

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Saxagliptin (Onglyza[®] 2,5 / 5 mg Filmtabletten)
Bristol-Myers Squibb GmbH & Co KGaA und
AstraZeneca GmbH

Modul 4 B

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

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	11
Abkürzungsverzeichnis	13
4 Modul 4 – allgemeine Informationen	15
4.1 Zusammenfassung der Inhalte von Modul 4.....	16
4.2 Methodik.....	29
4.2.1 Fragestellung.....	29
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	31
4.2.3 Informationsbeschaffung.....	34
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	34
4.2.3.2 Bibliografische Literaturrecherche.....	35
4.2.3.3 Suche in Studienregistern.....	37
4.2.3.4 Selektion relevanter Studien.....	38
4.2.4 Bewertung der Aussagekraft der Nachweise.....	38
4.2.5 Informationssynthese und -analyse.....	41
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	41
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	41
4.2.5.3 Meta-Analysen.....	45
4.2.5.4 Sensitivitätsanalysen.....	48
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	49
4.2.5.6 Indirekte Vergleiche.....	50
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	52
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	53
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	53
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	53
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	55
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	57
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	58
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	59
4.3.1.2.1 Studiendesign und Studienpopulationen.....	59
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	63
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	64
4.3.1.3.1 Hypoglykämien unter Berücksichtigung des HbA1c-Wertes – RCT.....	65
4.3.1.3.1.1 Ausmaß der Blutzuckerkontrolle (HbA1c) – RCT.....	65
4.3.1.3.1.2 Hypoglykämien – RCT.....	70
4.3.1.3.1.1 Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien.....	74
4.3.1.3.1.2 Gewichtsveränderung – RCT.....	77

4.3.1.3.1.3	Unerwünschte Ereignisse – RCT.....	80
4.3.1.3.1.4	Kardiale Ereignisse – RCT	83
4.3.1.3.2	Subgruppenanalysen – RCT.....	85
4.3.1.3.2.1	Ausmaß der Blutzuckerkontrolle (HbA1c), stratifiziert nach dem Alter – RCT	87
4.3.1.3.2.2	Ausmaß der Blutzuckerkontrolle (HbA1c), stratifiziert nach dem Geschlecht – RCT.....	89
4.3.1.3.2.3	Hypoglykämien, stratifiziert nach dem Alter – RCT	91
4.3.1.3.2.4	Hypoglykämien, stratifiziert nach dem Geschlecht – RCT.....	93
4.3.1.3.2.5	Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien, stratifiