFronzo berichtet wie auch die Studie D1690C00005 die Gesamtrate der Patienten mit mindestens einer Hypoglykämie und konnten somit für den indirekten Vergleich herangezogen werden. Es ergaben sich für die Zielpopulation keine signifikanten Unterschiede zwischen Sulfonylharnstoff+Dapagliflozin und Sulfonylharnstoff+Metformin (OR [95%-KI] durch den indirekten Vergleich mit der DeFronzo-Studie: 0,36 [0,09;1,39]); für die Studienpopulation war die Gesamtrate der Patienten mit mindestens einer Hypoglykämie unter Sulfonylharnstoff+Dapagliflozin signifikant kleiner als im Sulfonylharnstoff+Metformin Arm (OR [95%-KI]: 0,27 [0,08;0,97]), vermutlich durch das Erreichen der Signifikanzschwelle aufgrund der höheren Patientenzahl in dieser Analyse. In der Publikation von Goldstein wird die Gesamtrate nicht berichtet. Schwerwiegende

Hypoglykämien sind in keiner der betrachteten Studien aufgetreten, ein indirekter Vergleich ließ sich daher für diesen Endpunkt nicht durchführen.

Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

Hinsichtlich der Änderung des HbA1c-Wertes ergab sich im indirekten Vergleich von DeFronzo und D1690C00005 ein Nachteil in der Blutzuckersenkung für Sulfonylharnstoff+Dapagliflozin. Die Gesamtrate von Hypoglykämien war unter Sulfonylharnstoff+Dapagliflozin niedriger, bei Berücksichtigung der Studienpopulation sogar statistisch signifikant niedriger. Insgesamt ergab sich in der Zusammenschau kein Vorteil für eine der beiden Behandlungsoptionen.

Gewichtsveränderungen:

Die Daten zur Veränderung des Gewichts in der Publikation von Goldstein waren unvollständig und konnten für eine Berechnung nicht herangezogen werden, allerdings zeigen die absoluten Zahlen kein Unterschied zu Placebo, also eine Gewichtneutralität. In der Publikation von DeFronzo wie auch für die Studie D1690C00005 wurden die Veränderungen des Gewichts detailliert berichtet, sodass diese Daten für den indirekten Vergleich herangezogen werden konnten. Es zeigte sich eine deutliche und signifikant größere Gewichtsreduktion unter Sulfonylharnstoff+Dapagliflozin als unter Sulfonylharnstoff+Metformin (MWD [95%-KI]: -2,55 kg [-3,41;-1,69]).

Blutdrucksenkung:

Die Daten zur Veränderung des Blutdrucks waren nur aus dem Studienbericht der DeFronzo Studie, sowie der Studie D1690C00005 zu entnehmen, sodass diese Daten für den indirekten Vergleich herangezogen werden konnten. Es fand sich unter Sulfonylharnstoff+Dapagliflozin eine signifikant höhere systolische Blutdrucksenkung als unter Sulfonylharnstoff+Metformin (MWD [95%-KI]: -4,84 mmHg [-7,66;-2,02]). Hinsichtlich des diastolischen Blutdrucks war der Effektschätzer zwar ebenfalls niedriger, erreichte aber keine statistische Signifikanz.

Unerwünschte Ereignisse:

Der indirekte Vergleich für die Gesamtrate unerwünschter Ereignisse wurde aufgrund von Hinweisen auf Heterogenität für beide Studien (DeFronzo und Goldstein) separat durchgeführt. Es ergaben sich keine Unterschiede zwischen den Behandlungsoptionen Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin, weder beim indirekten Vergleich mit der Studie von DeFronzo (OR [95%-KI]: 0,65 [0,31;1,37]), noch beim Vergleich mit der Studie von Goldstein (OR [95%-KI]: 1,35 [0,60;3,03]). Die Gesamtrate schwerwiegender unerwünschter Ereignisse wurde zwar in der Publikation von Goldstein berichtet, es fanden jedoch in keiner Behandlungsgruppe Ereignisse statt, so dass sich kein Effektmaß berechnen ließ und somit auch kein indirekter Vergleich durchführbar war. Die Gesamtrate schwerwiegender unerwünschter Ereignisse zwischen den Behandlungsoptionen Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin, war beim indirekten Vergleich mit DeFronzo Daten nicht unterschiedlich (OR [95%-KI]: 2,50 [0,58;10,82]). Studienabbrüche aufgrund von unerwünschten Ereignissen wurden in beiden Publikationen und in der Studie mit dem zu bewertenden Arzneimittel berichtet. Es ergaben sich keine

signifikanten Unterschiede zwischen Sulfonylharnstoff+Dapagliflozin und Sulfonylharnstoff+Metformin (OR [95%-KI] des indirekten Vergleiches mit der Studie von DeFronzo: 0,56 [0,04;8,76] und von Goldstein 0,11 [0,01;1,74]).

Ergebnisse des indirekten Vergleichs Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor

Ausmaß der Blutzuckerkontrolle:

Für den Endpunkt „Ausmaß der Blutzuckerkontrolle“ standen alle 5 Studien des Vergleiches Sulfonylharnstoff+Placebo versus Sulfonylharnstoff+DPP-4-Inhibitor zur Verfügung. Die Meta-Analyse der Studien wies auf eine moderate Heterogenität der Studienergebnisse hin ($I^2=39\%$; $p=0.16$), die Größenordnung der Behandlungseffekte war aber in allen Studien ähnlich. Der indirekte Vergleich ergab keine statistisch signifikanten Unterschiede zwischen der zu bewertenden Therapie und Sulfonylharnstoff +DPP-4-Inhibitoren als Gruppe (MWD [95%-KI]: -0,12% [-0,33;0,09]) oder Sulfonylharnstoff +Sitagliptin als Einzelsubstanz (MWD [95%-KI]: -0,17% [-0,48;0,14]).

Hypoglykämien:

Die Gesamtrate der Hypoglykämien wurde in allen für den indirekten Vergleich als relevant identifizierten Studien berichtet. Es zeigten sich keine Hinweise auf bedeutsame Heterogenität der Ergebnisse ($I^2=0\%$; $p=0.59$), sodass die Zusammenfassung der Ergebnisse in einen gemeinsamen Effektschätzer sinnvoll erschien. Es fanden sich im indirekten Vergleich keine Unterschiede zwischen der zu bewertenden Therapie und Sulfonylharnstoff+DPP-4-Inhibitoren als Gruppe (OR [95%-KI]: 1,65 [0,52;5,18]) oder Sulfonylharnstoff+Sitagliptin als Einzelsubstanz (OR [95%-KI]: 0,80 [0,14;4,52]). Für den Endpunkt „Gesamtrate der Patienten mit mindestens einer schwerwiegenden Hypoglykämie“ war kein indirekter Vergleich möglich, da lediglich in der Studie von Pratley et al. eine schwerwiegende Hypoglykämie berichtet wurde (in der Gruppe Sulfonylharnstoff+Placebo) und daher keine Berechnung von Effektschätzern möglich war.

Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

Es fanden sich für den indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor weder Unterschiede im Ausmaß der Blutzuckerkontrolle, noch in der Häufigkeit von Hypoglykämien. Insgesamt ergab sich bei der gemeinsamen Betrachtung der Blutzuckerkontrolle und der Hypoglykämien kein Vorteil für eine der beiden Behandlungsoptionen.

Gewichtsveränderungen:

Für den Endpunkt „Gewichtsveränderung“ standen alle identifizierten DPP-4-Studien des Vergleiches Sulfonylharnstoff+Placebo versus Sulfonylharnstoff+DPP-4-Inhibitor zur Verfügung, für zwei der Studien (Chacra 2009; Lewin 2012) mussten die Standardabweichung des Behandlungseffektes aber geschätzt werden. Alle Studien wiesen in Bezug auf Veränderungen des Körpergewichts auf einen Nachteil von Sulfonylharnstoff+DPP-4-Inhibitoren gegenüber der Kontrollgruppe

Sulfonylharnstoff+Placebo hin. Es gab keine Hinweise auf bedeutsame Heterogenität zwischen den DPP-4-Inhibitor-Studien ($I^2=0\%$; $p=0.64$), die Meta-analytische Zusammenfassung der Ergebnisse war daher gerechtfertigt. Im indirekten Vergleich zeigte sich ein statistisch signifikanter Vorteil von der zu bewertenden Therapie im Vergleich zu Sulfonylharnstoff+DPP-4-Inhibitoren als Gruppe (MWD [95%-KI]: -2,36 kg [-3,14;-1,58]) oder Sulfonylharnstoff+Sitagliptin als Einzelsubstanz (MWD [95%-KI]: -2,82 kg [-3,98;-1,66]). Eine Sensitivitätsanalyse unter Ausschluss der Studien, für die die Standardabweichung des Behandlungseffektes geschätzt werden musste (Chacra 2009; Lewin 2012) ergab sehr ähnliche Ergebnisse wie die Hauptanalyse.

Blutdrucksenkung:

Die vorliegenden Daten zu den Sulfonylharnstoff+DPP-4-Inhibitor-Studien ließen keinen indirekten Vergleich in Bezug auf den Endpunkt Blutdrucksenkung zu.

Unerwünschte Ereignisse:

Bezüglich des Endpunktes „**Gesamtrate unerwünschter Ereignisse**“ fanden sich keine Unterschiede zwischen der zu bewertenden Therapie und Sulfonylharnstoff+DPP-4-Inhibitoren als Gruppe (OR [95%-KI]: 0,94 [0,52;1,69]) oder Sulfonylharnstoff+Sitagliptin als Einzelsubstanz (OR [95%-KI]: 0,59 [0,28;1,24]). Auch die **Gesamtrate schwerwiegender unerwünschter Ereignisse** unterschied sich nicht zwischen Dapagliflozin und Sulfonylharnstoff+DPP-4-Inhibitoren als Gruppe (OR [95%-KI]: 1,14 [0,31;4,16]) oder Sulfonylharnstoff+Sitagliptin als Einzelsubstanz (OR [95%-KI]: 1,51 [0,28;8,21]). Für den Endpunkt „**Studienabbrüche aufgrund unerwünschter Ereignisse**“ fanden sich ebenfalls keine statistisch signifikanten Unterschiede gegenüber Sulfonylharnstoff+DPP-4-Inhibitoren als Gruppe (OR [95%-KI]: 0,36 [0,03;4,58]) oder Sulfonylharnstoff+Sitagliptin als Einzelsubstanz (OR [95%-KI]: 0,14 [0,01;3,96]).

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 und sekundären Fragestellung lagen keine direkt vergleichenden Studien vor, daher wurden adjustierte indirekte Vergleiche Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin bzw. Sulfonylharnstoff+DPP-4-Inhibitoren durchgeführt.

Die in den adjustierten indirekten Vergleichen eingeschlossenen Studien wiesen alle ein niedriges Verzerrungspotenzials auf Studienebene auf und hatten aufgrund ihres Designs (doppelblinde, randomisierte, kontrollierte Studien) eine hohe Evidenzstufe. Die betrachteten Endpunkte sind in ihren Operationalisierungen valide, wiesen in den Studien ein geringes endpunktspezifisches Verzerrungspotential auf und stellen, wie in Abschnitt 4.2.5.2

begründet, patientenrelevante Endpunkte dar. Die Begründung für die Darstellung von Surrogatparametern ist in Abschnitt 4.5.4 gegeben. Die Validität der herangezogenen Endpunkte ist somit insgesamt als hoch einzuschätzen.

Einschränkungen in Bezug auf die Aussagekraft der Nachweise ergaben sich aber durch eine zum Teil hohe Heterogenität der Studien und der Studienergebnisse, sowie durch die Tatsache dass indirekte Vergleiche stets anfällig für Verzerrungen durch systematische Unterschiede in den einzelnen eingeschlossenen Studien sind. Für alle diskutierten Endpunkte können deswegen, im Falle statistisch signifikanter Ergebnisse in den adjustierten indirekten Vergleichen, Aussagen gemacht werden, die maximal als Hinweis gewertet werden können.

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, 2011a). 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.

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.

Die hier betrachtete Patientenpopulation schließt erwachsene Patienten mit manifestem Diabetes mellitus Typ 2 ein, die unter einer Sulfonylharnstoff-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen. Um den Blutzucker wieder in ausreichendem Maße kontrollieren zu können, wird bei diesen Patienten noch vor der Umstellung auf eine Insulintherapie in der Regel die Kombinationstherapie mit zwei verschiedenen oralen Antidiabetika (OAD) als Zweitlinientherapie bzw. OAD+Inkretin-Mimetika empfohlen. Als OAD werden Alpha-Glucosidasehemmer (Acarbose), DPP-4-Inhibitoren (Gliptine), Glitazone sowie Metformin genannt, soweit keine Kontraindikationen oder Unverträglichkeiten gegen Metformin bestehen (Matthaei et al., 2009).

Glitazone sind aufgrund erhöhten Risikos für Frakturen und kardiovaskuläre Ereignisse in Deutschland nicht mehr zu Lasten der Gesetzlichen Krankenversicherung verordnungsfähig (Verordnungsausschluss der Glitazone). Alpha-Glucosidasehemmer 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). DPP-4-Inhibitoren bewirken die Stimulation der Insulinsekretion und sind damit ebenso wie die Sulfonylharnstoffe abhängig von der Insulinverfügbarkeit (Matthaei et al., 2009). Schließlich kann aber auch die Kombination von Sulfonylharnstoff und Metformin wegen des unklaren Sicherheitsrisikos bezüglich kardiovaskulärer Endpunkte als kritisch angesehen werden (Matthaei et al., 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 und das einzige 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 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, 2012b). Das Risiko steigt etwas im Vergleich zu Placebo in der Kombinationstherapie mit Sulfonylharnstoff bzw. Insulin, so wie das auch für Metformin in Kombination mit diesen Stoffen bekannt ist (Sandoz Pharmaceuticals GmbH, 2009; Bristol-Myers Squibb und AstraZeneca GmbH, 2012b).

Obwohl der Schwerpunkt dieses Abschnitts auf der Beschreibung des Zusatznutzens liegt, soll der Nutzen von Sulfonylharnstoff+Dapagliflozin ebenfalls dargestellt werden.

Die Darstellung des medizinischen Nutzens von Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Placebo erfolgte mittels der Methodik der evidenzbasierten Medizin und mittels patientenrelevanter Endpunkte. Dapagliflozin wurde im Rahmen der Erteilung der Zulassung bereits ein positives Nutzen-Risiko-Verhältnis auf Basis der in den Studien prädefinierten primären Endpunkte bestätigt (European Medicines Agency, 2012).

Im Dossier betrachtete patientenrelevante Endpunkte sind das Auftreten von Hypoglykämien, Gewichtsveränderungen, Therapiezufriedenheit, unerwünschte Ereignisse sowie der Endpunkt „kardiale Morbidität und Mortalität“. Darüber hinaus wurden Angaben zum HbA1c-Wert als Maß für die Blutzuckerkontrolle extrahiert und dargestellt. Diese Angaben dienen zur Interpretation der Ergebnisse hinsichtlich patientenrelevanter Endpunkte, insbesondere bezüglich des Auftretens von Hypoglykämien.

Die Änderungen im systolischen und diastolischen Blutdruck werden als Surrogatparameter für das kardiovaskuläre Risiko berücksichtigt.

Die in der folgenden Beschreibung des Nutzens und des Zusatznutzens betrachteten Ergebnisse beziehen sich auf die für Dapagliflozin empfohlene Dosis von 10 mg und die Zielpopulation gemäß Zulassung (Typ-2-Diabetiker, die unter einer Sulfonylharnstoff-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, Kreatininclarence <60 mL/min und Patienten, die ≥ 75 Jahre alt sind bzw. mit Schleifendiuretika behandelt werden).

Nutzen von Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Placebo

Die Ergebnisse zum Nutzen des zu bewertenden Arzneimittels Dapagliflozin in Kombination mit Sulfonylharnstoff basieren auf einer RCT der Phase III (D1690C00005), die mit der freien Kombination von Glimepirid+Dapagliflozin durchgeführt wurde (Abschnitt 4.3.1).

Eine optimierte Blutzuckerkontrolle (und damit die HbA1c-Einstellung) wird von der 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 zusätzliche Gabe von Insulin 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). In der *United Kingdom Prospective Diabetes Study* (UKPDS), einer großen klinischen Studie, 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). Eine optimale Blutzuckereinstellung stellt damit ein wichtiges Therapieziel dar.

Für den Endpunkt „**Ausmaß der Blutzuckerkontrolle (HbA1c)**“ zeigte die Studie D1690C00005 nach 24 Wochen einen deutlichen statistisch signifikanten Unterschied zugunsten von Sulfonylharnstoff+Dapagliflozin (MWD [95%-KI]: -0,74% [-0,92;-0,55]), welcher auch noch nach 48 Wochen bestand (MWD [95%-KI]: -0,77% [-1,02;-0,52]). Studienabbrüche aufgrund einer unzureichenden glykämischen Kontrolle, oder Patienten, die aufgrund einer unzureichenden glykämischen Kontrolle eine Notfallmedikation erhielten, gab es in der Sulfonylharnstoff+Placebo-Gruppe sowohl nach 24 als auch nach 48 Wochen signifikant häufiger als in der Gruppe, die mit Sulfonylharnstoff+Dapagliflozin behandelt wurde (OR [95%-KI]: 0,07 [0,02;0,30] / 0,18 [0,10;0,33]). Hinsichtlich der Gesamtrate der **Hypoglykämien** (Anzahl der Patienten mit mindestens einer Hypoglykämie) gab es sowohl nach 24 als auch nach 48 Wochen keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen.

Hinsichtlich der **Gewichtsveränderungen** zeigt sich eine signifikant stärkere Gewichtsabnahme unter der Sulfonylharnstoff+Dapagliflozin-Behandlung sowohl nach 24 Wochen (MWD [95%-KI]: -1,72 kg [-2,43;-1,01]) als auch nach 48 Wochen (MWD [95%-KI]: -1,80 kg [-2,80;-0,80]). Weiterhin erreichen sowohl nach 24 Wochen (OR [95%-KI]: 7,42 [2,79;19,78]) als auch nach 48 Wochen (OR [95%-KI]: 3,47 [1,35;8,93]) signifikant mehr Patienten aus der Sulfonylharnstoff+Dapagliflozin-Gruppe eine Gewichtsabnahme $\geq 5\%$. Innerhalb eines halben Jahres konnte 25,5% der Patienten in der Sulfonylharnstoff+Dapagliflozin 10 mg Gruppe eine **Gewichtsreduktion von $\geq 5\%$** erreichen.

Die **Therapiezufriedenheit** der Patienten wurde mittels DTSQ gemessen. Dieser kann Werte von 0-36 annehmen, wobei hohe Werte eine höhere Therapiezufriedenheit bedeuten. Bereits zu Beginn hatten die Patienten in beiden Gruppen mit etwa 27-29 Punkten hohe Ausgangswerte. Nach 24 Wochen stieg die erzielte Punktzahl leicht an (+3,6 in der

Sulfonylharnstoff+Dapagliflozin Gruppe/+2,2 in der Sulfonylharnstoff+Placebo-Gruppe) d. h. beide Behandlungsgruppen konnten sich leicht verbessern. Während die Therapiezufriedenheit in der Sulfonylharnstoff+Placebo-Gruppe nach 48 Wochen relativ konstant blieb (+2,4 Punkte), stieg die Punktzahl im DTSQ bei den mit Sulfonylharnstoff+Dapagliflozin behandelten Patienten um 4,5 Punkte deutlich, aber nicht signifikant an.

Die **Blutdrucksenkung** war in der Sulfonylharnstoff+Dapagliflozin-Gruppe höher als in der Sulfonylharnstoff+Placebo-Gruppe. Der systolische Blutdruck konnte sowohl nach 24 Wochen als auch nach 48 Wochen signifikant stärker in der Sulfonylharnstoff+Dapagliflozin-Gruppe (-5,5 mmHg/-4,11 mmHg) als in der Sulfonylharnstoff+Placebo-Gruppe (-1,1 mmHg/+0,35 mmHg) gesenkt werden (MWD [95%-KI]: -4,4 mmHg [-7,2;-1,6] nach 24 Wochen/-4,46 mmHg [-8,62;-0,31] nach 48 Wochen). Auch der diastolische Blutdruck konnte sowohl nach 24 als auch nach 48 Wochen signifikant stärker gesenkt werden (MWD [95%-KI]: -2,0 mmHg [-3,8;-0,2] nach 24 Wochen/-3,12 mmHg [-5,68;-0,56] nach 48 Wochen).

Im Hinblick auf die unerwünschten Ereignisse gab es weder für die **Gesamtrate der unerwünschten Ereignisse** noch für die Rate der **schwerwiegenden unerwünschten Ereignisse** und der **Studienabbrüche wegen unerwünschter Ereignisse** statistisch signifikante Unterschiede. Auch für die Anzahl der kardialen Ereignisse war kein Unterschied zwischen den beiden Behandlungsgruppen zu sehen. Allgemein war das Auftreten dieser Ereignisse sehr selten. Ergänzend hierzu wurden die Anzeichen auf Genital- oder Harnwegsinfektionen als unerwünschte Ereignisse von besonderem Interesse betrachtet. Für die Zielpopulation konnte zu keinem Zeitpunkt ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (Sulfonylharnstoff+Dapagliflozin gegen Sulfonylharnstoff+Placebo) beobachtet werden. Darüber hinaus waren die Ereignisse nur mild bis moderat in ihrer Intensität, traten in den meisten Fällen nur einmal auf, und waren und mit den gängigen Therapieoptionen einfach zu behandeln.

Für keinen der betrachteten Endpunkte gab es einen Hinweis auf eine fazitrelevante Modifikation des Behandlungseffektes von Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Placebo durch die Faktoren Alter, Geschlecht oder Abstammung. Daher kann man von einem über alle Gruppen hinweg konstanten Behandlungseffekt ausgehen.

Zusammenfassend kann man sagen, dass die Behandlung mit Dapagliflozin bei Patienten, deren Blutzucker mit Sulfonylharnstoffen alleine nicht mehr ausreichend kontrolliert werden kann, sicher und wirksam ist. Die Kombination von Sulfonylharnstoff+Dapagliflozin senkt den Blutzucker effektiv ohne die Gesamtrate an Hypoglykämien statistisch signifikant zu erhöhen. Darüber hinaus kommt es zu einer signifikanten Gewichtsreduktion und der Blutdruck wird gesenkt. Die Gesamtrate an unerwünschten Ereignissen war in beiden Behandlungsgruppen ähnlich.

Zusatznutzen von Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Metformin bzw. Sulfonylharnstoff+DPP-4-Inhibitor

Im Folgenden werden die Ergebnisse zum medizinischen Zusatznutzen genauer betrachtet. Sie basieren auf adjustierten indirekten Vergleichen zwischen Dapagliflozin in Kombination mit einem Sulfonylharnstoff (AstraZeneca, 2010a) und der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin bzw. der alternativen Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor bei Vorliegen einer Metformin-Unverträglichkeit.

Die systematischen Literaturrecherchen identifizierten zwei relevante Studien mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin (DeFronzo und Goodman, 1995; Goldstein et al., 2003), welche für den indirekten Vergleich relevant waren. In den Studien wurden die Sulfonylharnstoffe Glipizid (Goldstein et al., 2003) und Glibenclamid (DeFronzo und Goodman, 1995) verwendet. Da die von DeFronzo et al. berichtete Studie zwecks Weiterentwicklung von Glukophage durchgeführt wurde, lag Bristol-Myers Squibb der Studienbericht vor. Aufgrund der klinischen Heterogenität zwischen den einzelnen Studien, sowie der Heterogenität der Einzelschätzer konnte kein Gesamtschätzer berechnet werden. Die Betrachtung erfolgte daher getrennt für jede Einzelstudie. Für den indirekten Vergleich mit der alternativen Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor (bei Vorliegen einer Metformin-Unverträglichkeit) wurden insgesamt 5 relevante Studien identifiziert, jeweils eine mit einem unterschiedlichen DPP-4-Inhibitor (Sitagliptin, Saxagliptin, Vildagliptin, Alogliptin und Linagliptin) in Kombination mit einem Sulfonylharnstoff. In jeweils zwei der Studien wurde Glimепirid (Hermansen et al. 2007, Garber et al. 2008) bzw. Glibenclamid (Chacra et al. 2009 und Pratley et al. 2009) als Sulfonylharnstoff benutzt. Die Studie von Lewin et al. 2012 verwendete diverse Sulfonylharnstoffe als Basistherapie. Alle fünf Studien waren placebokontrolliert.

Gemeinsame Betrachtung der Blutzuckerkontrolle und Auftreten von Hypoglykämien

Die Vermeidung von Hypoglykämien ist ein wichtiges Therapieziel bei Typ-2-Diabetikern. Es konnte allerdings in mehreren Studien gezeigt werden, dass eine intensiverte Therapie zur Blutzuckerkontrolle ein höheres Risiko für das Erleiden schwerwiegender Hypoglykämien zur Folge haben kann (UK Prospective Diabetes Study Group, 1998c; Patel et al., 2008; Duckworth et al., 2009; Ismail-Beigi et al., 2010). Dabei gilt es zu berücksichtigen, dass insbesondere eine Sulfonylharnstoff-Therapie ein hohes Risiko für das Auftreten von Hypoglykämien birgt, was auf den Wirkmechanismus der Sulfonylharnstoffe zurückzuführen ist. Dieser beruht auf einer Stimulation der Insulinsekretion der Betazellen des Pankreas durch Blockade der ATP-abhängigen Kaliumkanäle (AkdÄ, 2009). Darüber hinaus ist die Kombination von Sulfonylharnstoff und Metformin wegen des unklaren Sicherheitsrisikos bezüglich kardiovaskulärer Endpunkte kritisch zu sehen (AkdÄ, 2009). Auch die Fachinformation von Metformin empfiehlt die Kombinationsbehandlung mit Sulfonylharnstoff nur unter Vorsicht und warnt vor vermehrter Auftreten von Hypoglykämien in der Kombination mit einem Sulfonylharnstoff (Sandoz Pharmaceuticals GmbH, 2009). Dies gilt auch für eine Kombination eines DPP-4-Inhibitors mit einem Sulfonylharnstoff (Bristol-Myers Squibb/AstraZeneca, 2011; Merck Sharp & Dohme (Italia) S.p.A., 2011; Novartis Europharm Limited, 2011b; Novartis Europharm Limited, 2011a). Eine

blutzuckersenkende Therapie, die das Auftreten von Hypoglykämien weitgehend vermeidet, stellt somit ein wichtiges Ziel der Diabetestherapie dar.

Der indirekte Vergleich zeigte, dass das „**Ausmaß der Blutzuckerkontrolle (HbA1c)**“ bei Behandlung mit Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Metformin sehr ähnlich ist. Beide Behandlungen konnten den HbA1c-Wert effektiv senken. Der Vergleich mit der Studie von DeFronzo et al. zeigte eine signifikant stärkere Blutzuckerkontrolle mit der zweckmäßigen Vergleichstherapie (MWD [95%-KI]: 1,14 [0,79;1,49]). Die Ergebnisse müssen aber hinterfragt werden, da der Ausgangs-HbA1c-Wert in der Studie von DeFronzo et al. deutlich höher war und *Carry-over*-Effekte aufgrund der hohen eingesetzten Sulfonylharnstoff-Dosis nicht auszuschließen sind. Für das Auftreten von Hypoglykämien (Gesamtrate) konnte ein indirekter Vergleich mit der Studie von DeFronzo et al. durchgeführt werden. Beim Auftreten von Hypoglykämien zeigte sich kein signifikanter Unterschied zwischen den Behandlungen. In der Studienpopulation konnte hier sogar ein Vorteil gesehen werden, dieser stellte sich in der Zielpopulation aber nicht mehr statistisch signifikant dar.

Der indirekte Vergleich mit Sulfonylharnstoff+Dapagliflozin ergab weder für den Endpunkt Blutzuckerkontrolle noch **Hypoglykämien** (Gesamtrate) statistisch signifikante Unterschiede zwischen der zu bewertenden Therapie und Sulfonylharnstoff+einem DPP-4-Inhibitor als Gruppe (HbA1c MWD [95%-KI]: -0,12 [-0,33;0,09]; OR für Hypoglykämien [95%-KI]: 1.65 [0,52;5,18]) bzw. Sulfonylharnstoff+Sitagliptin (HbA1c MWD [95%-KI]: -0,17 [-0,48;0,14]; OR für Hypoglykämien [95%-KI]: 0,80 [0,14;4,52]).

Gemessen an dem Ziel, eine adäquate Blutzuckerkontrolle bei gleichzeitiger Hypoglykämievermeidung zu erreichen, ergibt sich kein Hinweis auf eine Über- oder Unterlegenheit weder für Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+Metformin noch für Sulfonylharnstoff+Dapagliflozin gegenüber Sulfonylharnstoff+ DPP-4-Inhibitor. Alle Behandlungen senken den Blutzucker effektiv ohne die Gesamtrate an Hypoglykämien zu erhöhen.

Betrachtung der Gewichtsveränderung

Die Adipositas (Fettleibigkeit), im Besonderen die viszerale/abdominale Adipositas, ist einer der wichtigsten Risikofaktoren zur Entwicklung eines Diabetes mellitus Typ 2 (Freemantle et al., 2008). Die Adipositas führt zusammen mit der Bewegungsarmut und Dyslipoproteinämie zur Insulinresistenz (unzureichende Insulinwirksamkeit), welche essentiell zur Entstehung eines Diabetes mellitus Typ 2 beiträgt (AkdÄ, 2009). 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 (Porzsolt 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 oder Glitazone (AkdÄ, 2009), weshalb auch in diesem Punkt ein klarer therapeutischer Bedarf besteht. Das

metabolische Syndrom¹³, dessen wichtigster Einzelfaktor die Adipositas darstellt, ist weiterhin wesentlich für das hohe kardiovaskuläre bzw. Makroangiopathierisiko verantwortlich, 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 (Perk et al., 2012) und es scheint, dass eine Verringerung der Fettmasse vor diesen Risikofaktoren schützen kann (Vega et al., 2006).

Der indirekte Vergleich versus Sulfonylharnstoff+Metformin zeigte, dass die Patienten, die mit Sulfonylharnstoff+Dapagliflozin behandelt wurden, eine signifikant höhere Gewichtsreduktion haben als Patienten, die mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin behandelt wurden (MWD [95%-KI] des indirekten Vergleiches mit DeFronzo: -2,55 kg [-3,41;-1,69]). Ein adjustierter indirekter Vergleich mit der Studie von Goldstein et al. war nicht möglich. Die Ergebnisse der Studie von Goldstein zeigten jedoch, dass kein Unterschied zwischen der Behandlungs- und Placebogruppe bestand, während die Patienten aus der Studie D1690C00005, die mit Sulfonylharnstoff+Dapagliflozin behandelt wurden, im Vergleich zur Sulfonylharnstoff+Placebo-Gruppe signifikant mehr Gewicht verloren. Der indirekte Vergleich versus einem Sulfonylharnstoff+einem DPP-4-Inhibitor im Endpunkt Gewichtsveränderung ergab einen statistisch signifikanten Unterschied zwischen Sulfonylharnstoff+Dapagliflozin und Sulfonylharnstoff+DPP-4-Inhibitor als Gruppe (MWD [95%-KI]: -2,36 [-3,14;-1,58]) wie auch der Kombination Sulfonylharnstoff+Sitagliptin (MWD [95%-KI]: -2,82 [-3,98;-1,66]).

Die signifikant größere Gewichtsreduktion unter Sulfonylharnstoff+Dapagliflozin trägt somit zu einer erfolgreichen Diabetestherapie bei. Damit liegt eine Verringerung nicht schwerwiegender Symptome bzw. eine relevante Vermeidung von Nebenwirkungen vor, sodass sich ein Hinweis auf einen geringen Zusatznutzen für die Behandlung mit Sulfonylharnstoff+Dapagliflozin sowohl im Vergleich zur zweckmäßigen wie auch zur alternativen Vergleichstherapie ergibt.

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, 2011b). Darüber hinaus existiert ein linearer Zusammenhang zwischen der Höhe des Blutdruckes und dem Auftreten kardio- 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,

¹³ 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).

1998b; 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% 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; Matthaei et al., 2009).

Der indirekte Vergleich versus Sulfonylharnstoff+Metformin zeigte, dass die Patienten, die mit Sulfonylharnstoff+Dapagliflozin behandelt wurden, eine signifikant stärkere Senkung des systolischen Blutdrucks haben als Patienten, die mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin behandelt wurden (MWD [95%-KI]: -4,84 [-7,66;-2,02]). Die Absenkung des diastolischen Blutdrucks war ebenfalls stärker im Sulfonylharnstoff+Dapagliflozin-Arm, war hier aber statistisch nicht signifikant.

Der indirekte Vergleich versus einem Sulfonylharnstoff+einem DPP-4-Inhibitor im Endpunkt Blutdrucksenkung konnte aufgrund mangelnder Daten für die Vergleichstherapie nicht durchgeführt werden.

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 Sulfonylharnstoff +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 und stellt eine Verringerung nicht schwerwiegender Symptome dar, sodass sich damit insgesamt einen Hinweis auf einen geringen Zusatznutzen ergibt.

Betrachtung der unerwünschten Ereignisse

Der indirekte Vergleich für die Gesamtrate der unerwünschten Ereignisse konnte mit beiden Studien zur zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Metformin durchgeführt werden wie auch mit allen fünf Studien zu Sulfonylharnstoff+einem DPP-4-Inhibitor. Es ergaben sich keine signifikanten Unterschiede zwischen Sulfonylharnstoff+Dapagliflozin und weder der zweckmäßigen noch der alternativen Vergleichstherapie. Der indirekte Vergleich für die Gesamtrate schwerwiegender unerwünschter Ereignisse wie auch Studienabbrüche aufgrund von unerwünschten Ereignissen ergaben ebenfalls keine signifikanten Unterschiede, weder für den Vergleich Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+Metformin noch Sulfonylharnstoff+Dapagliflozin versus Sulfonylharnstoff+DPP-4-Inhibitor

Damit kann gezeigt werden, dass die Behandlung mit Sulfonylharnstoff+Dapagliflozin im Vergleich zur Behandlung mit Sulfonylharnstoff+Metformin bzw. Sulfonylharnstoff+DPP-4-Inhibitor in den Sicherheitsendpunkten weder vor- noch nachteilhaft ist.

Auch wenn für den Endpunkt „kardiale Ereignisse“ keine Daten aus dem indirekten Vergleich vorliegen, gilt es ebenfalls zu berücksichtigen, dass beiden deutschen Therapieleitlinien die Kombination von Sulfonylharnstoff und Metformin wegen des unklaren Nutzen-Risiko-

Verhältnisses bezüglich kardiovaskulärer Endpunkte nur mit Einschränkungen empfehlen (AkdÄ, 2009; Matthaai et al., 2009). Weiterhin wird in beiden Leitlinien angemerkt, dass speziell die Kombination aus Metformin und Glibenclamid nur dann verabreicht werden soll, wenn äquivalent wirksame Alternativen nicht eingesetzt werden können und wenn der Patient eingehend über das mögliche Risiko sowie alternative Therapieoptionen aufgeklärt wurde.

Gesamtbetrachtung des Zusatznutzens

Die Ergebnisse zum medizinischen Nutzen zeigen, dass die Kombination von Dapagliflozin mit Sulfonylharnstoffen wirksam und sicher ist.

Die Ergebnisse zum Zusatznutzen zeigen, dass es für die Endpunkte „Ausmaß der Blutzuckerkontrolle“, „Hypoglykämien“, und „unerwünschte Ereignisse“ keinen Hinweis auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen (Sulfonylharnstoff+Dapagliflozin gegen Sulfonylharnstoff+Metformin bzw. Sulfonylharnstoff+DPP-4-Inhibitoren) gibt.

Für den Endpunkt „Gewichtsveränderungen“ konnte gezeigt werden, dass die Behandlung mit Sulfonylharnstoff+Dapagliflozin im Vergleich zu Sulfonylharnstoff+Metformin und Sulfonylharnstoff+DPP-4-Inhibitoren zu einer signifikant größeren Gewichtsreduktion führt. Im Vergleich zur zweckmäßigen Vergleichstherapie konnte auch der Blutdruck signifikant gesenkt werden, was beides auf den neuartigen Wirkmechanismus von Dapagliflozin zurückzuführen ist. Dabei gilt es zu berücksichtigen, dass der Gewichtsverlust vor allem auf die Reduktion von Körperfettmasse – inklusive viszerale Fettmasse – zurückzuführen ist (Bolinder et al., 2012), welche zusammen mit einer Hypertonie, der wichtigste Risikofaktor für die Entwicklung und das Fortschreiten der Insulinresistenz ist. Gemessen an dem Ziel Gewichtszunahmen zu vermeiden und unter Berücksichtigung der Tatsache, dass eine Verminderung des Körpergewichts für die Patienten eine ähnlich wichtige Rolle spielt wie die Blutzuckereinstellung (Porzolt et al., 2010), ergibt sich sowohl gegenüber Sulfonylharnstoff+Metformin als auch gegenüber Sulfonylharnstoff+DPP-4-Inhibitoren insgesamt ein Hinweis auf einen geringen Zusatznutzen von Sulfonylharnstoff+Dapagliflozin.

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 4.4.1 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 4-136: 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 Sulfonylharnstoff-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 oder Kreatinin-clearance <60 mL/min und von Patienten, die ≥75 Jahre sind und Patienten, die mit Schleifendiuretika behandelt werden.	Geringer Zusatznutzen
Erwachsene Patienten mit Diabetes mellitus Typ 2, die unter einer Sulfonylharnstoff-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 oder Kreatinin-clearance <60 mL/min und von Patienten, die ≥75 Jahre sind und Patienten, die mit Schleifendiuretika behandelt werden und für die Metformin nicht geeignet ist und bei denen die Anwendung von Insulin noch nicht angezeigt ist (Teilpopulation).*	Geringer Zusatznutzen

* Gegenüber der alternativen Vergleichstherapie: Sulfonylharnstoff+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 4-137: 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.

Es liegen keine direkten Vergleichsstudien zwischen dem zu bewertenden Arzneimittel (Sulfonylharnstoff+Dapagliflozin) und der zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+ Metformin) bzw. der alternativen Vergleichstherapie (Sulfonylharnstoff+DPP-4-Inhibitor) vor.

Mit Sulfonylharnstoff+Placebo als Brückenkomparator wurde deshalb ein indirekter Vergleich mit der zweckmäßigen (bzw. alternativen) Vergleichstherapie durchgeführt. Für den indirekten Vergleich wurden ausschließlich RCT verwendet, da diese mit der geringsten Ergebnisunsicherheit behaftet sind.

Der durchgeführte adjustierte indirekte Vergleich erfüllt damit den momentanen Stand der Wissenschaft hinsichtlich der Anforderungen an eine solche Analyse und 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 2005¹⁴, Molenberghs 2010¹⁵). 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 2006¹⁶) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und 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 2006¹⁷) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

¹⁴ Burzykowski T (Ed.: *The evaluation of surrogate endpoints*. New York: Springer; 2005.

¹⁵ 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.

¹⁶ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

¹⁷ Weir CJ, Walley RJ. *Statistical evaluation of biomarkers as surrogate endpoints: a literature review*. *Stat Med* 2006; 25(2): 183-203.

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 wurde 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 sich daraus abgeleiteten Effekte auf das kardiovaskuläre Risiko der Patientinnen und 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 dem Auftreten kardio- 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 60-80% 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% Wahrscheinlichkeit nach 10

Jahren mit einem kardiovaskulären Ereignis zu rechnen ist. Das Risiko ist bei Patienten mit zusätzlicher Nephropathie und insbesondere Mikroalbuminurie dabei nochmals erhöht (Bundesärztekammer et al., 2011).

Die Makroangiopathie ist nach Anlage 1 der RSA-ÄndV in Form der koronaren Herzkrankheit das Hauptproblem der Patientinnen / Patienten mit Diabetes mellitus Typ 2. Die Senkung eines erhöhten Blutdrucks reduziert die kardio- und zerebrovaskuläre Morbidität und Mortalität innerhalb weniger Jahre. Das *Disease Management Programm* (DMP) empfiehlt daher eine geeignete Risikoabschätzung hinsichtlich makroangiopathischer Komplikationen (IQWiG, 2011b).

Das IQWiG stellt in der Zusammenfassung der deutschen, europäischen und außereuropäischen Leitlinien die Bedeutung der Risikoreduktion speziell in Bezug auf den Bluthochdruck bei Patienten / Patientinnen 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	Titel	Im Dossier verwendete Datenquellen
Chacra 2009	<i>Saxagliptin added to a submaximal dose of sulphonylurea improves glycaemic control compared with uptitration of sulphonylurea in patients with type 2 diabetes: a randomised controlled trial</i>	(Chacra et al., 2009)
D1690C00005	<i>A 24-Week, International, Randomized, Double-blind, Parallel-group, Multi-center, Placebo-controlled Phase III Study with a 24-Week Extension Period to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Glimpiride (a Sulphonylurea) in Subjects with Type 2 Diabetes who Have Inadequate Glycemic Control on Glimpiride Therapy Alone</i>	(AstraZeneca, 2010a) (AstraZeneca, 2010b)
DeFronzo 1995	<i>Efficacy of metformin in patients with non-insulin-dependent diabetes mellitus</i>	(DeFronzo und Goodman, 1995) (Lipha Pharmaceutical Inc., 1992)
Garber 2008	<i>Effects of vildagliptin on glucose control in patients with type 2 diabetes inadequately controlled with a sulphonylurea</i>	(Garber et al., 2008)
Goldstein 2003	<i>Multicenter, Randomized, Double-Masked, Parallel-Group Assessment of Simultaneous Glipizide/Metformin as Second-Line Pharmacologic Treatment for Patients with Type 2 Diabetes Mellitus That Is Inadequately Controlled by a Sulfonylurea</i>	(Goldstein et al., 2003)
Hermansen 2007	<i>Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, in patients with type 2 diabetes mellitus inadequately controlled on glimepiride alone or on glimepiride and metformin</i>	(Hermansen et al., 2007)
Lewin 2012	<i>Efficacy and tolerability of linagliptin added to a sulfonylurea regimen in patients with inadequately controlled type 2 diabetes mellitus: An 18-week, multicentre, randomized, double-blind, placebo-controlled trial</i>	(Lewin et al., 2012)
Pratley 2009	<i>Efficacy and safety of the dipeptidyl peptidase-4 inhibitor alogliptin in patients with type 2 diabetes inadequately controlled by glyburide monotherapy</i>	(Pratley et al., 2009)

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] AstraZeneca. 2010a. A 24-Week, International, Randomized, Double-blind, Parallel-group, Multi-center, Placebo-controlled Phase III Study with a 24-Week Extension Period to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Glimepiride (a Sulphonylurea) in Subjects with Type 2 Diabetes who Have Inadequate Glycemic Control on Glimepiride Therapy Alone. Report for the first 24-week treatment period.
- [4] AstraZeneca. 2010b. A 24-Week, International, Randomized, Double-blind, Parallel-group, Multi-center, Placebo-controlled Phase III Study with a 24-Week Extension Period to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Glimepiride (a Sulphonylurea) in Subjects with Type 2 Diabetes who Have Inadequate Glycemic Control on Glimepiride Therapy Alone Report for the 24-week short-term treatment period plus the 24-week longterm extension period.
- [5] 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*.
- [6] 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.
- [7] 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.
- [8] Bristol-Myers Squibb Company. 2010. Characterization of the Kinetics of Renal Glucose Reabsorption in Response to Dapagliflozin in Healthy Subjects and in Subjects with Type 2 Diabetes Mellitus.
- [9] Bristol-Myers Squibb und AstraZeneca GmbH 2012a. Post-hoc Analysen der einzelnen Endpunkte und Subgruppen für die Zielpopulation von Dapagliflozin (Studie D1690C00005).
- [10] Bristol-Myers Squibb und AstraZeneca GmbH. 2012b. Fachinformation Forxiga (Dapagliflozin).

- [11] Bristol-Myers Squibb/AstraZeneca. 2011. Fachinformation Onglyza.
- [12] Brooks, A. M. & Thacker, S. M. 2009. Dapagliflozin for the treatment of type 2 diabetes. *The Annals of pharmacotherapy*, 43, 1286-93.
- [13] Bundesärztekammer, KBV & AWMF. 2011. Nationale Versorgungsleitlinie: Nierenerkrankungen bei Diabetes im Erwachsenenalter – Version 1.3: Langfassung.
- [14] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV) & Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) 2012. Nationale VersorgungsLeitlinie Therapieplanung bei Typ-2-Diabetes. 1.0. Konsultationsfassung.
- [15] Chacra, A. R., Tan, G. H., Apanovitch, A., et al. 2009. Saxagliptin added to a submaximal dose of sulphonylurea improves glycaemic control compared with up-titration of sulphonylurea in patients with type 2 diabetes: a randomised controlled trial. *International journal of clinical practice*, 63, 1395-406.
- [16] Chacra, A. R., Tan, G. H., Ravichandran, S., et al. 2011. Safety and efficacy of saxagliptin in combination with submaximal sulphonylurea versus up-titrated sulphonylurea over 76 weeks. *Diabetes & vascular disease research : official journal of the International Society of Diabetes and Vascular Disease*, 8, 150-9.
- [17] Clinicaltrials.gov. 2012a. *Randomized, Double-blind (db), Placebo-controlled 18 Week Study of Linagliptin (BI 1356) in Type 2 Diabetic Patients With Insufficient Glycaemic Control on a Sulfonylurea Drug* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00819091> [Aufgerufen am 30.10.2012].
- [18] Clinicaltrials.gov. 2012b. *Study of Alogliptin Combined With Sulfonylurea in Subjects With Type 2 Diabetes Mellitus* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00286468> [Aufgerufen am 30.10.2012].
- [19] Clinicaltrials.gov. 2012c. *Efficacy and Safety of Vildagliptin in Combination With Glimepiride in Patients With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00099944> [Aufgerufen am 30.10.2012].
- [20] Clinicaltrials.gov. 2012d. *A Study of Saxagliptin in Subjects With Type 2 Diabetes Who Have Inadequate Blood Sugar Control With Sulfonylureas* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00313313> [Aufgerufen am 30.10.2012].
- [21] Clinicaltrials.gov. 2012e. *Sulfonylurea Add-on Study in Patients With Type 2 Diabetes Mellitus (0431-035)* [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT00106704> [Aufgerufen am 30.10.2012].
- [22] Clinicaltrials.gov. 2012f. *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].
- [23] Clinicaltrials.gov. 2012g. *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].
- [24] Clinicaltrials.gov. 2012h. *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].

- [25] Clinicaltrials.gov. 2012i. *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].
- [26] Clinicaltrials.gov. 2012j. *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].
- [27] Clinicaltrials.gov. 2012k. *Renal Mechanism of Action/Splay vs. TmG (MOA)* [Online]. Adresse: <http://clinicaltrials.gov/show/NCT00726505> [Aufgerufen am 12.11.2012].
- [28] Clinicaltrials.gov. 2012l. *Efficacy and Safety of Dapagliflozin in Combination With Glimpiride (a Sulphonylurea) in Type 2 Diabetes Patients* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT00680745> [Aufgerufen am 30.10.2012].
- [29] DeFronzo, R. A. & Goodman, A. M. 1995. Efficacy of metformin in patients with non-insulin-dependent diabetes mellitus. The Multicenter Metformin Study Group. *The New England journal of medicine*, 333, 541-9.
- [30] 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.
- [31] Duckworth, W., Abraira, C., Moritz, T., et al. 2009. Glucose control and vascular complications in veterans with type 2 diabetes. *The New England journal of medicine*, 360, 129-39.
- [32] European Medicines Agency. 2012. European Public Assessment Report (EPAR) Dapagliflozin.
- [33] FDA 2008. Guidance for Industry: Diabetes Mellitus: Developing Drugs and Therapeutic Biologics for Treatment and Prevention (DRAFT GUIDANCE).
- [34] 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.
- [35] Garber, A. J., Foley, J. E., Banerji, M. A., et al. 2008. Effects of vildagliptin on glucose control in patients with type 2 diabetes inadequately controlled with a sulphonylurea. *Diabetes, obesity & metabolism*, 10, 1047-56.
- [36] Glanville, J. M., Lefebvre, C., Miles, J. N., et al. 2006. How to identify randomized controlled trials in MEDLINE: ten years on. *Journal of the Medical Library Association : JMLA*, 94, 130-6.
- [37] Glenny, A. M., Altman, D. G., Song, F., et al. 2005. Indirect comparisons of competing interventions. *Health Technol Assess*, 9, 1-134, iii-iv.
- [38] Goldstein, B. J., Pans, M. & Rubin, C. J. 2003. Multicenter, randomized, double-masked, parallel-group assessment of simultaneous glipizide/metformin as second-line pharmacologic treatment for patients with type 2 diabetes mellitus that is inadequately controlled by a sulphonylurea. *Clinical therapeutics*, 25, 890-903.

- [39] 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.
- [40] Hein 2012. Niederschrift (finale Fassung) zum Beratungsgespräch gemäß § 8 AM-NutzenV, Beratungsanforderung 2011-B-35, Dapagliflozin zur Behandlung des Diabetes mellitus Typ 2
- [41] Hermansen, K., Kipnes, M., Luo, E., et al. 2007. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, in patients with type 2 diabetes mellitus inadequately controlled on glimepiride alone or on glimepiride and metformin. *Diabetes, obesity & metabolism*, 9, 733-45.
- [42] Higgins, J. P., Thompson, S. G., Deeks, J. J., et al. 2003. Measuring inconsistency in meta-analyses. *BMJ*, 327, 557-60.
- [43] Higgins, J. P. & Green, S. 2008. *Cochrane Handbook for Systematic Reviews of Interventions*. , The Cochrane Collaboration.
- [44] ICTRP Search Portal. 2012a. *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].
- [45] ICTRP Search Portal. 2012b. *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].
- [46] ICTRP Search Portal. 2012c. *A 24-Week Randomised, Double-blind, Parallel-group, Multi-centre, Placebo-controlled Phase III Study to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Glimepiride (a Sulphonylurea) in Subjects with Type 2 Diabetes who Have Inadequate Glycaemic Control on Glimepiride Therapy Alone* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00680745> [Aufgerufen am 30.10.2012].
- [47] ICTRP Search Portal. 2012d. *A 24-Week Randomised, Double-blind, Parallel-group, Multi-centre, Placebo-controlled Phase III Study to Evaluate the Efficacy and Safety of Dapagliflozin in Combination with Glimepiride (a Sulphonylurea) in Subjects with Type 2 Diabetes who Have Inadequate Glycaemic Control on Glimepiride Therapy Alone* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2007-005931-27-HU> [Aufgerufen am 30.10.2012].
- [48] IQWiG 2009a. Glinide zur Behandlung des Diabetes mellitus Typ 2 - Abschlussbericht A05-05C. Version 1.0.
- [49] IQWiG. 2009b. Urin- und Blutzuckerselbstmessung bei Diabetes mellitus Typ 2
- [50] IQWiG. 2011a. Allgemeine Methoden Version 4.0 vom 23.09.2011. Verfügbar: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf.
- [51] IQWiG 2011b. Systematische Leitlinienrecherche und -bewertung sowie Extraktion neuer und relevanter Empfehlungen für das DMP Diabetes mellitus Typ 2: Vorbericht.

- [52] 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.
- [53] 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).
- [54] Lewin, A. J., Arvey, L., Liu, D., et al. 2012. Efficacy and tolerability of linagliptin added to a sulfonylurea regimen in patients with inadequately controlled type 2 diabetes mellitus: an 18-week, multicenter, randomized, double-blind, placebo-controlled trial. *Clinical therapeutics*, 34, 1909-1919.
- [55] Lipha Pharmaceutical Inc. 1992. Final Report Study No. 87-2D-6023.
- [56] Mancía, G. 2007. Blood pressure reduction and cardiovascular outcomes: past, present, and future. *The American journal of cardiology*, 100, 3J-9J.
- [57] Matthaëi, S., Bierwirth, R., Fritsche, A., et al. 2009. Medikamentöse antihyperglykämische Therapie des Diabetes mellitus Typ 2. *Diabetologie* 2009, 4, 32-64.
- [58] Mengel K. 2011. Antidiabetika. In: SCHWABE U. & PFAFFRATH D. (Hrsg.) *Arzneiverordnungsreport 2011*. Springer.
- [59] Merck Sharp & Dohme (Italia) S.p.A. 2011. Fachinformation Xelevia.
- [60] Novartis Europharm Limited. 2011a. Fachinformation Jalra.
- [61] Novartis Europharm Limited. 2011b. Fachinformation Galvus.
- [62] 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.
- [63] 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*.
- [64] Porzsolt, 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.
- [65] Pratley, R. E., Kipnes, M. S., Fleck, P. R., et al. 2009. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor alogliptin in patients with type 2 diabetes inadequately controlled by glyburide monotherapy. *Diabetes, obesity & metabolism*, 11, 167-76.
- [66] Sandoz Pharmaceuticals GmbH. 2009. Fachinformation Metformin.
- [67] 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.

- [68] Standl, E., Fuchs, C. H., Parandeh-Shab, F., et al. 2000. Management der Hypertonie beim Patienten mit Diabetes mellitus.
- [69] 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.
- [70] Strojek, K., Yoon, K. H., Hrubá, V., et al. 2011. Effect of dapagliflozin in patients with type 2 diabetes who have inadequate glycaemic control with glimepiride: a randomized, 24-week, double-blind, placebo-controlled trial. *Diabetes, obesity & metabolism*, 13, 928-38.
- [71] UK Prospective Diabetes Study Group 1998a. 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.
- [72] UK Prospective Diabetes Study Group 1998b. 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.
- [73] 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.
- [74] 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.
- [75] 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.
- [76] Wong, S. S., Wilczynski, N. L. & Haynes, R. B. 2006. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *Journal of the Medical Library Association : JMLA*, 94, 451-5.

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 ¹⁸] – 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

¹⁸ 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 Bewertung des Zusatznutzens von Dapagliflozin wurden zwei indirekte Vergleiche durchgeführt (Abschnitt 4.3.2.1). Dazu wurde eine Literaturrecherche für das Anwendungsgebiet *Add-on* zu Sulfonylharnstoff unter der Fragestellung, wie in Abschnitt 4.2.3.2 beschrieben, durchgeführt.

Indirekter Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin

In der Studie des pharmazeutischen Unternehmens D1690C00005 wurde das zu bewertende Arzneimittel Dapagliflozin in Kombination mit dem Sulfonylharnstoff Glimепirid verabreicht und mit der Kombinationsbehandlung Glimепirid und Placebo verglichen. Die Suchstrategien sind in Tabelle 4-138, Tabelle 4-139 und Tabelle 4-140 dargestellt. Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Relevante Studien aus dieser systematischen Literaturrecherche bilden zusammen mit den in Studienregistern identifizierten und vom pharmazeutischen Unternehmer durchgeführten Studien die Datengrundlage für die Bewertung des medizinischen Nutzens und Zusatznutzens von Dapagliflozin im Abschnitt 4.4.2 ab. Die Suche wurde für Abschnitt 4.3.2.1.1.2 durchgeführt.

Datenbankname	Medline Direct
Suchoberfläche	DIMDI
Datum der Suche	23.10.2012
Zeitsegment	1950 bis 2012 Woche 43
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- <i>Strategy minimizing difference between sensitivity and specificity</i> (Wong et al., 2006) (angepasst für DIMDI)

Tabelle 4-138 (Anhang): DIMDI: Suchstrategie MEDLINE 1950 bis heute: am 23. Oktober 2012

#	Suchbegriffe	Ergebnis
1	ME60	22204568
2	FT=diabetes mellitus OR (CT D "diabetes mellitus" OR UT="diabetes mellitus" OR IT="diabetes mellitus" OR SH="diabetes mellitus")	318910
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 02"; "DIABETES MELLITUS, NON INSULIN DEPENDENT"; "DIABETES MELLITUS, NONINSULIN DEPENDENT"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, TYPE II"; "DIABETES MELLITUS, TYPE 2")	153513
4	(CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM") OR FT=NIDDM	73996
5	(CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM") OR FT=T2DM	4350
6	2 OR 3 OR 4 OR 5	319385
7	(CT D "metformin" OR UT="metformin" OR IT="metformin" OR SH="metformin") OR FT=metformin	8569

#	Suchbegriffe	Ergebnis
8	(CTG D "glimepirid" OR UTG="glimepirid" OR ITG="glimepirid" OR SHG="glimepirid") OR FT= glimepirid	10
9	(CT D "glimepiride" OR UT="glimepiride" OR IT="glimepiride" OR SH="glimepiride") OR FT=glimepiride	773
10	(CT D "glimepirid"? OR UT="glimepirid"? OR IT="glimepirid"? OR SH="glimepirid?") OR FT=glimepirid?	777
11	8 OR 9 OR 10	777
12	7 AND 11	223
13	(CT D "glimepiride/Metformin" OR UT="glimepiride/Metformin" OR IT="glimepiride/Metformin" OR SH="glimepiride/Metformin") OR FT=glimepiride Metformin	27
14	(CT D "glimepiride-Metformin" OR UT="glimepiride-Metformin" OR IT="glimepiride-Metformin" OR SH="glimepiride-Metformin") OR FT=glimepiride-Metformin	27
15	13 OR 14	27
16	12 OR 15	223
17	6 AND 16	193
18	DT="RANDOMIZED CONTROLLED TRIAL"	332229
19	randomized/(Ti;Ab;Ct;Te) OR CT=Randomized?	393756
20	placebo/(Ti;Ab;Ct;TE) OR CT=placebo?	157764
21	18 OR 19 OR 20	620874
22	17 AND 21	96
23	(22) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	94

Datenbankname Embase

Suchoberfläche DIMDI

Datum der Suche 23.10.2012

Zeitsegment 1947 bis 2012 Woche 43

Suchfilter Filter für randomisierte kontrollierte Studien nach Wong 2006-
Strategy minimizing difference between sensitivity and specificity
(Wong et al., 2006) (angepasst für DIMDI)

Tabelle 4-139 (Anhang): DIMDI: Suchstrategie EMBASE 1947 bis heute: am 23. Oktober 2012

#	Suchbegriffe	Ergebnis
1	EM47	26087574
2	FT=diabetes mellitus OR (CT D "diabetes mellitus" OR UT="diabetes mellitus" OR IT="diabetes mellitus" OR SH="diabetes mellitus")	496286
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 2"; "DIABETES MELLITUS TYPE II"; "DIABETES MELLITUS, NONINSULIN-DEPENDENT"; "DIABETES MELLITUS, TYPE 2"; "DIABETES MELLITUS, TYPE II"; "DIABETES MELLITUS,NON INSULIN DEPENDENT")	364603
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	106986
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	4463
6	2 OR 3 OR 4 OR 5	496504
7	FT=metformin OR (CT D "metformin" OR UT="metformin" OR IT="metformin" OR SH="metformin")	25692
8	CT=("METFORMIN"; "METFORMIN HYDROCHLORIDE")	24809
9	7 OR 8	25692
10	FT=glimepiride OR (CT D "glimepiride" OR UT="glimepiride" OR IT="glimepiride" OR SH="glimepiride")	3678
11	FT=glimepirid? OR (CT D "glimepirid"? OR UT="glimepirid"? OR IT="glimepirid"? OR SH="glimepirid"?)	3681
12	10 OR 11	3681
13	CT="GLIMEPIRIDE PLUS METFORMIN"	32
14	FT=Glimepiride-Metformin OR (CT D "Glimepiride-Metformin" OR UT="Glimepiride-Metformin" OR IT="Glimepiride-Metformin" OR SH="Glimepiride-Metformin")	23
15	FT=Glimepiride Metformin OR (CT D "Glimepiride/Metformin" OR UT="Glimepiride/Metformin" OR IT="Glimepiride/Metformin" OR SH="Glimepiride/Metformin")	23
16	13 OR 14 OR 15	50
17	9 AND 12	2398
18	16 OR 17	2398
19	6 AND 18	2282
20	random?/(Ti;Ab)	674995
21	double-blind?/(Ti;Ab)	124799
22	placebo?/(Ti;Ab;Ct) OR CT=placebo?	278963
23	20 OR 21 OR 22	869920
24	19 AND 23	805
25	(24) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	784

#	Suchbegriffe	Ergebnis
26	DT="REVIEW"	1827278
27	25 NOT 26	380
28	27 NOT SU=Medline	375

Datenbankname Cochrane
Suchoberfläche DIMDI
Datum der Suche 23.10.2012
Zeitsegment 1948 bis 2012 Woche 43
Suchfilter Aufgrund der geringen Trefferzahl wurde kein Filter verwendet

Tabelle 4-140 (Anhang): DIMDI: Suchstrategie Cochrane 1948 bis heute: am 23. Oktober 2012

#	Suchbegriffe	Ergebnis
1	CCTR93	674332
2	FT=diabetes mellitus OR (CT D "diabetes mellitus" OR UT="diabetes mellitus" OR IT="diabetes mellitus" OR SH="diabetes mellitus")	15908
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 02"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, NON INSULIN DEPENDENT"; "DIABETES MELLITUS, NONINSULIN DEPENDENT"; "DIABETES MELLITUS, TYPE 2"; "DIABETES MELLITUS, TYPE II")	8805
4	(CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM") OR FT=NIDDM	6784
5	(CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM") OR FT=T2DM	365
6	2 OR 3 OR 4 OR 5	15992
7	(CT D "metformin" OR UT="metformin" OR IT="metformin" OR SH="metformin") OR FT=metformin	1654
8	(CT D "glimepirid" OR UT="glimepirid" OR IT="glimepirid" OR SH="glimepirid") OR FT=glimepirid	3
9	(CT D "glimepirid"? OR UT="glimepirid"? OR IT="glimepirid"? OR SH="glimepirid?") OR FT=glimepirid?	207
10	(CT D "glimepiride" OR UT="glimepiride" OR IT="glimepiride" OR SH="glimepiride") OR FT=glimepiride	206
11	8 OR 9 OR 10	207
12	(CT D "glimepiride-Metformin" OR UT="glimepiride-Metformin" OR IT="glimepiride-Metformin" OR SH="glimepiride-Metformin") OR FT=glimepiride-Metformin	15
13	(CT D "glimepiride Metformin" OR UT="glimepiride Metformin" OR IT="glimepiride Metformin" OR SH="glimepiride Metformin") OR	15

#	Suchbegriffe	Ergebnis
	FT=glimepiride Metformin	
14	12 OR 13	15
15	7 AND 11	85
16	14 OR 15	85
17	6 AND 16	78

Um die Evidenzlage zu erhöhen wurde die Literaturrecherche erweitert. Es wurde nach randomisierten, kontrollierten Studien mit den zweckmäßigen Vergleichstherapien Metformin in Kombination mit den Sulfonylharnstoffen Glibenclamid oder Glipizid unter der Fragestellung, wie in Abschnitt 4.2.3.2 beschrieben, gesucht. Die Suchstrategien sind in Tabelle 4-141, Tabelle 4-142 und Tabelle 4-143 dargestellt. Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Relevante Studien aus dieser systematischen Literaturrecherche bilden zusammen mit den in Studienregistern identifizierten und vom pharmazeutischen Unternehmer durchgeführten Studien die Datengrundlage für die Bewertung des medizinischen Nutzens und Zusatznutzens von Dapagliflozin im Abschnitt 4.4.2 ab. Die Suche wurde für Abschnitt 4.3.2.1.1.2 durchgeführt.

Datenbankname Medline Direct
Suchoberfläche DIMDI
Datum der Suche 23.10.2012
Zeitsegment 1950 bis 2012 Woche 43
Suchfilter Filter für randomisierte kontrollierte Studien nach Wong 2006-
 Strategy minimizing difference between sensitivity and specificity
 (Wong et al., 2006) (angepasst für DIMDI)

Tabelle 4-141 (Anhang): DIMDI: Suchstrategie MEDLINE 1950 bis heute: am 23. Oktober 2012

#	Suchbegriffe	Ergebnis
1	ME60	22204568
2	FT=diabetes mellitus OR (CT D "diabetes mellitus" OR UT="diabetes mellitus" OR IT="diabetes mellitus" OR SH="diabetes mellitus")	318910
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 02"; "DIABETES MELLITUS, NON INSULIN DEPENDENT"; "DIABETES MELLITUS, NONINSULIN DEPENDENT"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, TYPE II"; "DIABETES MELLITUS, TYPE 2")	153513
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	73996
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR	4350

#	Suchbegriffe	Ergebnis
	SH="T2DM")	
6	2 OR 3 OR 4 OR 5	319385
7	FT=Metformin OR (CT D "Metformin" OR UT="Metformin" OR IT="Metformin" OR SH="Metformin")	8569
8	FT=Glibenclamide OR (CT D "Glibenclamide" OR UT="Glibenclamide" OR IT="Glibenclamide" OR SH="Glibenclamide")	8302
9	FT=Glibenclamid? OR (CT D "Glibenclamid"? OR UT="Glibenclamid"? OR IT="Glibenclamid"? OR SHG="Glibenclamid"?)	8315
10	FT=Glibenclamid OR (CTG D "Glibenclamid" OR UTG="Glibenclamid" OR ITG="Glibenclamid" OR SHG="Glibenclamid")	5372
11	FT=Glyburide OR (CT D "Glyburide" OR UT="Glyburide" OR IT="Glyburide" OR SH="Glyburide")	5782
12	FT=Glyburid? OR (CT D "Glyburid"? OR UT="Glyburid"? OR IT="Glyburid"? OR SH="Glyburid"?)	5782
13	FT=Glyburid OR (CTG D "Glyburid" OR UTG="Glyburid" OR ITG="Glyburid" OR SHG="Glyburid")	5354
14	8 OR 9 OR 10 OR 11 OR 12 OR 13	8676
15	7 AND 14	549
16	FT=Glibenclamide Metformin OR (CT D "Glibenclamide Metformin" OR UT="Glibenclamide Metformin" OR IT="Glibenclamide Metformin" OR SH="Glibenclamide Metformin")	23
17	FT=Glibenclamide-Metformin OR (CT D "Glibenclamide-Metformin" OR UT="Glibenclamide-Metformin" OR IT="Glibenclamide-Metformin" OR SH="Glibenclamide-Metformin")	23
18	FT=Glyburide Metformin OR (CT D "Glyburide Metformin" OR UT="Glyburide Metformin" OR IT="Glyburide Metformin" OR SH="Glyburide Metformin")	31
19	FT=Glyburide-Metformin OR (CT D "Glyburide-Metformin" OR UT="Glyburide-Metformin" OR IT="Glyburide-Metformin" OR SH="Glyburide-Metformin")	31
20	16 OR 17 OR 18 OR 19	53
21	15 OR 20	549
22	FT=Glipizide OR (CT D "Glipizide" OR UT="Glipizide" OR IT="Glipizide" OR SH="Glipizide")	901
23	FT=Glipizid? OR (CT D "Glipizid"? OR UT="Glipizid"? OR IT="Glipizid"? OR SH="Glipizid"?)	906
24	FT=Glipizid OR (CTG D "Glipizid" OR UTG="Glipizid" OR ITG="Glipizid" OR SHG="Glipizid")	604
25	22 OR 23 OR 24	906
26	FT=Glipizide Metformin OR (CT D "Glipizide Metformin" OR UT="Glipizide Metformin" OR IT="Glipizide Metformin" OR SH="Glipizide Metformin")	6
27	FT=Glipizide-Metformin OR (CT D "Glipizide-Metformin" OR UT="Glipizide-Metformin" OR IT="Glipizide-Metformin" OR SH="Glipizide-Metformin")	6

#	Suchbegriffe	Ergebnis
28	26 OR 27	6
29	7 AND 25	118
30	28 OR 29	118
31	21 OR 30	618
32	6 AND 31	523
33	DT="Randomized controlled trial"	332229
34	Randomized/(Ti;Ab;Ct;TE) OR CT=Randomized?	393756
35	Placebo/(Ti;Ab;Ct;TE) OR CT=Placebo?	157764
36	33 OR 34 OR 35	620874
37	32 AND 36	202
38	37 AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	194

Datenbankname	Embase
Suchoberfläche	DIMDI
Datum der Suche	23.10.2012
Zeitsegment	1947 bis 2012 Woche 43
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- <i>Strategy minimizing difference between sensitivity and specificity</i> (Wong et al., 2006) (angepasst für DIMDI)

Tabelle 4-142 (Anhang): DIMDI: Suchstrategie EMBASE 1947 bis heute: am 23. Oktober 2012

#	Suchbegriffe	Ergebnis
1	EM47	26087574
2	FT=diabetes mellitus OR (CT D "diabetes mellitus" OR UT="diabetes mellitus" OR IT="diabetes mellitus" OR SH="diabetes mellitus")	496286
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 2"; "DIABETES MELLITUS TYPE II"; "DIABETES MELLITUS, NONINSULIN-DEPENDENT"; "DIABETES MELLITUS, TYPE 2"; "DIABETES MELLITUS, TYPE II"; "DIABETES MELLITUS,NON INSULIN DEPENDENT")	364603
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	106986
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	4463
6	2 OR 3 OR 4 OR 5	496504
7	FT=metformin OR (CT D "metformin" OR UT="metformin" OR IT="metformin" OR SH="metformin")	25692
8	CT=("METFORMIN"; "METFORMIN HYDROCHLORIDE")	24809
9	7 OR 8	25692
10	FT=glibenclamide OR (CT D "glibenclamide" OR UT="glibenclamide" OR	18845

#	Suchbegriffe	Ergebnis
	IT="glibenclamide" OR SH="glibenclamide")	
11	FT=glibenclamid? OR (CT D "glibenclamid"? OR UT="glibenclamid"? OR IT="glibenclamid"? OR SH="glibenclamid"?)	18856
12	FT=glyburide OR (CT D "glyburide" OR UT="glyburide" OR IT="glyburide" OR SH="glyburide")	18034
13	FT=glyburid? OR (CT D "glyburid"? OR UT="glyburid"? OR IT="glyburid"? OR SH="glyburid"?)	18166
14	10 OR 11 OR 12 OR 13	18994
15	CT="GLIBENCLAMID PLUS METFORMIN"	329
16	(CT D "glibenclamide-metformin" OR UT="glibenclamide-metformin" OR IT="glibenclamide-metformin" OR SH="glibenclamide-metformin") OR FT=glibenclamide-metformin	37
17	(CT D "glibenclamide/metformin" OR UT="glibenclamide/metformin" OR IT="glibenclamide/metformin" OR SH="glibenclamide/metformin") OR FT=glibenclamide metformin	37
18	CT=("GLYBURIDE AND METFORMIN HYDROCHLORIDE"; "GLYBURIDE PLUS METFORMIN"; "GLYBURIDE PLUS METFORMIN HYDROCHLORIDE")	329
19	FT=glyburide-metformin OR (CT D "glyburide-metformin" OR UT="glyburide-metformin" OR IT="glyburide-metformin" OR SH="glyburide-metformin")	36
20	FT=glyburide metformin OR (CT D "glyburide/metformin" OR UT="glyburide/metformin" OR IT="glyburide/metformin" OR SH="glyburide/metformin")	36
21	15 OR 16 OR 17 OR 18 OR 19 OR 20	378
22	9 AND 14	4499
23	21 OR 22	4499
24	FT=glipizide OR (CT D "glipizide" OR UT="glipizide" OR IT="glipizide" OR SH="glipizide")	4253
25	FT=glipizid? OR (CT D "glipizid"? OR UT="glipizid"? OR IT="glipizid"? OR SH="glipizid"?)	4258
26	24 OR 25	4258
27	CT=("GLIPIZIDE AND METFORMIN HYDROCHLORIDE"; "GLIPIZIDE PLUS METFORMIN"; "GLIPIZIDE PLUS METFORMIN HYDROCHLORIDE")	102
28	FT=glipizide-metformin OR (CT D "glipizide-metformin" OR UT="glipizide-metformin" OR IT="glipizide-metformin" OR SH="glipizide-metformin")	10
29	FT=glipizide metformin OR (CT D "glipizide/metformin" OR UT="glipizide/metformin" OR IT="glipizide/metformin" OR SH="glipizide/metformin")	10
30	27 OR 28 OR 29	110
31	9 AND 26	1995
32	30 OR 31	1995
33	23 OR 32	5093

#	Suchbegriffe	Ergebnis
34	6 AND 33	4631
35	random?/(Ti;Ab)	674995
36	double-blind?/(Ti;Ab)	124799
37	placebo?/(Ti;Ab;Ct) OR CT=placebo?	278963
38	35 OR 36 OR 37	869920
39	34 AND 38	1250
40	(39) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	1195
41	DT="REVIEW"	1827278
42	40 NOT 41	641
43	42 NOT SU=Medline	628

Datenbankname Cochrane
Suchoberfläche DIMDI
Datum der Suche 23.10.2012
Zeitsegment 1948 bis 2012 Woche 43
Suchfilter Aufgrund der geringen Trefferzahl wurde kein Filter verwendet

Tabelle 4-143 (Anhang): DIMDI: Suchstrategie Cochrane 1948 bis heute: am 23. Oktober 2012

#	Suchbegriffe	Ergebnis
1	CCTR93	674332
2	FT=Diabetes mellitus OR (CT D "Diabetes mellitus" OR UT="Diabetes mellitus" OR IT="Diabetes mellitus" OR SH="Diabetes mellitus")	15908
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 02"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, NON INSULIN DEPENDENT"; "DIABETES MELLITUS, NONINSULIN DEPENDENT"; "DIABETES MELLITUS, TYPE 2"; "DIABETES MELLITUS, TYPE II")	8805
4	(CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM") OR FT=NIDDM	6784
5	(CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM") OR FT=T2DM	365
6	2 OR 3 OR 4 OR 5	15992
7	FT=glibenclamid OR (CT D "glibenclamid" OR UT="glibenclamid" OR IT="glibenclamid" OR SH="glibenclamid")	37
8	FT=glibenclamide OR (CT D "glibenclamide" OR UT="glibenclamide" OR IT="glibenclamide" OR SH="glibenclamide")	687
9	FT=glibenclamid? OR (CT D "glibenclamid"? OR UT="glibenclamid"? OR IT="glibenclamid"? OR SH="glibenclamid"?)	692

10	FT=Glyburid OR (CTG D "Glyburid" OR UTG="Glyburid" OR ITG="Glyburid" OR SHG="Glyburid")	0
11	FT=Glyburide OR (CT D "Glyburide" OR UT="Glyburide" OR IT="Glyburide" OR SH="Glyburide")	513
12	FT=Glyburid? OR (CT D "Glyburid"? OR UT="Glyburid"? OR IT="Glyburid"? OR SH="Glyburid"?)	513
13	7 OR 8 OR 9 OR 10 OR 11 OR 12	726
14	FT=metformin OR (CT D "metformin" OR UT="metformin" OR IT="metformin" OR SH="metformin")	1654
17	FT=glyburide-metformin OR (CT D "glyburide-metformin" OR UT="glyburide-metformin" OR IT="glyburide-metformin" OR SH="glyburide-metformin")	16
18	FT=glyburide metformin OR (CT D "glyburide metformin" OR UT="glyburide metformin" OR IT="glyburide metformin" OR SH="glyburide metformin")	16
19	15 OR 16 OR 17 OR 18	25
20	13 AND 14	182
21	19 OR 20	182
22	FT=glipizide OR (CT D "glipizide" OR UT="glipizide" OR IT="glipizide" OR SH="glipizide")	192
23	FT=glipizid? OR (CT D "glipizid"? OR UT="glipizid"? OR IT="glipizid"? OR SH="glipizid"?)	192
24	FT=glipizid OR (CT D "glipizid" OR UT="glipizid" OR IT="glipizid" OR SH="glipizid")	0
25	FT=glipizide-metformin OR (CT D "glipizide-metformin" OR UT="glipizide-metformin" OR IT="glipizide-metformin" OR SH="glipizide-metformin")	4
26	FT=glipizide metformin OR (CT D "glipizide metformin" OR UT="glipizide metformin" OR IT="glipizide metformin" OR SH="glipizide metformin")	4
27	22 OR 23 OR 24	192
28	25 OR 26	4
29	14 AND 27	32
30	28 OR 29	32
31	21 OR 30	207
32	6 AND 31	195
33	32 AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	190

Indirekter Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4-Inhibitor

Die Suchstrategien sind in Tabelle 4-144, Tabelle 4-145 und Tabelle 4-146 dargestellt. Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Relevante Studien aus dieser systematischen Literaturrecherche bilden zusammen mit den in Studienregistern identifizierten und vom pharmazeutischen Unternehmer durchgeführten Studien die Datengrundlage für die Bewertung des medizinischen Nutzens und Zusatznutzens von Dapagliflozin im Abschnitt 4.4.2 ab. Die Suche wurde für Abschnitt 4.3.2.1.1.2 durchgeführt.

Datenbankname	Ovid MEDLINE(R)
Suchoberfläche	Ovid
Datum der Suche	17.10.2012
Zeitsegment	1946 bis 2012 Woche 42
Suchfilter	Filter für randomisierte kontrollierte Studien nach Glanville 2006 – Suchstrategie E und F (Glanville et al., 2006)

Tabelle 4-144 (Anhang): DIMDI: Suchstrategie MEDLINE 1946 bis heute: am 17. Oktober 2012

#	Suchbegriffe	Ergebnis
1	diabetes mellitus.af.	274126
2	NIDDM.af	6740
3	T2DM.af.	4303
4	1 or 2 or 3	274587
5	Dipeptidyl-Peptidase IV Inhibitor*.af.	1326
6	DPP-4 Inhibitor*.af.	77
7	Saxagliptin*.af.	182
8	Linagliptin*.af.	155
9	Sitagliptin*.af.	682
10	Vildagliptin*.af.	408
11	5 or 6 or 7 or 8 or 9 or 10	1793
12	glimepirid*.af.	784
13	glibenclamid*.af.	6574
14	glyburid*.af.	5866
15	sulphonylurea.af.	1298
16	sulfonylurea.af.	7228
17	12 or 13 or 14 or 15 or 16	14857
18	11 and 17	219

#	Suchbegriffe	Ergebnis
19	4 and 18	191
20	Clinical trial.pt.	475203
21	Randomized controlled trial.pt.	339719
22	20 or 21	600922
23	19 and 22	30
24	(ENG or GER).lg.	18344643
25	23 and 24	30
26	..l/ 25 hu=y	30

Datenbankname Embase Classic+Embase
Suchoberfläche Ovid
Datum der Suche 17.10.2012
Zeitsegment 1947 bis 16.10.2012
Suchfilter Filter für randomisierte kontrollierte Studien nach Wong 2006-
Strategy minimizing difference between sensitivity and specificity
(Wong et al., 2006)

Tabelle 4-145 (Anhang): DIMDI: Suchstrategie EMBASE 1947 bis heute: am 17. Oktober 2012

#	Suchbegriffe	Ergebnis
1	diabetes mellitus.af.	528458
2	NIDDM.af.	8063
3	T2DM.af.	7967
4	1 or 2 or 3	528850
5	Dipeptidyl-Peptidase IV Inhibitor*.af.	2481
6	DPP-4 Inhibitor*.af.	166
7	Saxagliptin*.af.	903
8	Linagliptin*.af.	371
9	Sitagliptin*.af.	2811
10	Vildagliptin*.af.	1620
11	5 or 6 or 7 or 8 or 9 or 10	4978
12	glimepirid*.af.	3931
13	glibenclamid*.af.	19697
14	glyburid*.af.	1678
15	sulphonylurea.af.	2090

#	Suchbegriffe	Ergebnis
16	sulfonylurea.af.	18519
17	12 or 13 or 14 or 15 or 16	35026
18	11 and 17	2042
19	4 and 18	1993
20	random*.ti. or random*.ab.	783029
21	double-blind*.ti. or double-blind*.ab.	141511
22	placebo*.ti. or placebo*.ab. or placebo*.sh.	304731
23	20 or 21 or 22	986478
24	19 and 23	860
25	(German or English).lg.	22791572
26	24 and 25	837
27	26 not Review.pt.	399
28	..1/ 27 hu=y	376

Datenbankname Cochrane
Suchoberfläche DIMDI
Datum der Suche 17.10.2012
Zeitsegment 1948 bis Oktober 2012
Suchfilter Aufgrund der geringen Trefferzahl wurde kein Filter verwendet

Tabelle 4-146 (Anhang): DIMDI: Suchstrategie Cochrane 1948 bis heute: am 17. Oktober 2012

#	Suchbegriffe	5 Ergebnis
1	CCTR93	673304
2	FT=diabetes mellitus OR (CT D "diabetes mellitus" OR UT="diabetes mellitus" OR IT="diabetes mellitus" OR SH="diabetes mellitus")	15874
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 02"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, NON INSULIN DEPENDENT"; "DIABETES MELLITUS, NONINSULIN DEPENDENT"; "DIABETES MELLITUS, TYPE 2"; "DIABETES MELLITUS, TYPE II")	8780
4	FT=NIDDM	930
5	FT=T2DM	363
6	2 OR 3 OR 4 OR 5	15958
7	FT=Saxagliptin? OR (CT D "Saxagliptin"? OR UT="Saxagliptin"? OR IT="Saxagliptin"? OR SH="Saxagliptin"?)	22
8	FT=Vildagliptin? OR (CT D "Vildagliptin"? OR UT="Vildagliptin"? OR	80

#	Suchbegriffe	5 Ergebnis
	IT="Vildagliptin"? OR SH="Vildagliptin"?)	
9	FT=Linagliptin? OR (CT D "Linagliptin"? OR UT="Linagliptin"? OR IT="Linagliptin"? OR SH="Linagliptin"?)	15
10	FT=Sitagliptin? OR (CT D "Sitagliptin"? OR UT="Sitagliptin"? OR IT="Sitagliptin"? OR SH="Sitagliptin"?)	78
11	7 or 8 or 9 or 10	194
12	12 FT=Glibenclamid? OR (CT D "Glibenclamid"? OR UT="Glibenclamid"? OR IT="Glibenclamid"? OR SH="Glibenclamid"?)	692
13	FT=Glyburid? OR (CT D "Glyburid"? OR UT="Glyburid"? OR IT="Glyburid"? OR SH="Glyburid"?)	513
14	FT=Glipizid? OR (CT D "Glipizid"? OR UT="Glipizid"? OR IT="Glipizid"? OR SH="Glipizid"?)	191
15	12 or 13 or 14	873
16	15 and 11	14
17	16 and 6	12

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.3 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 für das zu bewertende Arzneimittel

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	29.10.2012
Suchstrategie	Dapagliflozin OR BMS-512148 [ALL-Fields] 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 Diabetes mellitus 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 4.3.1.1.3 durchgeführt.

Studienregistersuche für die zweckmäßige Vergleichstherapie – Sulfonylharnstoff plus Metformin

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	24.10.2012
Suchstrategie	metformin AND (glimepiride OR glyburide OR Glibenclamid OR glipizide) [ALL-Fields] AND „Diabetes Mellitus“ [Condition] AND (Phase II OR Phase III OR Phase IV) [Phase]
Treffer	106

Die Suche nach registrierten Studien mit Metformin und Sulfonylharnstoffen zur Behandlung des Diabetes mellitus erfolgte anhand spezifischer Stichwörter. Die Datenbankabfrage über *ClinicalTrials.gov* ergab anhand der Suchanfrage 106 Treffer.

Da das Studienregister *clinicalstudyresults.org* nicht mehr zur Verfügung steht, wurde hier keine Suche vorgenommen.

Da die Webseite des WHO ICTRP nur sehr einfache Suchanfragen und kaum Verknüpfungen zwischen Suchbegriffen zulässt, musste für jede Kombination aus Sulfonylharnstoff und Metformin eine eigene Suche durchgeführt werden:

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	24.10.2012
Suchstrategie	metformin AND glimepiride [Intervention] AND „Diabetes Mellitus“ [Condition] AND „All“ [Recruitment Status]
Treffer	57

Die Datenbankabfrage über das ICTRP ergab anhand der Suchanfrage „metformin AND glimepiride“ 57 Treffer.

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	24.10.2012
Suchstrategie	metformin AND glibenclamide [Intervention] AND „Diabetes Mellitus“ [Condition] AND „All“ [Recruitment Status]
Treffer	16

Die Datenbankabfrage über das ICTRP ergab anhand der Suchanfrage „metformin AND Glibenclamide“ 16 Treffer.

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	24.10.2012
Suchstrategie	metformin AND glipizide [Intervention] AND „Diabetes Mellitus“ [Condition] AND „All“ [Recruitment Status]
Treffer	2

Die Datenbankabfrage über das ICTRP ergab anhand der Suchanfrage „metformin AND glipizide“ 2 Treffer.

Die Suche über die „Advanced Search“-Funktionen der zwei berücksichtigten Studienregister wurde für Abschnitt 4.3.2.1.1.3 durchgeführt.

Studienregistersuche für die alternative Vergleichstherapie – Sulfonylharnstoff plus DPP-4-Inhibitor

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	17.10.2012
Suchstrategie	(sitagliptin OR vildagliptin OR saxagliptin OR linagliptin OR alogliptin) AND (glimepiride OR glyburide OR Glibenclamid OR glipizide OR sulfonylurea) [ALL-Fields] AND „Diabetes Mellitus“ [Condition] AND (Phase II OR Phase III OR Phase IV) [Phase]
Treffer	71

Die Suche nach registrierten Studien mit DPP-4-Inhibitoren und Sulfonylharnstoffen zur Behandlung des Diabetes mellitus erfolgte anhand spezifischer Stichwörter. Die Datenbankabfrage über *ClinicalTrials.gov* ergab anhand der Suchanfrage 71 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	17.10.2012
Suchstrategie	(sitagliptin OR vildagliptin OR saxagliptin OR linagliptin OR alogliptin) AND (glimepiride OR glyburide OR Glibenclamid OR glipizide OR sulfonylurea) [Title] AND „Diabetes Mellitus“ [Condition] AND „All“ [Recruitment Status]
Treffer	210

Die Datenbankabfrage über das ICTRP ergab 210 Treffer (267 Datenbankeinträge für 210 Studien).

Die Suche über die „*Advanced Search*“-Funktionen der zwei berücksichtigten Studienregister wurde für Abschnitt 4.3.2.1.1.3 durchgeführt.

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 4-147 (Anhang): Im Volltext ausgeschlossene nicht relevante Publikationen

	Im Volltext ausgeschlossenen Publikation	Grund
Literatursuche für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+Metformin		
1	Charpentier G, Fleury F, Kabir M, Vaur L, Halimi S. Improved glycaemic control by addition of glimepiride to metformin monotherapy in Type 2 diabetic patients. <i>Diabetic Medicine</i> . 2001; 18: 828-834	In dieser Studie wurde die falsche Patientenpopulation untersucht, die Patienten waren mit Metformin-Monotherapie vorbehandelt
2	Bruce S, Park JS, Fiedorek FT, Howlett HC. Beta-cell response to metformin-Glibenclamid combination tablets (Glucovance) in patients with type 2 diabetes. <i>International journal of clinical practice</i> . 2006; 60(7): 783-90	In dieser Studie wurde die falsche Patientenpopulation untersucht (keine Vorbehandlung).
3	Marena S, Tagliaferro V, Montegrosso G, Pagano A, Scaglione L, Pagano G. Metabolic effects of metformin addition to chronic Glibenclamid treatment in type 2 diabetes. <i>Diabète & metabolism</i> . 1994; 20(1): 15-9	Die Studiendauer betrug <18 Wochen.
4	Tosi F, Muggeo M, Brun E, Spiazzi G, Perobelli L, Zanolin E, Gori M, Coppini A, Moghetti P. Combination treatment with metformin and Glibenclamid versus single-drug therapies in type 2 diabetes mellitus: a randomized, double-blind, comparative study. <i>Metabolism: clinical and experimental</i> . 2003; 52(7): 862-7	In dieser Studie wurde die falsche Patientenpopulation untersucht (keine Vorbehandlung).
5	Erle G, Lovise S, Stocchiero C, Lora L, Coppini A, Marchetti P, Merante D. A comparison of preconstituted, fixed combinations of low-dose glyburide plus metformin versus high-dose glyburide alone in the treatment of type 2 diabetic patients. <i>Acta diabetologica</i> . 1999 ; 36(1-2) : 61-5	In dieser Studie wurde die falsche Patientenpopulation untersucht (keine einheitliche Vorbehandlung).
6	Garber A, Marre M, Blonde L, Allavoine T, Howlett H, Lehert P, Cornes M.. Influence of initial hyperglycaemia, weight and age on the blood glucose lowering efficacy and incidence of hypoglycaemic symptoms with a single-tablet metformin-Glibenclamid therapy (Glucovance®) in type 2 diabetes. <i>Diabetes, Obesity and Metabolism</i> . 2003; 5(3): 171-179	In dieser Studie wurde die falsche Patientenpopulation untersucht (keine Vorbehandlung). Zudem befindet sich der Placebo-Arm nicht im selben Stratum wie Glibenclamid.

	Im Volltext ausgeschlossenen Publikation	Grund
7	Abbasi F, Kamath V, Rizvi AA, Carantoni M, Chen YD, Reaven GM. Results of a placebo-controlled study of the metabolic effects of the addition of metformin to sulfonylurea-treated patients. Evidence for a central role of adipose tissue. Diabetes care. 1997; 20(12): 1863-9	Die Endpunkte dieser Studie entsprechen nicht den Einschlusskriterien.
8	Gregorio F, Ambrosi F, Manfrini S, Verluzzi M, Carle F, Testa R, Merante D, Filliponi P. Poorly controlled elderly Type 2 diabetic patients: the effects of increasing sulphonylurea dosages or adding metformin. Diabetic medicine : a journal of the British Diabetic Association. 1999; 16(12): 1016-24	In dieser Studie wurde die falsche Patientenpopulation untersucht (Alter im Durchschnitt ≥ 75 und $HbA1c \geq 10$). Zudem enthält die Studie keinen Placebo-Arm.
9	Guillas-Herrero A, Aguilar-Salinas CA, Gómez-Pérez FJ, Rull JA. The combination metformin/glyburide exerts its hypoglycemic effect mainly by increasing insulin secretion: a controlled, randomized, double-blind, crossover study. Diabetes, nutrition & metabolism. 2003; 16(5-6) 268-76	In dieser Studie wurde die falsche Patientenpopulation untersucht (keine Vorbehandlung)
10	Herrmann LS, Bitzen PO, Kjellstron T, Lindgarde F, Schersten B. Comparative efficacy of metformin and Glibenclamid in patients with non-insulin-dependent diabetes mellitus. Diabete et Metabolisme. 1991; 17(1BIS): 201-208	In dieser Studie wurde die falsche Patientenpopulation untersucht (keine einheitliche Vorbehandlung)
11	Hermann, LS, Scherstén, B, Bitzén, PO, Kjellström, T, Lindgårde, F & Melander, A. Therapeutic comparison of metformin and sulfonylurea, alone and in various combinations. A double-blind controlled study. Diabetes care. 1994; 17, 1100-9.	In dieser Studie wurde die falsche Patientenpopulation untersucht (keine einheitliche Vorbehandlung)
12	Higginbotham L, Martin FIR. Double-blind trial of metformin in the therapy of non-ketotic diabetics. Medical Journal of Austria. 1997; 2:154-156	Die Studiendauer betrug <18 Wochen.
Literatursuche für indirekten Vergleich Sulfonylharnstoff+Dapagliflozin vs. Sulfonylharnstoff+DPP-4- Inhibitor		
1	Allen E, Fleming D, Berglind N, Donovan M, Ohman P. Long-term tolerability of saxagliptin as add-on therapy in type 2 diabetes (T2D): Pooled analysis. Diabetes 2011; 60:A298-A299.	Review / Meta-Analyse ohne neue Daten
2	Allen E, Karyekar C, Ohman P. Safety profile of saxagliptin (SAXA) in combination with 2 other agents: Data from dual-therapy trials in patients receiving rescue treatment. Diabetes 2011; 60:A619-A620.	Review / Meta-Analyse ohne neue Daten
3	Barnett A, Huisman H, Jones R, Von EM, Patel S,	Studiendesign

	Im Volltext ausgeschlossenen Publikation	Grund
	Woerle H-J. Efficacy and safety of linagliptin in elderly patients (≥ 70 years) with type 2 diabetes. Diabetes 2012;61:A260-A261.	(Behandlungsgruppen) nicht passend; keine Vollpublikation
4	Bellary S. For type 2 diabetes poorly controlled by metformin monotherapy, the addition of any non-insulin antidiabetic drug reduces HbA1c to a similar extent, but with differing effects on weight and hypoglycaemic risk. Evid -Based Med 2011; 16(2):39-40.	Review / Meta-Analyse ohne neue Daten
5	Busse FP, Stumvoll M. ADOPT: Evidence for early use of thiazolidinediones. Expert Rev Endocrinol Metab 2007; 2(2):121-124.	Review / Meta-Analyse ohne neue Daten
6	Chacra AR, Tan GH, Ravichandran S, List J, Chen R. Safety and efficacy of saxagliptin in combination with submaximal sulphonylurea versus up-titrated sulphonylurea over 76 weeks. Diabets Vasc Dis Res 2011; 8(2):150-159.	Die Studie von Chacra et al., 2011 berichtete nur Ergebnisse für den Zeitraum 76 Wochen und war daher für einen indirekten Vergleich ungeeignet.
7	Chien M-N, Lee C-C, Chen W-C, Liu S-C, Leung C-H, Wang C-H. Effect of sitagliptin as add-on therapy in elderly type 2 diabetes patients with inadequate glyceimic control in Taiwan. Int J Gerontol 2011; 5(2):103-106.	Studiendesign (Behandlungsgruppen) nicht passend
8	Del PS, Taskinen M-R, Owens D, et al. Efficacy and safety of linagliptin in patients with type 2 diabetes and poor glyceimic control. Diabetes 2011; 60:A293.	Review / Meta-Analyse ohne neue Daten
9	Derosa G, Maffioli P, Ferrari I, et al. Effects of one year treatment of vildagliptin added to pioglitazone or glimepiride in poorly controlled type 2 diabetic patients. Horm Metab Res 2010; 42(9):663-669.	Studiendesign (Behandlungsgruppen) nicht passend
10	Doucet J, Chacra A, Maheux P, Lu J, Harris S, Rosenstock J. Efficacy and safety of saxagliptin in older patients with type 2 diabetes mellitus. Curr Med Res Opin 2011; 27(4):863-869.	Review / Meta-Analyse ohne neue Daten
11	Doucet J, Maheux P, Allen E, Ravichandran S, Harris S, Chen R. Comparison of efficacy and safety of saxagliptin 5 mg once-daily therapy in young and old patients with type 2 diabetes mellitus. Fundam Clin Pharmacol 2011; 25:19.	Review / Meta-Analyse ohne neue Daten
12	Drucker D, Easley C, Kirkpatrick P. Sitagliptin. Nat Rev Drug Discov 2007; 6(2):109-110.	Review / Meta-Analyse ohne neue Daten
13	Fairman KA, Curtiss FR. Call for comparative effectiveness research: Lowering A1c with sitagliptin,	Keine RCT

	Im Volltext ausgeschlossenen Publikation	Grund
	saxagliptin, or cinnamon. J Managed Care Pharm 2009; 15(8):696-700.	
14	Freeman MK. Efficacy and safety of linagliptin (Tradjenta) in adults with type-2 diabetes mellitus. P T 2011; 36(12):807.	Review / Meta-Analyse ohne neue Daten
15	Gautier J-F, Sauvanet J-P. Efficacy of saxagliptin as an add-on to oral monotherapy in the phase 3 clinical development program: Predictive factors of the treatment response in type 2 diabetes. Ann Endocrinol 2011; 72(4):287-295.	Review / Meta-Analyse ohne neue Daten
16	Giugliano D, Esposito K. Adding noninsulin antidiabetic drugs to metformin therapy for type 2 diabetes. J Am Med Assoc 2010; 304(4):405.	Review / Meta-Analyse ohne neue Daten
17	Gomis R, Owens DR, Taskinen M-R, et al. Long-term safety and efficacy of linagliptin as monotherapy or in combination with other oral glucose-lowering agents in 2121 subjects with type 2 diabetes: Up to 2 years exposure in 24-week phase III trials followed by a 78-week open-label extension. Int J Clin Pract 201266(8):731-740.	Studiendesign (Behandlungsgruppen) nicht passend
18	Gysling E. Liraglutide. Pharma-Krit 2010; 32(1):3-4.	Review / Meta-Analyse ohne neue Daten
19	Iqbal N, Allen E, Donovan M, Ohman P. Long-term safety and tolerability of saxagliptin add-on therapy in elderly patients with type 2 diabetes. Diabetes 201261:A299.	Review / Meta-Analyse ohne neue Daten; keine Vollpublikation
20	Karagiannis T, Paschos P, Paletas K, Matthews DR, Tsapas A. Dipeptidyl peptidase-4 inhibitors for treatment of type 2 diabetes mellitus in the clinical setting: systematic review and meta-analysis. BMJ. 2012 Mar 12;344:e1369. doi: 10.1136/bmj.e1369.	Review / Meta-Analyse ohne neue Daten
21	Karyekar C, Frederich R, Donovan M, Ravichandran S. Achieving reductions in HbA1c $\geq 1\%$ without hypoglycemia with saxagliptin combination therapy: Post HOC analysis of 4 randomized controlled studies in patients with type 2 diabetes. Diabetes 2011; 60:A594-A595.	Review / Meta-Analyse ohne neue Daten
22	Kaufman. Rx report. Dial Transplant 2008; 37(1):32-33.	Review / Meta-Analyse ohne neue Daten
23	Kikuchi M, Haneda M, Koya D, et al. Efficacy and tolerability of vildagliptin as an add-on to glimepiride in Japanese patients with Type 2 diabetes mellitus. Diabetes	Beobachtungsdauer zu kurz

	Im Volltext ausgeschlossenen Publikation	Grund
	Res Clin Pract 2010; 89(3):216-223.	
24	Lukashevich V, Schweizer A, Shao Q, Groop P-H, Kothny W. Safety and efficacy of vildagliptin versus placebo in patients with type 2 diabetes and moderate or severe renal impairment: A prospective 24-week randomized placebo-controlled trial. Diabestes Obes Metab 2011; 13(10):947-954.	Studiendesign (Behandlungsgruppen) nicht passend
25	Meier JJ. Sitagliptin (Januvia). Gynakol Prax 2008; 32(2):361-365.	Review / Meta-Analyse ohne neue Daten
26	Morita S, Shimajiri Y, Kadoya Y, Furuta M, Yamada S, Sanke T. Enhancing effects of DPP-IV inhibitor on glycemic control depends on combined SU: A pilot cross-over study. Diabetes 2012; 61:A614.	Studiendesign (Behandlungsgruppen) nicht passend
27	Novel oral antidiabetic agents show promise. Br J Diabetes Vasc Dis 2005; 5(5):294.	Review / Meta-Analyse ohne neue Daten
28	Patel CG, Kornhauser D, Vachharajani N, et al. Saxagliptin, a potent, selective inhibitor of DPP-4, does not alter the pharmacokinetics of three oral antidiabetic drugs (metformin, glyburide or pioglitazone) in healthy subjects. Diabestes Obes Metab 2011; 13(7):604-614.	Falsche Population
29	Rendell M, Chrysant S, Emser A, et al. Linagliptin effectively reduces HbA1c independent of age in patients with type 2 diabetes. Pharmacotherapy 2011; 31(10):337e-338e.	Review / Meta-Analyse ohne neue Daten; keine Vollpublikation
30	Schwartz SL. Saxagliptin for the treatment of type 2 diabetes mellitus: Focus on recent studies. Ann Med 2012; 44(2):157-169.	Review / Meta-Analyse ohne neue Daten
31	Scott LJ. Linagliptin: In type 2 diabetes mellitus. Drugs 2011; 71(5):611-624.	Review / Meta-Analyse ohne neue Daten
32	Sitagliptin combined with sulphonylureas: Other treatments are preferable. Prescrire Int 2009; 18(99):14-15.	Review / Meta-Analyse ohne neue Daten
33	Tajima N, Kadowaki T, Odawara M, Nishii M, Taniguchi T, Ferreira JCA. Addition of sitagliptin to ongoing glimepiride therapy in Japanese patients with type 2 diabetes over 52 weeks leads to improved glycemic control. Diabetol Int 2011; 12(1):32-44.	Beobachtungsdauer zu kurz
34	Traynor K. Linagliptin approved for type 2 diabetes. Am J Health-Syst Pharm 2011; 68(12):1059.	Keine RCT
35	Vildagliptin and vildagliptin/metformin: Same limited efficacy as sitagliptin in type 2 diabetes. Prescrire Int	Review / Meta-Analyse ohne neue Daten

	Im Volltext ausgeschlossenen Publikation	Grund
	2008; 17(97):188.	
36	Vinik A. Too Many Notes: Up and Down the Scales of Diabetes Therapy. Clin Ther 2007; 29(6 part 1):1227-1235.	Review / Meta-Analyse ohne neue Daten
37	Von Eynatten M, Barnett AH, Patel S, Neubacher D, Woerle H-J. Efficacy and safety of linagliptin as add-on therapy to sulphonylurea (SU) in type 2 diabetes patients with moderate or severe renal impairment (RI). Diabetologia 2012; 55(3):S350-S351.	Review / Meta-Analyse ohne neue Daten
38	Wiedeman PE, Trevillyan JM. Dipeptidyl peptidase IV inhibitors for type 2 diabetes and metabolic syndrome. Drug Discov Today Ther Strateg 2005; 2(2):143-149.	Review / Meta-Analyse ohne neue Daten

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.

Nicht zutreffend

Tabelle 4-148 (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						
MB102 020	nein	ja	nein	ja (Bristol-Myers Squibb Company, 2010)	ja [NCT00726505] (Clinicaltrials.gov, 2012k)	nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
--						
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 4-149 (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, 2012j)	nein
MB102-073	nein	ja	nein	nein	ja [NCT01137474] (Clinicaltrials.gov, 2012i)	nein
MB102-077	nein	ja	nein	nein	ja [NCT01195662] (Clinicaltrials.gov, 2012h)	nein
D1690C 00018 (LT II)	nein	ja	nein	nein	ja [NCT01031680] (Clinicaltrials.gov, 2012g) (ICTRP Search Portal, 2012a)	nein
D1690C 00019 (LT II)	nein	ja	nein	nein	ja [NCT01042977] (Clinicaltrials.gov, 2012f) (ICTRP Search Portal, 2012b)	nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
--						

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])
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben. b: Studie, für die der Unternehmer Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist. 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 aber noch 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 4-150 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 4-150 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-150 (Anhang): Studiendesign und -methodik für Studie D1690C00005

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Primäres Ziel (nach 24 Wochen): Beurteilung der Wirksamkeit von Dapagliflozin versus Placebo als <i>Add-on</i>-Therapie zu Glimепirid (durch die Veränderung des HbA1c vom Ausgangswert bis Woche 24) bei Patienten mit Typ-2-Diabetes</p> <p>Sekundäre Ziele (nach 24 Wochen): Reduktion des Gewichtes durch Dapagliflozin im Gegensatz zu Placebo jeweils als <i>Add-on</i>-Therapie zu Glimепirid Reduktion des <i>2-host-challenge</i> Blutzuckerwertes durch Dapagliflozin+Glimепirid Darstellung eines höheren Anteils an Patienten mit glykämischer <i>Response</i> durch die Behandlung mit Dapagliflozin+Glimепirid Reduktion des Gewichtes durch Dapagliflozin im Gegensatz zu Placebo jeweils als <i>Add-on</i>-Therapie zu Glimепirid bei Patienten mit einem Ausgangswert BMI $\geq 27\text{kg/m}^2$ Reduktion in FPG mit Dapagliflozin+Glimепirid im Gegensatz zu Dapagliflozin+Placebo</p> <p>Andere sekundäre Endpunkte (nach 24 Wochen): Vergleich der Wirksamkeit von Dapagliflozin+Glimепirid mit Dapagliflozin+Placebo an zusätzlichen Gewicht-, glykämischen und Patientenberichteten Variablen Beurteilung der Sicherheit und Verträglichkeit von Dapagliflozin+Glimепirid im Vergleich zu Dapagliflozin+Placebo durch die Erhebung von unerwünschten Ereignissen, Laborwerten, EKG, Puls, Blutdruck, Hypoglykämien, Kreatininclearance und durch die Erhebung des Befunds einer körperlichen Untersuchung</p> <p>Explorative Endpunkte (nach 24 Wochen): Vergleich der Wirksamkeit von Dapagliflozin+Glimепirid mit Dapagliflozin+Placebo an explorativen Endpunkten</p> <p>Hauptziele (nach 48 Wochen): Beurteilung der Sicherheit und Verträglichkeit der Behandlung in einem 48-wöchigen Zeitraum Beurteilung der Beständigkeit der</p>

Item ^a	Charakteristikum	Studieninformation
		Wirksamkeit von jeder Dosis Dapagliflozin+Glimpirid versus Glimpirid (allein) in der 48-wöchigen Behandlungsdauer
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Internationale, multizentrische, randomisierte (1:1:1:1), parallele, placebokontrollierte und doppelblinde 24-wöchige Phase-III-Studie mit einer 24-wöchigen Verlängerungsphase
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 1. Visite (<i>Enrollment</i>):</p> <ul style="list-style-type: none"> - <i>Informed Consent</i> - Männer oder Frauen ≥ 18 Jahre - Frauen im gebärfähigem Alter müssen entsprechende Methoden zur Verhütung benutzen - Diagnose: Typ-2-Diabetes mellitus - Behandlung mit Glimpirid (Dosis nicht höher als 4 mg/Tag) für mindestens 8 Wochen vor Visite 1 <p>Einschlusskriterien bei 2. Visite (Laborwerte von 1. Visite):</p> <ul style="list-style-type: none"> - $HbA1c \geq 7,0\%$ und $\leq 10,0\%$ - $FPG \leq 270$ mg/dL - C-Peptid Level $\geq 1,0$ ng/mL <p>Patienten mit einer stabilen Dosis von 4 mg/Tag (Glimpirid) konnten die <i>Lead-in</i>-Phase und Qualifikationsphase überspringen:</p> <p>Einschlusskriterien bei 5. Visite (Randomisierungsphase):</p> <ul style="list-style-type: none"> - $HbA1c \geq 7,0\%$ und $\leq 10,0\%$ - $FPG \leq 270$ mg/dL - C-Peptid Level $\geq 1,0$ ng/mL <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> - Typ-1-Diabetes, Diabetes insipidus, kortikosteroid-induzierter Diabetes oder vorbelastet mit diabetischer Ketoazidose oder hyperosmolares nicht ketotisches Koma - Symptome eines schlecht eingestellten Diabetes mellitus - Höhere Dosis als 4mg/Tag von

Item ^a	Charakteristikum	Studieninformation
		<p>Glimepirid 8 Wochen vor Studienbeginn</p> <ul style="list-style-type: none"> - AST >3 x ULN - ALT >3 x ULN - Gesamtbilirubin im Serum > 34 µmol/L - Berechnete Kreatininclearance < 50 mL/min - Verhältnis von Eiweiß im Urin zu Kreatinin >1800 mg/g - Kreatininkinase >3 x ULN - Hämoglobin ≤100 g/L für Männer und ≤95 g/L für Frauen - Unkontrollierter Bluthochdruck (systolischer Blutdruck ≥180 mmHg und/oder diastolischer Blutdruck ≥ 110 mmHg) - Abnormale thyreoidstimulierende Hormonwerte - Bedeutende, kardiovaskuläre Vorgeschichte innerhalb der letzten 6 Monate vor der 1. Visite - Bösartiger Tumor innerhalb der letzten 5 Jahre - Vorgeschichte eines von Blutfett induzierten Xanthoms oder einer Hypertriglyzeridämie-induzierten Pankreatitis - Vorgeschichte einer labilen oder einer schnell fortschreitenden Nierenkrankheit - Wissen einer erblich bedingten renalen Glukosurie - Vorgeschichte einer chronischen hämolytischen Anämie - Patienten (aus Sicht der Prüfarzte) mit Risiko einer Dehydratation - Schwanger oder stillend - BMI >45,0 kg/m² - Insulintherapie innerhalb von 24 Wochen vor dem Einschluss - Verdacht, dass Patient infiziert (WHO Risikofaktor 2 bis 4) - Adipositaschirurgie - Gewichtsreduzierende Medikation - Behandlung mit antiviralen

Item ^a	Charakteristikum	Studieninformation
		Medikamenten/HIV-Behandlung <ul style="list-style-type: none"> - Überempfindlichkeit gegen Dapagliflozin, Glimepirid, Unverträglichkeit / Kontraindikation gegenüber Sulfonylharnstoffe, - Drogen-induzierte Myopathie oder Kreatinkinase-Bewertung - Schwere Lebererkrankungen - Vorgeschichte von Drogen-induzierten Leberenzym-Werten - Vorgeschichte von <u>Hämoglobinopathie</u> - Drogen- oder Alkoholmissbrauch innerhalb der letzten 5 Jahre - Klinisch bedeutende Abnormalitäten, schwere labile medizinische oder seelische Verfassung - Beteiligung an der Planung und Durchführung der Studie - Bereits Teilnahme an dieser Studie - Teilnahme an einer klinischen Studie mit Dapagliflozin und/oder einem SGLT-2 Inhibitor - Erhalt eines Prüfpräparats innerhalb 12 Wochen vor der 1. Visite
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische, internationale Studie durchgeführt an 84 Zentren in 7 Ländern
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<i>Open-label</i> Glimepirid+Placebo versus <i>Open-label</i> Glimepirid+2,5 mg Dapagliflozin versus <i>Open-label</i> Glimepirid+5 mg Dapagliflozin versus <i>Open-label</i> Glimepirid+10 mg Dapagliflozin
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 Ziel (nach 24 Wochen): Veränderung des HbA1c vom Ausgangswert Sekundäre Ziele (nach 24 Wochen): Veränderung des Gewichts Veränderung des oralen Blutzuckertests (OGTT) Anteil der Patienten, die eine therapeutisch glykämische <i>Response</i> (HbA1c <7%) erreichen

Item ^a	Charakteristikum	Studieninformation
		<p>Veränderung des Körpergewichts an Patienten mit einem BMI-Ausgangswert $\geq 27 \text{ kg/m}^2$</p> <p>Veränderung des FPG</p> <p>Andere sekundäre Endpunkte (nach 24 Wochen):</p> <p>Anteil der Patienten, die eine therapeutisch glykämische <i>Response</i> (HbA1c <6,5%) erreichen</p> <p>Veränderung des HbA1c bei Patienten mit einem BMI $\geq 27 \text{ kg/m}^2$ und einem BMI $\geq 30 \text{ kg/m}^2$</p> <p>Veränderung des HbA1c bei Patienten mit einem HbA1c-Ausgangswert <8%, $\geq 8\%$, <9% und $\geq 9\%$</p> <p>Veränderung des FPG nach einer Woche</p> <p>Veränderung des Bauchumfangs</p> <p>Veränderung des Körpergewichts bei Patienten mit einem BMI-Ausgangswert $\geq 30 \text{ kg/m}^2$</p> <p>Patientenzufriedenheit (DTSQc und DTSQs)</p> <p>Forschungsendpunkte (nach 24 Wochen):</p> <p>Prozentuale Veränderung des Körpergewichts</p> <p>Prozentuale Veränderung der <i>Fasting</i> Lipids</p> <p>Prozentuale Veränderung der freien Fettsäuren (FFS)</p> <p>Veränderung des systolischen und diastolischen Blutdrucks</p> <p>Veränderung des systolischen Blutdrucks bei Patienten mit einem Ausgangswert von >140 mmHg</p> <p>Haupt-Wirksamkeitsendpunkte (nach 48 Wochen im Vergleich zum Ausgangswert):</p> <p>Veränderung des HbA1c von Woche 24-48 und von Woche 0-48</p> <p>Anteil der Patienten mit glykämischer <i>Response</i> (HbA1c <7%)</p> <p>Veränderung des Körpergewichts vom Ausgangswert nach 48 Wochen</p> <p>Veränderung des Körpergewichts bei Patienten mit einem BMI-Ausgangswert $\geq 27 \text{ kg/m}^2$</p> <p>Veränderung im 2 h-<i>post-challenge</i> Plasma-Glukose-Level</p> <p>Anteil der Patienten mit HbA1c $\leq 6,5\%$ in Woche 48</p> <p>Veränderung des HbA1c bei Patienten mit einem BMI $\geq 27 \text{ kg/m}^2$ und einem BMI $\geq 30 \text{ kg/m}^2$</p> <p>Veränderung des HbA1c bei Patienten mit</p>

Item ^a	Charakteristikum	Studieninformation
		<p>einer <i>Baseline</i> HbA1c <8%, ≥8%, <9% und ≥9%</p> <p>Veränderung des FPG</p> <p>Veränderung des Bauchumfangs</p> <p>Veränderung des Körpergewichts</p> <p>Veränderung des Körpergewichts bei Patienten mit einer <i>Baseline</i> BMI ≥30 kg/m²</p> <p>Patientenzufriedenheit (DTSQc und DTSQs)</p> <p>Prozentuale Veränderung des Körpergewichts</p> <p>Prozentuale Veränderung der Nüchtern-Lipide</p> <p>Prozentuale Veränderung der FFS</p> <p>Anteil der Patienten, die die Studie abbrechen aufgrund unzureichender Blutzuckerkontrolle oder eine für den Notfall bestimmte Medikation aufgrund unzureichender Blutzuckerkontrolle genommen haben (in Woche 32, 40 und 48)</p> <p>Veränderung des systolischen und diastolischen Blutdrucks</p> <p>Veränderung des systolischen Blutdrucks bei Patienten mit einem Ausgangswert von >140 mmHg</p> <p>Sicherheitsendpunkte (sowohl nach 24 als auch nach 48 Wochen)</p> <p>Unerwünschte Ereignisse</p> <p>Hypoglykämien</p> <p>Laborwerte</p> <p>Kreatinin-Clearance</p> <p>EKG</p> <p>Vitalwerte (Puls und Blutdruck)</p> <p>Körperliche Untersuchung</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Annahmen, um eine Differenz des HbA1c von 0,5%-Punkten zwischen jeder Dapagliflozin-Gruppe und Placebo nach 24 Wochen zu entdecken</p> <p>Standardabweichung 1,1%, zweiseitiger Test mit Signifikanzniveau von 0,019 bei einer Power von 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

Item ^a	Charakteristikum	Studieninformation
		unterschiedenen <i>Informed Consent</i>
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Balancierte Blockrandomisierung von fortlaufenden Randomisierungs-codes
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	Randomisierung erfolgt streng nach der Reihenfolge innerhalb eines Zentrums
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 zur Verfügung
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 aussehender Studienmedikation in identischer Anzahl für Patienten und medizinisches Personal
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>24 Wochen Studie: Der primäre Endpunkt (Veränderung des HbA1c nach 24 Wochen) wird mit der Methode von Dunnett analysiert. (Gesamt-$\alpha = 0,05$). Das Signifikanzlevel für das paarweise Testen ist 0,019. Sekundäre Endpunkte: Tests werden der Reihe nach durchgeführt, um das α innerhalb jeder Dapagliflozin-Gruppe zu kontrollieren. Die statistische / nicht-statistische Signifikanz vom primären Endpunkt bestimmt die statistische Inferenz, die für die sekundären Endpunkte durchgeführt wird. Stetige Wirksamkeitsendpunkte wurden mittels einer ANCOVA, dichotome Endpunkte wurden mittels Anteile mit der Methode von Zhang Tsiatis und Davidian durchgeführt (<i>Full analysis set</i>). Zusammenfassungen der Sicherheitsendpunkte wurde mittels <i>Safety analysis set</i> gemacht.</p> <p>48 Wochen Studie: Langzeitergebnisse wurden nur als eine orientierende Untersuchung betrachtet, LOCF wurde nicht verwendet, stetige Wirksamkeitsvariablen wurden mittels eines</p>

Item ^a	Charakteristikum	Studieninformation
		Modells für wiederholte Messungen untersucht, es wurden keine p-Werte für Gruppenvergleiche angegeben. Wirksamkeitsendpunkte wurden anhand des <i>Full analysis set</i> untersucht, Zusammenfassungen der Sicherheitsendpunkte wurde mittels <i>Safety analysis set</i> gemacht.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Subgruppen: Veränderung des HbA1c nach Geschlecht, Alter, Rasse, Region, <i>Baseline</i> HbA1c und <i>Baseline</i> eGFR
-	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 berücksichtigt wurden	Dapagliflozin 2,5 mg vs 5 mg vs.10 mg vs. Placebo a) 154 vs 145 vs. 151 vs. 146 b) 154 vs 142 vs. 151 vs. 145 c)154 vs 142 vs. 151 vs. 145
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flowchart</i>
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studienbeginn: 30.04.2008 Studienende: 20.05.2010
14b	Informationen, warum die Studie endete oder beendet wurde	Studie endete regulär
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

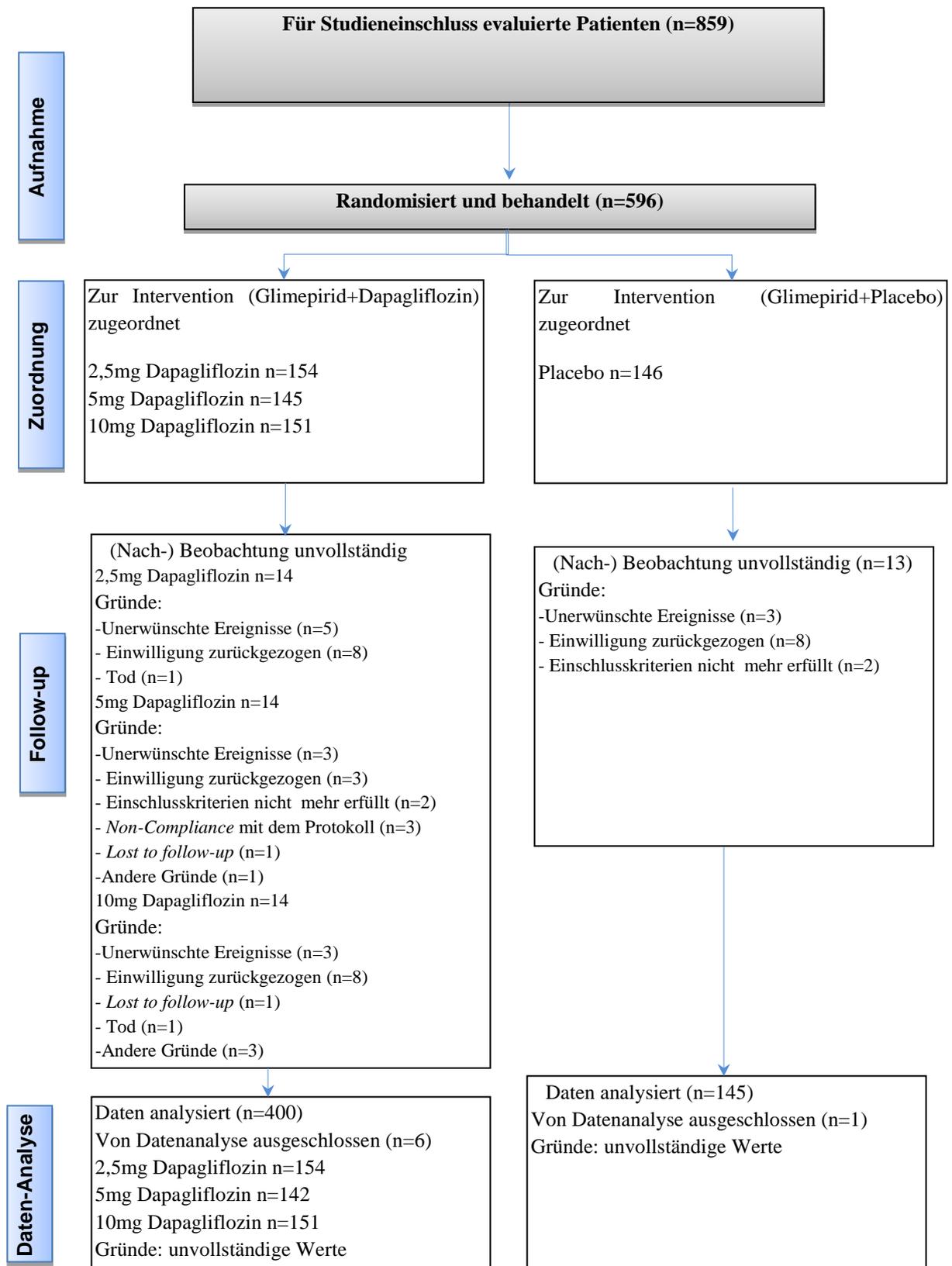


Abbildung 19: Flussdiagramm Studiendesign und -methodik für Studie D1690C00005

Tabelle 4-151 (Anhang): Studiendesign und -methodik für Studie DeFronzo et al. 1995

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Primäres Ziel: Beurteilung der Wirksamkeit von Metformin als Therapiemöglichkeit bei mäßig adipösen Patienten mit Typ-2-Diabetes, anhand von Blutproben (nüchtern; zur Messung von glykosiliertem Hämoglobin, Nüchternblutzucker und Laktat), sowie der Messung von Lipiden und Insulin bevor, während und nach Beendigung der Studie.</p> <p><i>Studienprotokoll 1:</i> Beurteilung der Wirksamkeit von Metformin bei Patienten, deren Diabetes nicht ausreichend durch Diät alleine kontrolliert werden konnte (Studienaufbau: Metformin vs. Placebo).</p> <p><i>Studienprotokoll 2:</i> Beurteilung der Wirksamkeit von Metformin bei Patienten, deren Diabetes nicht ausreichend durch Diät und die Gabe von Glibenclamid kontrolliert werden konnte (Studienaufbau: Glibenclamid+Metformin-Placebo vs Metformin+Glibenclamid-Placebo vs Metformin+Glibenclamid).</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>Zwei randomisierte, placebokontrollierte, multizentrische Studienprotokolle.</p> <p><i>Studienprotokoll 1:</i> Vorrandomisierungsphase (Phase I), Dauer 8 Wochen; Metformin-Titrationsphase (Phase 2), Dauer 6 Wochen; Aktive Metformin-Behandlungsphase (Phase III), Dauer 23 Wochen Zuteilungsverhältnis: 1:1 für Metformin:Placebo</p> <p><i>Studienprotokoll 2:</i> Vorrandomisierungsphase (Phase I), Dauer 5 Wochen; Metformin-Titrationsphase (Phase 2), Dauer 6 Wochen; Aktive Metformin-Behandlungsphase (Phase III), Dauer 23 Wochen Zuteilungsverhältnis 1:1:1 für Glibenclamid+Metformin-Placebo: Metformin+Glibenclamid-Placebo: Metformin+Glibenclamid</p>
3b	Relevante Änderungen der Methodik nach	Nicht zutreffend

Item ^a	Charakteristikum	Studieninformation
	Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien bei 1. Visite (<i>Enrollment</i>):</p> <ul style="list-style-type: none"> - <i>Informed Consent</i> - Diagnose: Typ-2-Diabetes - Männer oder Frauen zwischen 40 und 70 Jahren - Mangelhaft kontrollierter Typ-2-Diabetes gemessen am FBG ≥ 140 mg/dl nach 8-wöchiger Diät (<i>Studienprotokoll 1</i>) oder mindestens 4-wöchiger Diät plus Behandlung mit 20 mg Glibenclamid täglich (<i>Studienprotokoll 2</i>) - Erhöhtes Gewicht (120-170% des Idealgewichts, gemessen an den <i>Metropolitan Life Insurance</i> Tabellen von 1983) - Normale Nierenfunktion (Serum Kreatinin $\leq 1,4$ mg/dl, 124 μmol/l bei Männern und $\leq 1,3$ mg/dl, 115 μmol/l bei Frauen; und $\leq 2+$ Proteinurie) - Normale Leberfunktion <p>Ausschlusskriterien bei 1. Visite:</p> <ul style="list-style-type: none"> - Symptomatischer Diabetes mellitus (Polyurie, Polydipsie, Gewichtsverlust) - Symptomatische kardiovaskuläre Erkrankungen - Diastolischer Blutdruck über 100 mmHg während der Behandlung mit blutdrucksenkenden Mitteln - Weitere Begleiterkrankung - Insulin-Behandlung innerhalb der letzten 6 Monate vor Studienbeginn - Einnahme von Medikamenten, die den Glukose-Metabolismus beeinflussen - Einnahme von 3 oder mehr alkoholischen Getränken am Tag (≥ 3 oz Alkohol am Tag) - Gebrauch von illegalen Drogen - Kürzlich Behandlung mit Metformin <p>Die Behandlung mit Östrogen, Progestin und Chlortalidon oder einem Thiaziddiuretikum war erlaubt unter der Voraussetzung, dass die Dosis während der Studie konstant blieb.</p>
4b	Studienorganisation und Ort der Studiendurchführung	Zwei randomisierten, parallele, doppelblinde, placebokontrollierte, multizentrische

Item ^a	Charakteristikum	Studieninformation
		Studienprotokolle, durchgeführt in den USA.
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p><i>Studienprotokoll 1:</i></p> <p>Vorrandomisierungsphase (Phase I): 8-wöchige Diät, mit 20% weniger Kalorien als der kalkulierte Energiebedarf des Patienten</p> <p>Randomisierung: Woche 0 Zuteilungsverhältnis 1:1 für Metformin:Placebo-Behandlung</p> <p>Metformin-Titrationsphase (Phase II): Woche 0-2: 850 mg Metformin oder Placebo zum Abendessen Woche 2-4: 850 mg Metformin oder Placebo zweimal täglich (zum Frühstück und Abendessen) Woche 5: 850 mg Metformin oder Placebo dreimal täglich (zum Frühstück, Mittagessen, Abendessen)</p> <p>Aktive Metformin-Behandlungsphase (Phase III): Einnahme der maximalen Dosis (3*850 mg = 2550 mg täglich) Metformin oder Placebo bis Woche 29; bei Eintreten von Nebenwirkungen Reduktion der Dosis; Kontrolle alle 4 Wochen</p> <p><i>Studienprotokoll 2:</i></p> <p>Vorrandomisierungsphase (Phase I): einwöchige Vorbehandlung mit 5 mg Glibenclamid täglich und Diät; anschließend 4-wöchige Vorbehandlung mit 10 mg Glibenclamid täglich und Diät</p> <p>Randomisierung: Woche 0 Zuteilungsverhältnis 1:1:1 für Glibenclamid+Metformin-Placebo: Metformin+Glibenclamid-Placebo: Metformin+Glibenclamid</p> <p>Metformin-Titrationsphase (Phase II): Woche 1: 500 mg Metformin oder Placebo täglich zum Abendessen; 4*5 mg Glibenclamid oder Placebo täglich (20 mg) Woche 2: 1000mg Metformin oder Placebo täglich (500 mg zum Frühstück, 500 mg zum Abendessen); 4*5 mg Glibenclamid oder Placebo täglich (20 mg) Woche 3: 1500 mg Metformin oder Placebo täglich (500 mg zum Frühstück, 500 mg zum Mittagessen, 500 mg zum Abendessen); 4*5 mg Glibenclamid oder Placebo täglich (20 mg) Woche 4: 2000 mg Metformin oder Placebo</p>

Item ^a	Charakteristikum	Studieninformation
		<p>täglich (500 mg zum Frühstück, 500 mg zum Mittagessen, 1000 mg zum Abendessen); 4*5 mg Glibenclamid oder Placebo täglich (20 mg)</p> <p>Woche 5: 2500 mg Metformin oder Placebo täglich (1000 mg zum Frühstück, 500 mg zum Mittagessen, 1000 mg zum Abendessen); 4*5 mg Glibenclamid oder Placebo täglich (20 mg)</p> <p>Diese Konzentration wurde beibehalten, solange die FBG bei ≥ 140 mg/dL blieb.</p> <p>Aktive Metformin-Behandlungsphase (Phase III):</p> <p>2500 mg Metformin oder Placebo täglich (1000 mg zum Frühstück, 500 mg zum Mittagessen, 1000 mg zum Abendessen); 4*5 mg Glibenclamid oder Placebo täglich (20 mg); bei Eintreten von Nebenwirkungen Reduktion der Dosis</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>Primäres Ziel (nach 29 Wochen): Beurteilung der Wirksamkeit von Metformin.</p> <p>Sekundäre Ziele (nach 29 Wochen): <i>Studienprotokoll 1 und 2:</i> Gemessen bei jedem Besuch:</p> <ul style="list-style-type: none"> - Veränderung des Gewichts - die <i>Compliance</i> der Patienten - Nebenwirkungen - interkurrente medizinische Ereignisse - Blutdruck - Blutproben (nüchtern; zur Messung von glykosiliertem Hämoglobin, Nüchternblutzucker und Laktat) - Veränderung Insulin <p>Gemessen in Woche 29</p> <ul style="list-style-type: none"> - oraler Glukose-Toleranz-Test - Plasma-Lipid-Konzentration - Serum Vitamin B12- Konzentration - Folsäure-Konzentration <p>Gemessen in Woche 9, 21 und 29:</p> <ul style="list-style-type: none"> - Plasma-Metformin
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Nicht berichtet
7b	Falls notwendig, Beschreibung von	Nicht zutreffend

Item ^a	Charakteristikum	Studieninformation
	Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierte, parallele, doppelblinde, placebokontrollierte, multizentrische Studienprotokolle Zuteilungsverhältnis 1:1 (Metformin:Placebo-Behandlung) für <i>Studienprotokoll 1</i> , 1:1:1 (Glibenclamid+Metformin-Placebo: Metformin+Glibenclamid-Placebo: Metformin+Glibenclamid) für <i>Studienprotokoll 2</i>
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?	a) ja b) ja c) ja Randomisierung und identische Studienmedikation
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Verwendung von gleich aussehenden Studienmedikationen in identischer Anzahl für Patienten und medizinisches Personal
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die primären Endpunkte wurden mit der ITT-Population analysiert. Absolute Werte und Veränderungen in der <i>Baseline</i> wurden für alle Wirksamkeitsmessungen berücksichtigt. Zwei Analysen wurden mit der ITT-Population durchgeführt: 1) nur mit den beobachteten Werten; 2) fehlende Werte wurden mit der LOCF-Methode ersetzt. Statistischer Vergleich wurde mit der SAS <i>Software</i> durchgeführt. Der Vergleich innerhalb von Gruppen wurde mit dem verbundenen zweiseitigem T-Test durchgeführt. Für stetige Variablen wurde der Vergleich

Item ^a	Charakteristikum	Studieninformation
		zwischen Gruppen mit einem linearen Model verwendet. Diese Modelle enthielten Behandlungseffekt, Zentrum (Varianz-Analyse), mit selektierten Modellen inklusive einem Effekt der <i>Baseline</i> -Werte (Kovarianz-Analyse). Alle Werte wurden im Mittel \pm SE angegeben.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Studienprotokoll 1 und 2
-	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 berücksichtigt wurden	<i>Studienprotokoll 1:</i> Randomisierung: Woche 0 Zuteilungsverhältnis 1:1 für Metformin:Placebo-Behandlung a) 143:146 b) 112:105 c) 217 Patienten beendeten die aktive Metformin Behandlungsphase; das Verhältnis war: 112:105 <i>Studienprotokoll 2:</i> Zuteilungsverhältnis 1:1:1 für Glibenclamid+Metformin-Placebo: Metformin+Glibenclamid-Placebo: Metformin+Glibenclamid a) 209:210:213 b) 174:173:192 c) 523 beendeten die aktive Metformin Behandlungsphase; das Verhältnis war: 174:157:192
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flowchart für Studienprotokoll 1 (Abbildung 20) und Studienprotokoll 2 (Abbildung 21)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Nicht berichtet
14b	Informationen, warum die Studie endete oder beendet wurde	Studie endete regulär
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

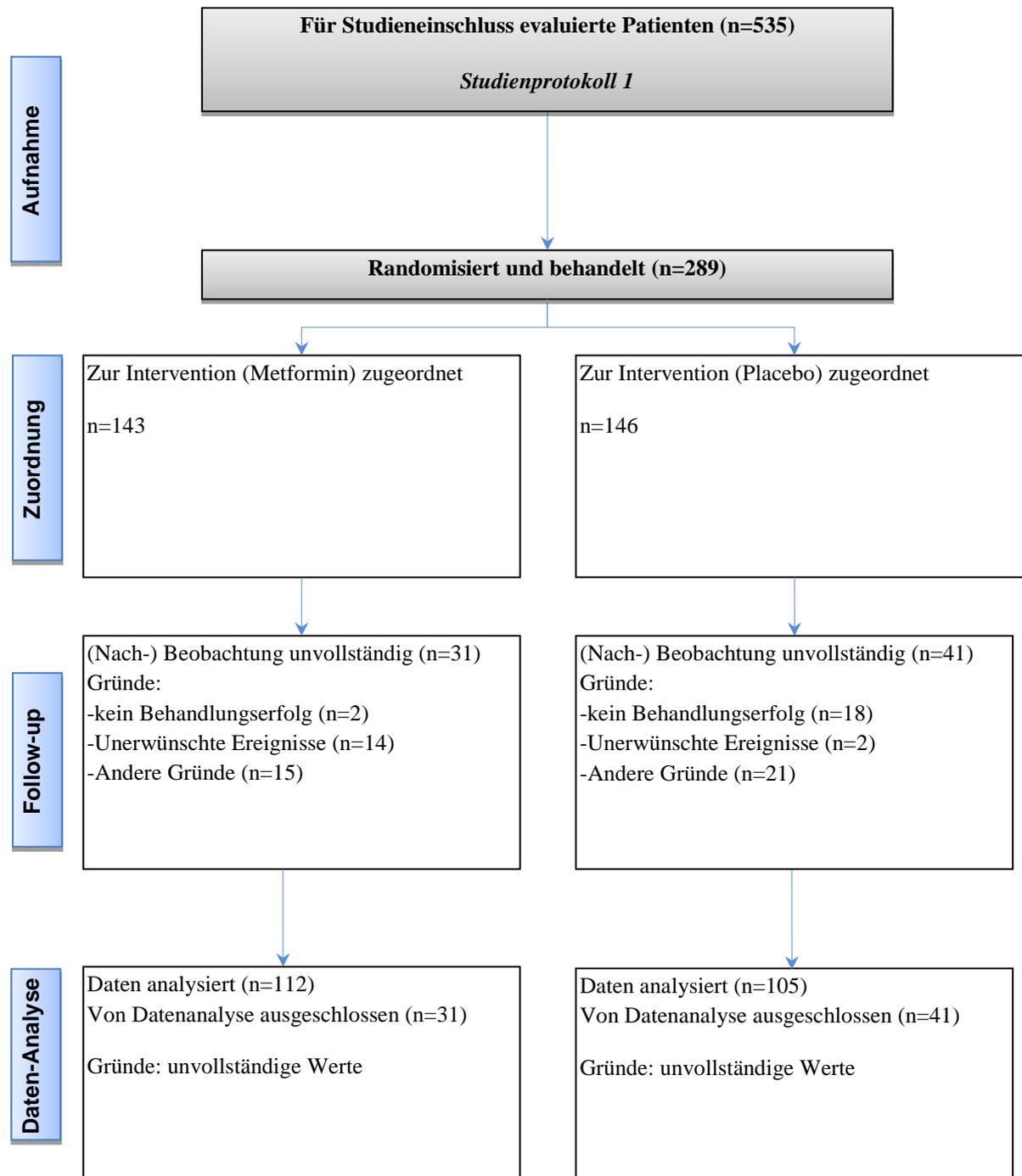


Abbildung 20: Flussdiagramm Studiendesign und -methodik für Studie DeFronzo 1995 (Protokoll 1)

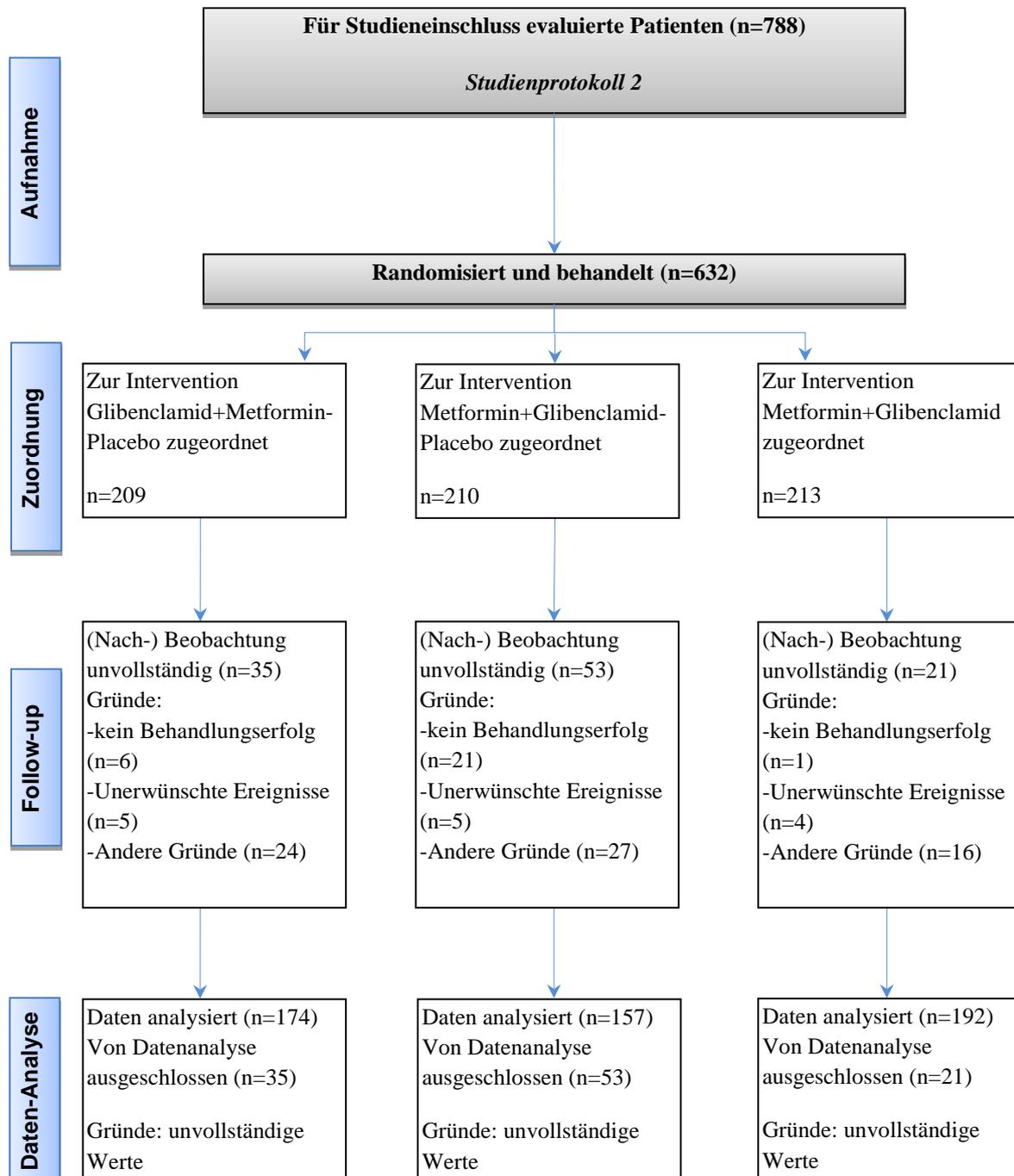


Abbildung 21: Flussdiagramm Studiendesign und –methodik für DeFronzo 1995 (Protokoll 2)

Tabelle 4-152 (Anhang): Studiendesign und –methodik für Studie Goldstein et al. 2003

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Primäres Ziel (nach 18 Wochen): Beurteilung der Wirksamkeit und Verträglichkeit von Glipizid+Metformin als Kombinationspräparat zur Anwendung als Zweitlinienmedikament bei Patienten mit Typ-2-Diabetes im Vergleich zu Metformin oder Glipizid als Monotherapie durch die Veränderung des HbA1c nach 18 Wochen.</p> <p>Sekundäre Ziele (nach 18 Wochen):</p> <ul style="list-style-type: none"> - <i>Mean daily glukose</i> (MDG) - Veränderung der Glucosewerte (Nüchternblutzucker (FBG) sowie 3-Stunden Post-prandialer Blutzucker (PBG) unter Berücksichtigung der Konzentrationszeitkurve (AUC), sowie 3-Stunden Post-prandiales Insulin (PPI) bei zunehmenden AUC - Veränderung des Nüchtern-Insulin-Spiegels - Veränderung der Serum Lipid Profile - Veränderung des Körpergewichts <p>Andere sekundäre Ziele (nach 18 Wochen):</p> <ul style="list-style-type: none"> - Vergleich der Verträglichkeit der verschiedenen Behandlungsarme - Beurteilung der Sicherheit und Verträglichkeit der 3 Behandlungsarme durch die Erhebung unerwünschter Ereignisse und veränderter Laborwerte
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>Randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte 18-wöchige Studie.</p> <p>2-wöchige einfachblinde Einführungsphase, in der die Patienten von der Sulphonylurea-Monotherapie zu 15 mg Glipizid täglich umgestellt werden.</p> <p>Anschließende randomisierte Behandlung für 18-wöchige Studie.</p> <p>Zuteilungsverhältnis: 1:1:1; 30 mg Glipizid : 500 mg Metformin : 5 mg Glipizid+500 mg Metformin.</p>
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 1. Visite (<i>Enrollment</i>):</p> <ul style="list-style-type: none"> - <i>Informed Consent</i> - Diagnose: Typ-2-Diabetes mellitus seit >3 Monaten - FPG < 300 mg/dL - HbA1c zwischen $\geq 7,5$ und $\leq 12,0\%$ - BMI zwischen ≥ 25 und ≤ 40 kg/m² - Bereitschaft und Fähigkeit zur Selbstüberprüfung des Blutzuckers - Mangelhaft kontrollierter Typ-2-Diabetes durch Behandlung mit Sulphonylharnstoff als Monotherapie <p>Ausschlusskriterien bei 1. Visite:</p> <ul style="list-style-type: none"> - Symptomatischer Diabetes mellitus (deutliche Polyurie und Polydipsie oder >10,0% Gewichtsverlust) - Signifikante renale, hepatische oder kardiovaskuläre Krankheiten - Bei Verwendung von antihyperglykämischen Mitteln (ausgenommen Sulphonylharnstoffe) innerhalb der letzten 8 Wochen vor Einschluss in die Studie - Vorgeschichte von diabetischer Ketoacidose, Hyperosmolarem nicht-ketotischem Koma oder Langzeitinsulintherapie
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische, nationale Studie in 108 ambulanten Krankenhäusern in den USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Woche -2-0 Vorbehandlung: Vor Beginn der Studie wurden die Patienten von ihrer vorherigen Behandlung mit Sulphonylharnstoff zu 15 mg Glipizid zweimal täglich umgestellt.</p> <p>Woche 0-18 Behandlungsphase: Glipizid vs Metformin vs Glipizid+Metformin Die Patienten, die für die Glipizid+Metformin (5+500 mg) oder die Metformin-Monotherapie (500 mg) randomisiert wurden, erhielten in der ersten Studienwoche eine Tablette täglich. Die Dosis wurde in der zweiten Woche zu zwei Tabletten (morgens und abends) täglich erhöht. In Woche 3, 5 und 8 konnte die Dosis erhöht werden, abhängig von der MDG (>130 mg/dL) und der kapillaren Blutglukosekonzentration (>100 mg/dL).</p> <p>Patienten, die für die Behandlung mit Glipizid als Monotherapie randomisiert wurden, erhielten 15 mg Glipizid zweimal täglich,</p>

		<p>jeweils 30 Minuten vor Nahrungsaufnahme.</p> <p>Die maximale Gesamtdosis von Glipizid war 30 mg, von Metformin 2000 mg und Glipizid+Metformin 20+2000 mg.</p> <p>Die Verringerung der Dosis wurde Patienten erlaubt, die bei mehr als einem Anlass einen Blutglukosewert von <60 mg/dL und Symptome einer Hypoglykämie hatten.</p> <p>Die Studie wurde abgebrochen, wenn die Blutglukosewerte ungenügend kontrolliert werden konnten, was durch den FPG >280 mg/dL in Woche 5, >240 mg/dL während der Wochen 5 bis 12, oder >200 mg/dL von Woche 12 an, bestimmt wurde.</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>Primäres Ziel (nach 18 Wochen):</p> <p>Beurteilung der Wirksamkeit und Verträglichkeit von Glipizid+Metformin als Kombinationspräparat zur Anwendung als Zweitlinienmedikament bei Patienten mit Typ-2-Diabetes, welche durch Monotherapie mit zumindest der Hälfte der maximal beschriebenen täglichen Dosis von Sulfonylharnstoff ungenügend behandelt sind.</p> <p>Die Wirksamkeit und Verträglichkeit von Glipizid +Metformin als Kombinationstherapie wird mit der Monotherapie mit Metformin oder Glipizid anhand der Veränderung des HbA1c in Woche 18 bei Patienten mit Typ-2-Diabetes beurteilt.</p> <p>Sekundäre Ziele (nach 18 Wochen):</p> <ul style="list-style-type: none"> - <i>Mean daily Glucose</i> (MDG) - Veränderung der Glukosewerte (Nüchternblutzucker (FBG) sowie 3-Stunden Post-prandialer Blutzucker (PBG) unter Berücksichtigung der Konzentrationszeitkurve (AUC), sowie 3-Stunden Post-prandiales Insulin (PPI) bei zunehmenden AUC - Veränderung des Nüchtern-Insulin-Spiegels - Veränderung der Serum Lipid Profile - Veränderung des Körpergewichts <p>Andere sekundäre Ziele (nach 18 Wochen):</p> <ul style="list-style-type: none"> - Vergleich der Verträglichkeit der verschiedenen Behandlungsarme - Beurteilung der Sicherheit und Verträglichkeit der drei Behandlungsarme durch die Erhebung unerwünschter Ereignisse

		und veränderter Laborwerte
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Fallzahl: mindestens 74 Patienten pro Gruppe Annahme: Überlegenheit von Glipizid+Metformin über jede der Monotherapien. HbA1c ist in der Glipizid+Metformin Gruppe um 0,9% geringer als bei einer der Monotherapien. Standardabweichung 1,5%, einseitiger Test mit Signifikanzniveau von 0,025 bei einer Power von 90%, 5%-ige Ausschlussrate .
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, die die Einschlusskriterien erfüllten, wurden mit einer fortlaufenden randomisierten Nummer ausgestattet. Diese wurde zur Anweisung einer verblindeten Therapie entsprechend eines Randomisierungsschema benutzt, welches an einer zentralen Stelle aufbewahrt wurde. Die Zuteilung erfolgte im Verhältnis 1:1:1 für 30 mg Glipizid-Monotherapie, 500 mg Metformin-Monotherapie oder 5+500 mg Glipizid+Metformin-Kombinationstherapie.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Blockrandomisierung Jedes Zentrum wurde mit allen Studienmedikationen in einem 3-er Block ausgestattet. Jeder Block enthielt die Versorgung für eine Behandlung. Dadurch war die Behandlung in jedem Zentrum ausgewogen.
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	Randomisierung erfolgte durch ein zentrales Randomisierungszentrum
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?	Patienten, die die Einschlusskriterien erfüllten, wurden mit einer fortlaufenden randomisierten Nummer ausgestattet. Diese wurden an einer zentralen Stelle aufbewahrt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die	a) ja b) ja

	Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	c) ja Randomisierung und Blockrandomisierung
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Jeder Block enthielt die Medikation für eine Behandlung.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Analyse des HbA1c und anderer Endpunkte wurde mit einer Kovarianz-Analyse mit Behandlungsgruppe und <i>Baseline</i> -Werten als Kovariablen durchgeführt.
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)	
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	30 mg Glipizid-Monotherapie : 500 mg Metformin-Monotherapie : 5+500 mg Glipizid+Metformin-Kombinationstherapie a) 84:76:87 (insgesamt 247) b) 61:50:67 (insgesamt 178) c.) 247 Patienten wurden randomisiert und 246 wurden behandelt, 178 Patienten beendeten ihre Therapie
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe <i>Flowchart</i> (Abbildung 22)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Nicht berichtet
14b	Informationen, warum die Studie endete oder beendet wurde	Studie endete regulär
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

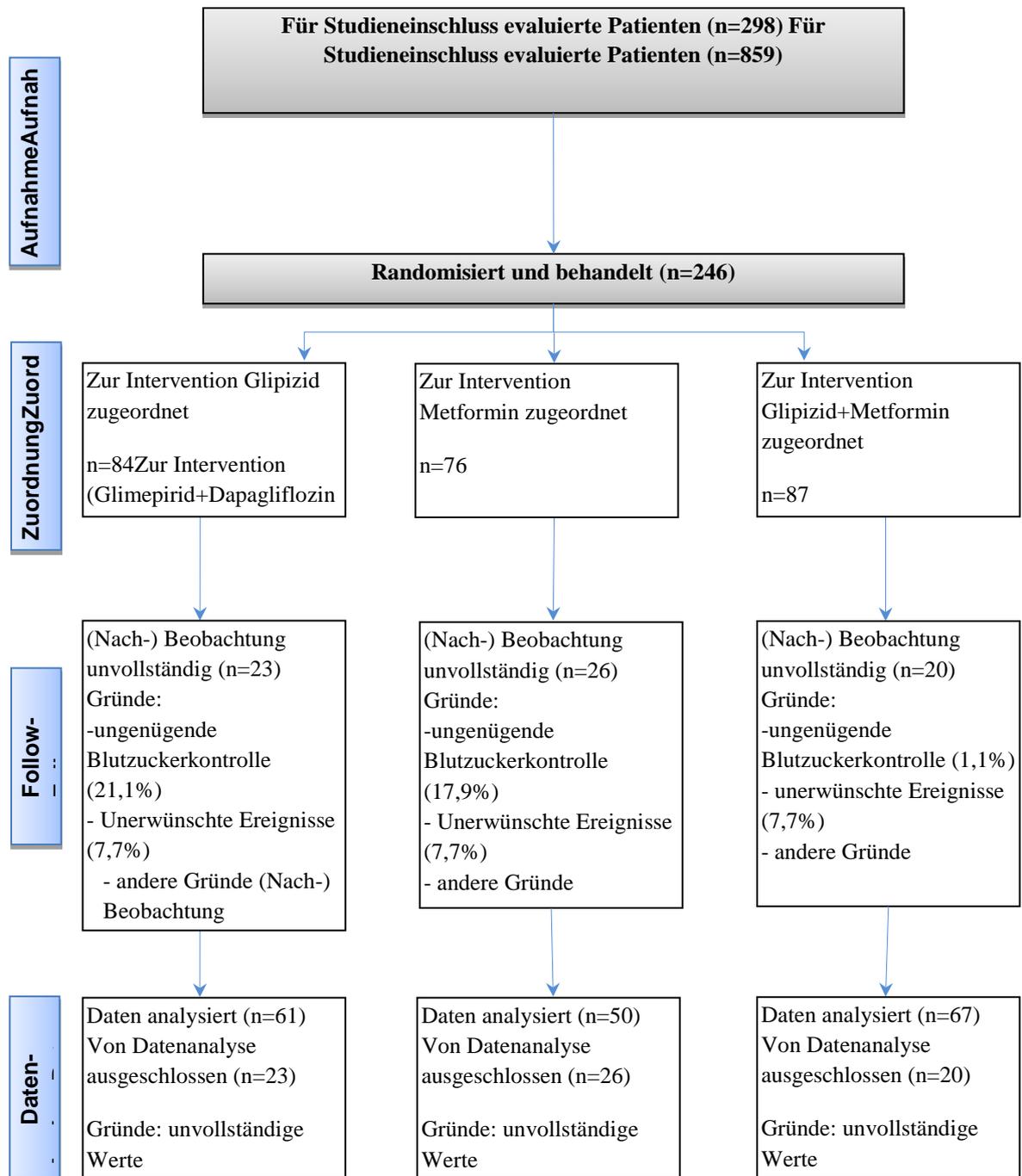


Abbildung 22: Flussdiagramm Studiendesign und-methodik für Goldstein 2003

Tabelle 4-153 (Anhang): Studiendesign und -methodik für Studie Hermansen et al. 2007

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Bestimmung von Wirksamkeit und Sicherheit des selektiven oralen DDP-4-Hemmers Sitagliptin, zusätzlich gegeben zu einer Glimepirid-Monotherapie oder einer Kombination von Glimepirid und Metformin
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multinationale, randomisierte, doppelblinde Parallelgruppenstudie mit einfach blinder Run-in Phase
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	keine Angabe
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <p>≥18 bis ≤75 Jahre Typ-2 Diabetes Eines der drei Folgenden:</p> <ul style="list-style-type: none"> - Bereits mit Glimepirid oder Glimepirid + Metformin behandelt - Einnahme eines oder ≤ dreier OAD - kein OAD in den letzten 8 Wochen <p>HbA1c ≥7,5% bis ≤10,5% am Ende der einfach blinden Placebo <i>Run-in</i> Phase Compliance ≥75% in der Placebo <i>Run-in</i> Phase</p> <p>Ausschlusskriterien</p> <p>Typ-1 Diabetes Insulingabe innerhalb 8 Wochen vor der Screening-Untersuchung Kreatininclearance <45ml/min (<60 ml/min bei Metforminmedikation) Bekannte Überempfindlichkeit, Unverträglichkeit oder Kontraindikationen gegen Glimepirid, Sulfonylharnstoffe, Metformin oder Pioglitazon (vorgesehene Rescue-Medikation)</p>
4b	Studienorganisation und Ort der Studiendurchführung	Multinational; keine weiteren Angaben
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<ul style="list-style-type: none"> - Glimepirid + Placebo - Glimepirid+Metformin+Placebo - Glimepirid+100mg Sitagliptin 1 x tgl. - Glimepirid+Metformin+ 100 mg Sitagliptin

Item ^a	Charakteristikum	Studieninformation
		<p>1 x tgl.</p> <p>Die Dosierung von Glimepirid und Metformin war jeweils die, mit der sich der geforderte HbA1c-Wert (7,5%-10,5%) 10 Wochen lang halten lassen.</p> <p>Die Studienmedikation wurde jeweils im Anschluss an eine zweiwöchige einfach verblindete Placebo-Run-in Phase gegeben.</p> <p>Patienten mit anderen OAD oder ohne OAD hatten nach der Umstellung eine maximal 4-wöchige Dosistitratonsphase mit Glimepirid ± Metformin und anschließend eine bis zu 10-wöchige Dosistabilisierungsphase durchlaufen</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>Primärer Endpunkt mittlere HbA1c-Änderung der Ausgangswerte zu Woche 24</p> <p>Sekundäre Endpunkte Änderung der Nüchternblutzuckerwerte zu Woche 24 Prozentuale Veränderung von</p> <ul style="list-style-type: none"> - Gesamtcholesterin - LDL-Cholesterin - Triglyzeriden - HDL-Cholesterin - Non HDL-Cholesterin <p>zu Woche 24 Weitere Zielvariable HOMA-β Proinsulin/Insulin-Quotient HOMA-IR QUICKI</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	keine Angaben
7	Fallzahl	<p>1098 Patienten im Screening 441 Patienten randomisiert:</p> <ul style="list-style-type: none"> - 106 Placebo + Glimepirid, 113 Placebo + Glimepirid + Metformin (n=219) - 106 Sitagliptin + Glimepirid, 116 Sitagliptin + Glimepirid + Metformin (n=222)
7a	Wie wurden die Fallzahlen bestimmt?	keine Angabe
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	keine Angaben zu Zwischenanalysen, kein vorzeitiger Abbruch
8	Randomisierung, Erzeugung der Behandlungsfolge	

Item ^a	Charakteristikum	Studieninformation
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung und die Stratifizierung wurden mittels eines interaktiven <i>Voice Response Systems</i> (IVRS) kontrolliert
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratum 1: Glimepirid allein Stratum 2: Glimepirid + Metformin
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	keine Angaben
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?	1:1 Randomisierung via IVRS zu 100mg Sitagliptin oder Placebo 1 x täglich
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?	doppelblind
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht relevant
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Kovarianzanalyse (ANCOVA), adjustiert nach - Ausgangswerten für HbA1c - Stratum Ausgewertete Population: Alle Patienten, die zumindest eine Dosis der Studienmedikation erhalten hatten und für die neben den Ausgangswerten zumindest ein weiterer Wirksamkeitswert vorlag. Für fehlende Daten: LOCF
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Zusätzliche Auswertung des primären und der wichtigsten sekundären Endpunkte für die einzelnen Strata separat. Prädefinierte Kriterien zur Bildung von Subgruppen: HbA1c-Ausgangswert (< vs. > Medianwert) <8% vs. ≥8 bis <9% vs. ≥9% OAD vor Studienbeginn: keine vs. Monotherapie vs. Kombinationstherapie Geschlecht, Ethnizität, Alter (≥ vs. > Median)

Item ^a	Charakteristikum	Studieninformation
		BMI (\leq vs. $>$ Median) Diabetesdauer (\leq vs. $>$ Median) HOMA-IR (\leq vs. $>$ Median) HOMA- β (\leq vs. $>$ Median) Status metabolisches Syndrom nach NCEP ATP III
-	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 berücksichtigt wurden	Screening: 1098 Randomisiert: 441 Placebo+Glimepirid 106 Placebo+Glimepirid+Metformin 113 Sitagliptin+Glimepirid 106 Sitagliptin+Glimepirid+Metformin 116 Beendet: 364 Placebo+Glimepirid 87 Placebo+Glimepirid+Metformin 92 Sitagliptin+Glimepirid 83 Sitagliptin+Glimepirid+Metformin 102
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Placebo+Glimepirid zu hohe Glykämiewerte 10, <i>Lost to Follow-up</i> 2, Einverständnis widerrufen 5, Protokollverletzung 2 Placebo+Glimepirid+Metformin unerwünschte Wirkung 2, zu hohe Glykämiewerte 5, Einverständnis widerrufen 8, Protokollverletzung 4, protokollgerechter Abbruch 1, Prüfzentrum geschlossen 1 Sitagliptin+Glimepirid unerwünschte Wirkung 4, zu hohe Glykämiewerte 5, <i>Lost to Follow-up</i> 4, Einverständnis widerrufen 5, Protokollverletzung 2, sonstige 3 Sitagliptin+Glimepirid+Metformin Unerwünschte Wirkung 2, unerwünschte Laborwerte 1, zu hohe Glykämiewerte 7, Einverständnis widerrufen 3, Protokollverletzung 1
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	keine Angaben
14b	Informationen, warum die Studie endete oder beendet wurde	keine Angaben

Item ^a	Charakteristikum	Studieninformation
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

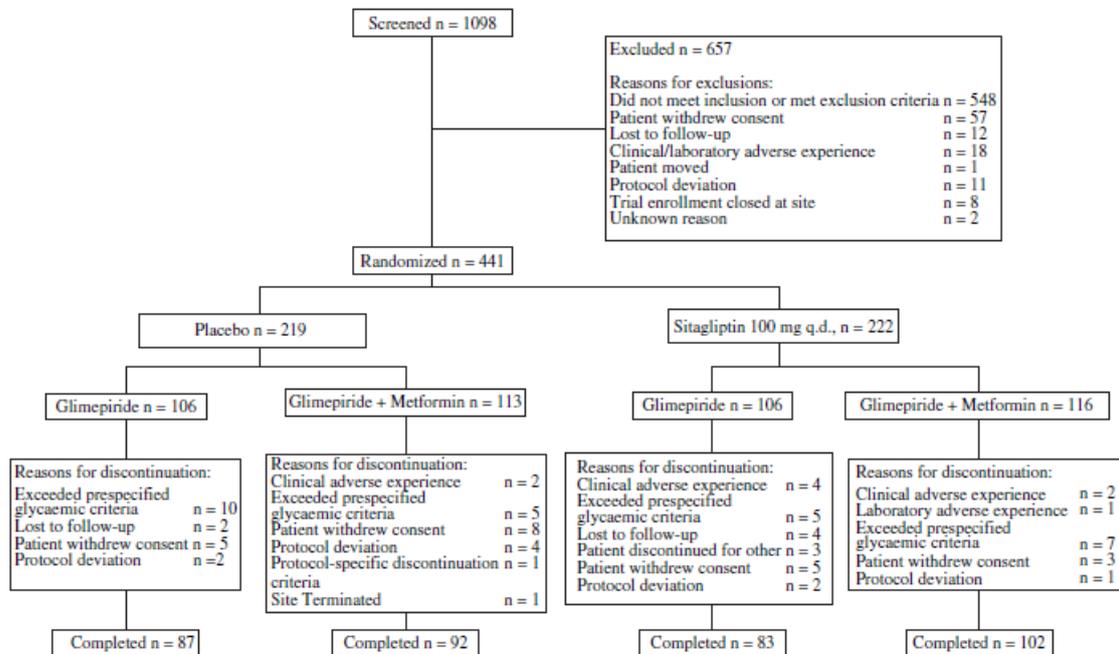


Abbildung 23: Flussdiagramm Studiendesign und-methodik für Hermansen 2007

Tabelle 4-154 (Anhang): Studiendesign und -methodik für Studie Chacra et al. 2009

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Bewertung von Wirksamkeit und Sicherheit von Saxagliptin in Kombination mit Glibenclamid im Vergleich zu auftitrierten Dosen Glibenclamid
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, doppelblinde, dreiarmige, multinationale Studie mit 4-wöchiger einfach blinder Run-in Phase
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	keine Angabe
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <p>18 bis 77 Jahre</p> <p>Typ-2 Diabetes</p> <p>Behandlung mit einem Sulfonylharnstoff in submaximaler Dosis seit mindestens zwei Monaten</p> <p>keine andere antihyperglykämische Medikation</p> <p>HbA1c $\geq 7,5$ - $\leq 10,0$ %</p> <p>Nüchtern C-Peptid $\geq 1,0$ ng/ml</p> <p>BMI 40 kg/m²</p> <p>Ausschlusskriterien</p> <p>Schlecht kontrollierter Diabetes</p> <p>Anamnestisch bekannte diabetische Ketoazidose oder hyperosmolares nicht-ketotisches Koma</p> <p>Insulintherapie innerhalb des letzten Jahres</p> <p>Ein kardiovaskuläres Ereignis in den letzten 6 Monaten</p> <p>Herzinsuffizienz NYHA III/IV</p> <p>LVEF ≤ 40 %</p> <p>Relevante Nieren- oder Lebererkrankung in der Anamnese</p> <p>Psychiatrische Erkrankung</p> <p>Anamn. Alkohol- oder Drogenmissbrauch während des vergangenen Jahres</p> <p>Behandlung mit potenten CYP-Inhibitoren oder -Induktoren.</p> <p>Eingeschränkte Immunkompetenz</p> <p>Aktive Lebererkrankung</p>

Item ^a	Charakteristikum	Studieninformation
		Auffällige Werte in den Tests zur Leber- und Nierenfunktion oder endokrinologischen, metabolischen oder hämatologischen Tests. Compliance in der Run-in Phase ≥ 80 % bis ≤ 120 %
4b	Studienorganisation und Ort der Studiendurchführung	115 Zentren: USA 54 Argentinien 4 Brasilien 10 Mexiko 9 Peru 4 Puerto Rico 4 Hong Kong 2 Korea 8 Phillipinen 3 Singapur 1 Taiwan 3 Israel 8 Südafrika 5
	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	7,5 mg/d Glibenclamid + 2,5 mg/d Saxagliptin 7,5 mg/d Glibenclamid + 5 mg Saxagliptin Placebo + titrierte Dosis Glibenclamid, beginnend mit 10 mg/d; Auftitration war in Woche 2 und Woche 4 erlaubt, Rescue-Medikation 500-2000 mg Metformin
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	Primärer Endpunkt HbA1c-Änderung der Ausgangswerte bis Woche 24 Sekundäre Endpunkte Änderung des Nüchtern BZ vom Ausgangswert bis Woche 24 Anteil Patienten, die bis Woche 24 einen HbA1c-Wert < 7 % aufweisen Veränderungen der postprandialen AUC ₀₋₁₈₀ für Glukose nach OGTT in Woche 24 Weitere Endpunkte Anteil Patienten mit HbA1c-Werten $\leq 6,5$ % Veränderungen der Nüchternwerte und der postprandialen AUC für Insulin, C-Peptid und Glukagon Parameter der β -Zellfunktion (HOMA-2 β) und der Insulinresistenz (HOMA-2IR) Anteil Patienten, die wegen unzureichender therapeutischer Wirkung die Studie

Item ^a	Charakteristikum	Studieninformation
		abbrechen oder eine Rescue-Medikation benötigen
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	keine Angaben
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl beruht auf der Annahme, dass mit einer Power von mindestens 90% ein Unterschied im HbA1c von 0,4 Prozentpunkten detektiert wird
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	keine Angaben zu Zwischenanalysen, kein vorzeitiger Abbruch
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	keine Angaben
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgte im Verhältnis 1 : 1 : 1 via <i>Interactive Voice Response System (IVRS)</i> im Double Dummy Design und in Dreierblocks (block size 3)
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	keine weiteren Angaben
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 erfolgte im Verhältnis 1 : 1 : 1 via IVRS im Double Dummy Design und in Dreierblocks (<i>block size 3</i>)
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?	keine weiteren Angaben
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht relevant
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Für den primären Endpunkt wurde jeder Vergleich zwischen der Saxagliptin-Gruppe und der Glibenclamid-Gruppe auf 0,027- α -Niveau (Dunnett's <i>Adjustment</i>) durchgeführt, so dass der Typ-1 Fehler auf dem Signifikanzniveau von 0,05 kontrolliert war. ANCOVA (LOCF) mit der Behandlungsgruppe als Effektgröße und Ausgangswerten als Kovariable.

Item ^a	Charakteristikum	Studieninformation
		Für die Veränderungen der Saxagliptingruppen gegenüber der Glibenclamidgruppe 95% Konfidenzintervall Gruppenvergleich mittels zweiseitigem Fisher Exact Test
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Inzidente Variable wurden mit deskriptiver Statistik ausgewertet und mittels nicht-parametrischem, nach Kovariablen adjustiertem Mantel-Haenszel-Test verglichen. Die Homöostaasemodelle zur β -Zellfunktion und Veränderungen von Variablen des Lipidstoffwechsels wurden im ANCOVA-Modell untersucht.
-	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 berücksichtigt wurden	2040 Patienten im Screening 959 Patienten in der <i>Run-in</i> Phase 768 Patienten randomisiert: 248 2,5 mg Saxagliptin 253 5 mg Saxagliptin 267 Placebo 621 Patienten Studie abgeschlossen: 192 2,5 mg Saxagliptin 253 5 mg Saxagliptin - 176 Placebo
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	2,5 mg Saxagliptin Abbrüche 56 Mangelnde Wirkung 41 Einverständnis widerrufen 7 <i>Lost to follow-up</i> 5 Schwangerschaft 1 UE 1 Mangelnde Compliance 1 5 mg Saxagliptin Abbrüche 58 Mangelnde Wirkung 38 UE 6 Einverständnis widerrufen 5 <i>Lost to follow-up</i> 4 sonstige Gründe 2 Schwangerschaft 1 Mangelnde Compliance 1 Studienkriterien treffen nicht mehr zu 1

Item ^a	Charakteristikum	Studieninformation
		Placebo 91 Abbrüche Mangel an Wirkung 66 Einverständnis widerrufen 14 UE 4 <i>Lost to follow-up</i> 3 sonstige Gründe 2 Tod 1 mangelnde Compliance 1
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Rekrutierung: 17. April 2006 bis 2. Februar 2007 Follow-up bis 14. September 2007 Ende der Studie September 2008
14b	Informationen, warum die Studie endete oder beendet wurde	
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

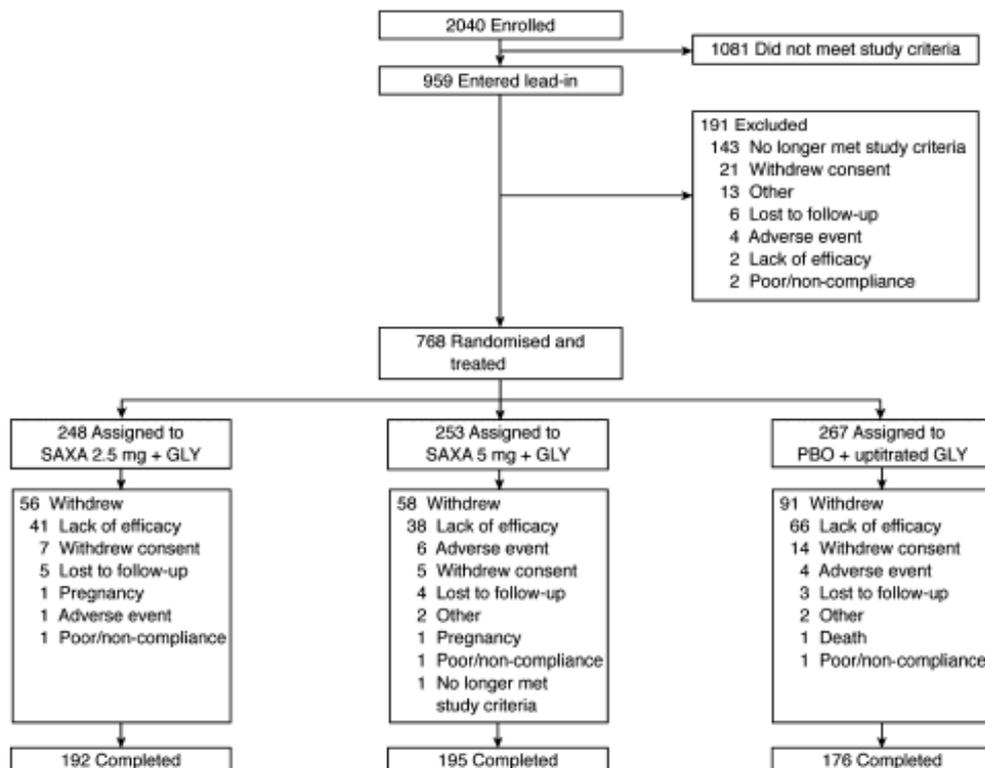


Abbildung 24: Flussdiagramm Studiendesign und-methodik für Chacra 2009

Tabelle 4-155 (Anhang): Studiendesign und -methodik für Studie Garber et al. 2008

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Bestimmung von Wirksamkeit und Verträglichkeit von Vildagliptin im Vergleich zu Placebo in Kombination mit 4mg Glimepirid 1 x täglich
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multinationale, randomisierte, doppelblinde, placebokontrollierte Studie mit vierwöchiger Run-in Phase mit 4 mg Glimepirid 1 x tgl.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	keine Angabe
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <p>18 – 80 Jahre BMI 22 – 45 kg/m² Typ 2 Diabetes Behandlung mit Glibenclamid $\geq 7,5$mg, ≥ 2mg Glimepirid oder einer äquipotenten Dosis Glipizid seit mindestens 3 Monaten stabile Dosis OAD seit mind. vier Wochen vor Screening-Visite HbA1c 7,5 – 11% Nüchtern BZ <15 mmol/l</p> <p>Ausschlusskriterien</p> <p>Typ-1 Diabetes oder sekundäre Diabetesform Myokardinfarkt, Instabile Angina oder CABG in den letzten sechs Monaten Herzinsuffizienz NYHA III/IV Schwere Lebererkrankung (Zirrhose, chronisch aktive Hepatitis) Andere OAD als SU in den letzten zwei Monaten erhöhte Werte für ALT oder AST (>3 x ULN) direktes Bilirubin (>1,3 x ULN) Serumkreatinin (>220 μmol/l) TSH Nüchterntriglyzeride (>7,9 mmol/l)</p>
4b	Studienorganisation und Ort der Studiendurchführung	Multinationale Studie in 114 Prüfzentren, davon 88 in USA,

Item ^a	Charakteristikum	Studieninformation
		9 in Schweden, 7 in Finnland, 6 in Argentinien 4 in Litauen
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	4 mg Glimepirid 1 x tgl. + 50 mg Vildagliptin 1 x tgl. 4 mg Glimepirid 1 x tgl. + 50 mg Vildagliptin 2 x tgl. 4 mg Glimepirid 1 x tgl. + Placebo
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	Primärer Endpunkt mittlere HbA1c-Änderung der Ausgangswerte bis Woche 24 (LOCF bei vorzeitigem Abbruch) Sekundäre Endpunkte Änderung der Nüchternblutzuckerwerte bis Woche 24 Anteil Patienten mit HbA1c <7% bis Woche 24 Anteil Patienten mit HbA1c-Reduktion um $\geq 0,7\%$ -Punkte bis Woche 24 Plasmalipide (nüchtern) Körpergewicht Sicherheitsprofil nach 24 Wochen HbA1c-Änderung bis Woche 24 bei Patienten mit hohen im Vergleich zu Patienten mit niedrigen HbA1c-Ausgangswerten in Subgruppe (ca.16%): Glukose, Insulin und C-Peptid im Verlauf bis 240 Minuten postprandial nach Testmahlzeit
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	keine Angaben
7	Fallzahl	515 Patienten randomisiert 408 Patienten als ITT-Population ausgewertet 332 Patienten Studie abgeschlossen Vildagliptin 50 mg: 170 randomisiert, 132 auswertbar, 113 Studie abgeschlossen Vildagliptin 100 mg 169 randomisiert, 132 auswertbar, 111 Studie abgeschlossen Placebo 176 randomisiert, 144 auswertbar, 108 Studie abgeschlossen
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl wurde so berechnet, dass eine klinisch relevante Differenz im HbA1c von 0,5 Prozentpunkten mit einer Power von

Item ^a	Charakteristikum	Studieninformation
		mindestens 90% bei einer <i>Dropout</i> -Rate von 10% detektiert werden konnte. Bei einer Standardabweichung von 1 und einem Signifikanzniveau von 5% wurde die Überlegenheit von Vildagliptin angenommen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	keine Angaben zu Zwischenanalysen, kein vorzeitiger Abbruch
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Randomisierung 1 : 1 : 1; Die Randomisierungsliste wurde nach einem von der zuständigen Gesundheitsbehörde geprüften, validierten System zur automatischen Randomisierung erstellt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	keine Angaben
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, Ärzte, Untersucher, Laborpersonal hatten keinen Zugang zu Randomisierungsdaten
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?	Randomisierung 1 : 1 : 1; Die Randomisierungsliste wurde nach einem von der zuständigen Gesundheitsbehörde geprüften, validierten System zur automatischen Randomisierung erstellt.
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?	doppelblind
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht relevant
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Kovariate Analyse; Klassifikationsvariable: Behandlung, Zentrum Kovariable: Ausgangswert Zweiseitige Tests, Signifikanzniveau 0,05 Für HbA1c und NBZ Mehrfachtestung: Step-up Prozedur nach Hochberg Vergleich der Anteile Patienten mit HbA1c < 7% mit χ^2 Test
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Präspezifizierte Subgruppen: - Patienten mit Ausgangs-HbA1c

Item ^a	Charakteristikum	Studieninformation
		≤9% >9% - Alter <65 Jahre ≥65 Jahre
-	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 berücksichtigt wurden	Screening: 1050 Randomisiert: 515 Ausgewertet: 408 Vildagliptin 50 mg: 132 Vildagliptin 100 mg: 132 Placebo: 144 Beendet: 332 Vildagliptin 50 mg: 113 Vildagliptin 100 mg: 111 Placebo: 108
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Vildagliptin 50 mg: UE: 4 Abweichende Laborwerte: 0 Mangelnder Therapieeffekt: 6 Protokollverletzung: 2 Einwilligung widerrufen: 4 Lost to Follow-up: 2 Administrativ: 1 Vildagliptin 100 mg: UE: 4 Abweichende Laborwerte: 1 Mangelnder Therapieeffekt: 5 Protokollverletzung: 0 Einwilligung widerrufen: 4 Lost to Follow-up: 26 Administrativ: 1 Placebo: UE: 2 Abweichende Laborwerte: 0 Mangelnder Therapieeffekt: 16 Protokollverletzung: 2 Einwilligung widerrufen: 12 Lost to Follow-up: 4 Administrativ: 0
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Mai 2004 bis Oktober 2005

Item ^a	Charakteristikum	Studieninformation
14b	Informationen, warum die Studie endete oder beendet wurde	keine Angaben

a: nach CONSORT 2010.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

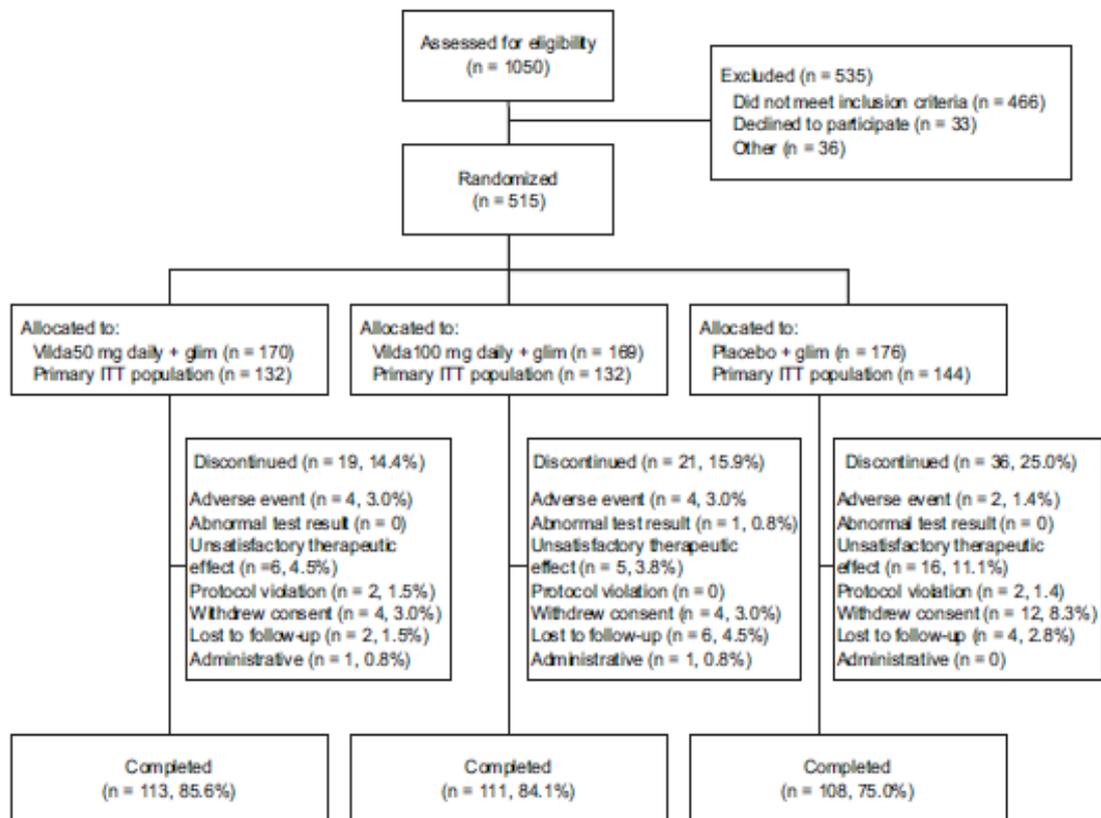


Abbildung 25: Flussdiagramm Studiendesign und-methodik für Garber 2008

Tabelle 4-156 (Anhang): Studiendesign und -methodik für Studie Pratley et al. 2009

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Bewertung von Wirksamkeit und Sicherheit des selektiven oralen DDP-4-Hemmers Alogliptin, zusätzlich gegeben zu einer Glibenclamid-Monotherapie
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, doppelblinde, placebokontrollierte, multinationale Studie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	keine Angabe
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <p>18 bis 80 Jahre Typ-2 Diabetes HbA1c 7,0-10% unter SU-Monotherapie Nüchternblutzuckerwerte <275 mg/dl, HbA1c 7,0-10,0%, Compliance \geq75% und seit 8 Wochen stabile SU-Dosis am Ende der Run-in Phase Behandlung mit einem Sulfonylharnstoff seit mindestens drei Monaten BMI 23-45 kg/m² Blutdruck \leq180/110 Hb \geq12 g/dl (Männer) bzw. \geq10g/dl (Frauen) Normale Leber- und Nierenfunktion Normale Schilddrüsenfunktion Zur BZ-Selbstmessung bereit und in der Lage</p> <p>Ausschlusskriterien</p> <p>Andere OAD als SU in den vergangenen drei Monaten Serumkreatinin >2,0 mg/dl Albumin/Kreatinin-Quotient im Urin >1000 μg/mg Laserbehandlung einer proliferativen diabetischen Retinopathie in den vergangenen sechs Monaten Anamnestisch bekannte Behandlung einer diabetischen Gastroparese Herzinsuffizienz NYHA III/IV Angioplastik, Koronarstent, CABG oder Myokardinfarkt in den letzten sechs Monaten Einnahme einer <i>investigational drug</i> oder Teilnahme an einer anderen Alogliptin-Studie</p>

Item ^a	Charakteristikum	Studieninformation
		in den letzten 30 Tagen Einnahme von Wirkstoffen zur Gewichtsreduktion, von Bosentan oder systemischen Glukokortikoiden (oral oder parenteral) Hepatitis B oder C-Infektion, HIV Infektion Malignom Hämoglobinopathie
4b	Studienorganisation und Ort der Studiendurchführung	124 Zentren in 16 Ländern: 54 USA weitere Zentren in Argentinien, Brasilien, Chile, Dominikanische Republik, Guatemala, Mexiko, Peru Australien, Neuseeland Indien Niederlande, Polen, UK Südafrika
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	1 x tgl. 12,5 mg Alogliptin + Glibenclamid* 1 x tgl. 25 mg Alogliptin + Glibenclamid* Placebo + Glibenclamid* *die Glibenclamid-Dosis wurde in einer 4-wöchigen <i>Run-in</i> Phase ermittelt und war ≤10 mg täglich
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	Primärer Endpunkt HbA1c-Änderung der Ausgangswerte bis Woche 26 Sekundäre Endpunkte HbA1c-Änderungen bis Woche 4, 8, 12, 16 und 20 Änderung von NBZ bis Woche 1, 2, 4, 8, 12, 16, 20 und 26 Anteil Patienten mit Hyperglykämien (NBZ ≥ 200 mg/dl) Änderungen der Proinsulinwerte bis Woche 4, 8, 12, 16, 20, 26 Änderungen der Insulinwerte bis Woche 4, 8, 12, 16, 20, 26 Änderungen des Proinsulin/Insulin-Quotienten bis Woche 4, 8, 12, 16, 20, 26 Änderungen der C-Peptidwerte bis Woche 4, 8, 12, 16, 20, 26 Anteile Patienten mit HbA1c ≤ 6,5 % / ≤ 7,0 % / ≤ 7,5 % Anteil Patienten mit HbA1c-Reduktion um mehr als 0,5 / 1,0 / 1,5 / 2,0 Prozentpunkte Änderung des Körpergewichts bis Woche 8, 12, 20 und 26

Item ^a	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	keine Angaben
7	Fallzahl	585 Patienten im Screening 500 Patienten randomisiert: - 203 12,5 mg Alogliptin - 198 25 mg Alogliptin - 99 Placebo 363 Patienten Studie abgeschlossen: - 153 12,5 mg Alogliptin - 148 25 mg Alogliptin - 62 Placebo
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahlberechnung beruhte auf der Annahme einer mittleren HbA1c- Differenz zwischen Verum und Placebo von 0,4 Prozentpunkten in Woche 26 (T-Test, Standardabweichung 0,8%, zweiseitiges Signifikanzniveau 0,05, Drop-out Rate unter 80%)
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	keine Angaben zu Zwischenanalysen, kein vorzeitiger Abbruch
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	keine Angaben
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgte im Verhältnis 2 : 2 : 1 (Alogliptin 12,5 mg, Alogliptin 25 mg, Placebo) und war stratifiziert nach dem HbA1c-Wert in Woche 1 (<8% vs. ≥8%) und geographischer Region
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	doppelblind
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?	2 : 2 : 1 Randomisierung zu 12,5 oder 25 mg Alogliptin oder Placebo 1 x täglich nach einem <i>Permuted Block</i> Schema
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?	Patient, Pflegepersonen, Prüfarzte und auswertende Personen waren verblindet
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht relevant

Item ^a	Charakteristikum	Studieninformation												
12	Statistische Methoden													
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Der Gruppenunterschied für den primären Endpunkt wurde mittels ANCOVA mit geographischer Region, Ausgangs-HbA1c-Werten und Anfangsdosis Glibenclamid als Kovariable ermittelt.</p> <p>Nach Testung auf Signifikanz im primären Endpunkt (zweiseitiges Signifikanzniveau 0,05 für 25 mg gegen Placebo) wurde für 12,5 mg ebenso getestet.</p> <p>Für die kontinuierlichen Messungen im Studienverlauf wurde wie für HbA1c getestet, jedoch mit den jeweiligen Ausgangswerten als Kovariable</p>												
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Inzidente Variable wurden mit deskriptiver Statistik ausgewertet und mittels nicht-parametrischen, nach Kovariablen adjustierten Mantel-Haenszel-Test verglichen.</p> <p>Die Homöostaasemodelle zur β-Zellfunktion und Veränderungen von Variablen des Lipidstoffwechsels wurden im ANCOVA-Modell untersucht.</p>												
-	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 berücksichtigt wurden	<p>Screening: 585</p> <p>Randomisiert: 500</p> <table> <tr> <td>12,5 mg Alogliptin</td> <td>203</td> </tr> <tr> <td>25 mg Alogliptin</td> <td>198</td> </tr> <tr> <td>Placebo</td> <td>99</td> </tr> </table> <p>Beendet: 363</p> <table> <tr> <td>12,5 mg Alogliptin</td> <td>153</td> </tr> <tr> <td>25 mg Alogliptin</td> <td>148</td> </tr> <tr> <td>Placebo</td> <td>62</td> </tr> </table>	12,5 mg Alogliptin	203	25 mg Alogliptin	198	Placebo	99	12,5 mg Alogliptin	153	25 mg Alogliptin	148	Placebo	62
12,5 mg Alogliptin	203													
25 mg Alogliptin	198													
Placebo	99													
12,5 mg Alogliptin	153													
25 mg Alogliptin	148													
Placebo	62													
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>12,5 mg Alogliptin</p> <p>Rescuebedürftige Hyperglykämie 31, UE 4, Protokollverletzung 1, <i>lost to follow-up</i> 2, Patientenwunsch 11, Prüfarztentscheidung 1</p> <p>25 mg Alogliptin</p> <p>Rescuebedürftige Hyperglykämie 30, UE 6, Protokollverletzung 3, <i>lost to follow-up</i> 1, Patientenwunsch 8, Prüfarztentscheidung 2</p> <p>Placebo</p> <p>Rescuebedürftige Hyperglykämie 28, UE 2, <i>lost to follow-up</i> 1, Patientenwunsch 3, Prüfarztentscheidung 3</p>												
14	Aufnahme / Rekrutierung													

Item ^a	Charakteristikum	Studieninformation
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	April 2006 bis Juni 2007
14b	Informationen, warum die Studie endete oder beendet wurde	Juni 2007

a: nach CONSORT 2010.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

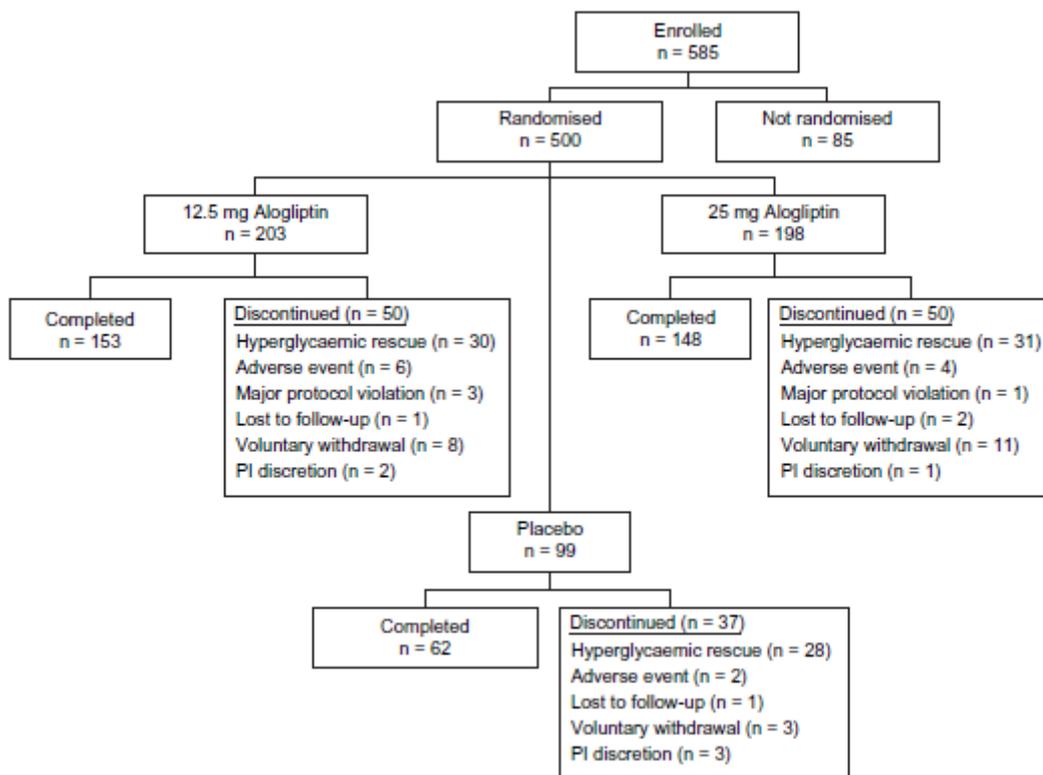


Abbildung 26: Flussdiagramm Studiendesign und-methodik für Pratley 2008

Tabelle 4-157 (Anhang): Studiendesign und -methodik für Studie Lewin et al. 2012

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Wirksamkeit und Sicherheit von BI 1356 im Vergleich zu Placebo bei Patienten mit Typ 2 Diabetes und unzureichender glykämischer Kontrolle unter der Behandlung mit einem Sulfonylharnstoff
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multinationale, randomisierte, doppelblinde, placebokontrollierte Studie mit zweiwöchigem Placebo <i>Run-in</i>
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	keine Angabe
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <p>18 – 80 Jahre Typ 2 Diabetes HbA1c 7 – 10%</p> <p>Ausschlusskriterien</p> <p>Myokardinfarkt, Schlaganfall oder TIA in den letzten sechs Monaten Behandlung mit Rosiglitazon, Pioglitazon, GLP-1-Analoga, Insulin oder einem Mittel zur Gewichtsreduktion in den vergangenen 3 Monaten Eingeschränkte Leberfunktion Schwere Nierenfunktionsstörung Behandlung mit systemischen Steroiden Veränderung in der Dosierung eines Schilddrüsenhormons Hereditäre Galaktose-Intoleranz</p>
4b	Studienorganisation und Ort der Studiendurchführung	Multinationale Studie in 45 Prüfzentren: USA 10 Argentinien 4 Russland 7 Polen 5 Ungarn 5 Indien 8 Japan 3
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	5 mg Linagliptin 1 x tgl. Placebo jeweils zusätzlich zur bisherigen SU-Medikation in konstanter Dosis

Item ^a	Charakteristikum	Studieninformation
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>Primärer Endpunkt mittlere HbA1c-Änderung der Ausgangswerte bis Woche 18</p> <p>Sekundäre Endpunkte Änderung der Nüchternblutzuckerwerte bis Woche 18 Anteil Patienten mit HbA1c <7% bis Woche 18 Anteil Patienten mit HbA1c <6,5% bis Woche 18 Anteil Patienten mit HbA1c-Reduktion um $\geq 0,5\%$-Punkte bis Woche 18 Mixed Model Repeated Measurement Analysis zur HbA1c-Änderung bis Woche 6 Woche 12 Woche 18 Änderung des NBZ bis Woche 6 Woche 12</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	keine Angaben
7	Fallzahl	245 Patienten in die Studie aufgenommen 228 Patienten Studie abgeschlossen
7a	Wie wurden die Fallzahlen bestimmt?	keine Angaben
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	keine Angaben zu Zwischenanalysen, kein vorzeitiger Abbruch
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Randomisierung 2 : 1 Linagliptin 5 mg: Placebo
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	keine Angaben
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	keine Angaben
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?	Randomisierung 2 : 1 Linagliptin 5 mg : Placebo
11	Verblindung	

Item ^a	Charakteristikum	Studieninformation
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?	doppelblind
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht relevant
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Kovarianzanalyse (ANCOVA) adjustiert nach Behandlungsarm, früheren OAD und HbA1c-Ausgangswerten
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	keine Angaben
-	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 berücksichtigt wurden	Randomisiert: 245 Linagliptin 161 Placebo 84 Abgeschlossen: 228 Linagliptin: 151 Placebo: 77
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Linagliptin 10 UE: 5 Protokollverletzung: 1 <i>Lost to Follow-up</i> : 0 Patiententenwunsch: 2 keine Angaben zum Abbruchgrund 2 Placebo 7 UE: 3 Protokollverletzung: 0 <i>Lost to Follow-up</i> : 2 Patiententenwunsch: 1 keine Angaben zum Abbruchgrund 1
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Dezember 2008 bis Januar 2010
14b	Informationen, warum die Studie endete oder beendet wurde	keine Angaben
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

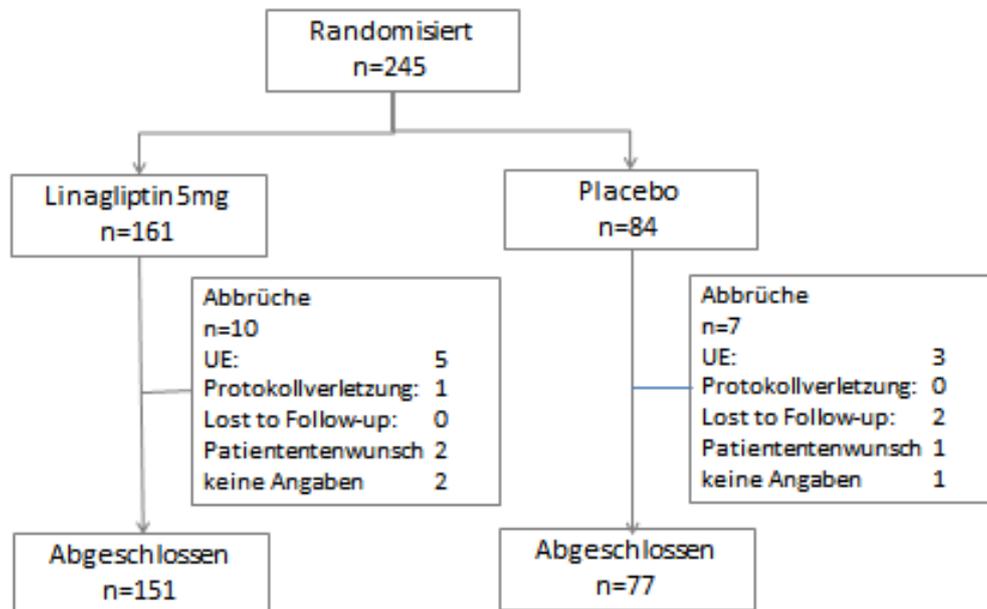


Abbildung 27: Flussdiagramm Studiendesign und-methodik für Lewin 2012

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 4-158 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D1690C00005

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: D1690C00005

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht (AstraZeneca, 2010a)	CSR05
Publikation (Strojek et al., 2011)	PubStrojek2011

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. (CSR05)

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 verblindete Behandlung erfolgte anhand einer *double-dummy* Technik;
Prüf- und Vergleichsmedikation waren identisch (CSR05).

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 (CSR05).

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 (CSR05).

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: 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 (CSR05).

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: 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 (CSR05).

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 (CSR05).

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:
Ungleiche Abbrecherraten zwischen den Gruppen (CSR05)

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 (CSR05).

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:
Unklare Therapiesituation der Patienten mit Hypertonie (CSR05)

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 (CSR05).

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 (CSR05).

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 4-159 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Publikation von DeFronzo 1995

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 1995

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Publikation (DeFronzo und Goodman, 1995)- (Protokoll 2)	PubDeFronzo1995
Studienbericht (Lipha Pharmaceutical Inc., 1992)	CSRDeFronzo1995

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 („allocationconcealment“)

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. alphaspendingapproach 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

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: 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

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 („fullanalysisset“). 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: 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

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“-). 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 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 4-160 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Goldstein 2003

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: Goldstein 2003

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Publikation (Goldstein et al., 2003)	PubGoldstein2003

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 („allocationconcealment“)

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:

Angaben aus PubDeFronzo1995.

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:

Angaben aus PubDeFronzo1995.

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubDeFronzo1995).

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. alphaspendingapproach 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

Laborwerte wurden ohne Kenntnis der Allokation bestimmt (PubDeFronzo1995).

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 („intentiontotreat“)-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 („fullanalysisset“). 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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubDeFronzo1995).

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

Einstufung erfolgte ohne Kenntnis der Behandlungsallokation (PubDeFronzo1995).

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubDeFronzo1995).

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

Bei Bestimmung des Gewichtes bestand keine Kenntnis über die Allokation (PubDeFronzo1995).

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 („fullanalysisset“). 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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden.

Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubDeFronzo1995).

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

Unerwünschte Ereignisse wurden verblindet erhoben. Ein unerwünschte Entblindung erscheint unwahrscheinlich (PubDeFronzo1995).

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“). 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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine adäquate Umsetzung des ITT-Prinzips nicht überprüft werden (PubDeFronzo1995).

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubDeFronzo1995).

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 4-161 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Hermansen K et al. 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: Hermansen 2007

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Publikation (Hermansen et al., 2007)	PubHermansen2007

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 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:

Laut Angaben in der Publikation zur Studie (PubHermansen2007)

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:

Angaben in der Publikation zur Studie (PubHermansen2007)

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubHermansen2007).

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: primärer Endpunkt HbA1c-Änderung

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

Alle Laborergebnisse wurden zentral und verblindet ermittelt (PubHermansen2007)

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:

Differenzierte Darstellung der Abbruchgründe unter 13b und in der Darstellung des Patientenflusses

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubHermansen2007).

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

Alle Angaben zu Hypoglykämien wurden verblindet ermittelt (PubHermansen2007)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubHermansen2007).

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

Alle Angaben zu Körpergewicht wurden verblindet ermittelt (PubHermansen2007)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubHermansen2007).

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

Alle Angaben zu unerwünschten Ereignissen wurden verblindet ermittelt
(PubHermansen2007)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubHermansen2007).

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 4-162 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Chacra AR et al. 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: Chacra 2009

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Publikation (Chacra et al., 2009)	PubChacra2009

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:

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:

Laut Angaben in der Publikation zur Studie (PubChacra2009)

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:

Angaben in der Publikation zur Studie (PubChacra2009)

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubChacra2009).

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 existieren keine Hinweise auf sonstige, endpunktübergreifende Aspekte, die das Verzerrungspotenzial beeinflussen (PubChacra2009)

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: primärer Endpunkt HbA1c-Änderung

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

Alle Laborergebnisse wurden zentral und verblindet ermittelt (PubChacra2009)

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:

Differenzierte Darstellung der Abbruchgründe unter 13b und in der Darstellung des Patientenflusses

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubChacra2009).

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

Alle Angaben zu Hypoglykämien wurden verblindet ermittelt (PubChacra2009)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubChacra2009).

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

Alle Angaben zu Körpergewicht wurden verblindet ermittelt (PubChacra2009)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubChacra2009).

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

Alle Angaben zu unerwünschten Ereignissen wurden verblindet ermittelt
(PubChacra2009)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubChacra2009).

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 4-163 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Garber AJ et al. 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: Garber 2008

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Publikation (Garber et al., 2008)	PubGarber2008

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 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:

Laut Angaben in der Publikation zur Studie (PubGarber2008)

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:

Angaben in der Publikation zur Studie (PubGarber2008)

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubGarber2008).

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 existieren keine Hinweise auf sonstige, endpunktübergreifende Aspekte, die das Verzerrungspotenzial beeinflussen (PubGarber2008)

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: primärer Endpunkt HbA1c-Änderung

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

Alle Laborergebnisse wurden zentral und verblindet ermittelt (PubGarber2008)

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:

Differenzierte Darstellung der Abbruchgründe unter 13b und in der Darstellung des Patientenflusses (PubGarber2008)

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubGarber2008).

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

Alle Angaben zu Hypoglykämien wurden verblindet ermittelt (PubGarber2008)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubGarber2008).

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

Alle Angaben zu Körpergewicht wurden verblindet ermittelt (PubGarber2008)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubGarber2008).

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

Alle Angaben zu unerwünschten Ereignissen wurden verblindet ermittelt
(PubGarber2008)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubGarber2008).

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 4-164 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Pratley et al. 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: Pratley 2009

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Publikation (Pratley et al., 2009)	PubPratley2009

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:

Laut Angaben in der Publikation zur Studie (PubPratley2009)

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:

Angaben in der Publikation zur Studie (PubPratley2009)

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubPratley2009).

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 existieren keine Hinweise auf sonstige, endpunktübergreifende Aspekte, die das Verzerrungspotenzial beeinflussen (PubPratley2009)

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: primärer Endpunkt HbA1c-Änderung

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

Alle Laborergebnisse wurden zentral und verblindet ermittelt (PubPratley2009)

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:

Differenzierte Darstellung der Abbruchgründe unter 13b und in der Darstellung des Patientenflusses

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubPratley2009).

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

Alle Angaben zu Hypoglykämien wurden verblindet ermittelt (PubPratley2009)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubPratley2009).

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

Alle Angaben zu Körpergewicht wurden verblindet ermittelt (PubPratley2009)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubPratley2009).

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

Alle Angaben zu unerwünschten Ereignissen wurden verblindet ermittelt
(PubPratley2009)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubPratley2009).

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 4-165 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Lewin 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: Lewin 2012

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Lewin (Lewin et al., 2012)	PubLewin2012

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:

Es gibt keine weiteren Angaben zum Randomisierungsverfahren, vermutlich handelte es sich aber um ein zentralisiertes Verfahren; daher keine Abwertung der Gesamtbewertung der Studie (PubLewin2012).

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:

Laut Angaben in der Publikation zur Studie (PubLewin2012)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubLewin2012).

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 existieren keine Hinweise auf sonstige, endpunktübergreifende Aspekte, die das Verzerrungspotenzial beeinflussen (PubLewin2012)

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: primärer Endpunkt HbA1c-Änderung

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

Alle Laborergebnisse wurden zentral und verblindet ermittelt (PubLewin2012)

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:

Differenzierte Darstellung der Abbruchgründe unter 13b und in der Darstellung des Patientenflusses

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 liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubLewin2012).

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

Alle Angaben zu Hypoglykämien wurden verblindet ermittelt (PubLewin2012)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubLewin2012).

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

Alle Angaben zu unerwünschten Ereignissen wurden verblindet ermittelt
(PubLewin2012)

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:

Es liegt nur die Publikation der Studie vor (kein Studienbericht), dadurch kann eine ergebnisunabhängige Berichterstattung nicht abschließend überprüft werden. Hinweise auf eine ergebnisunabhängige Berichterstattung fanden sich nicht (PubLewin2012).

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: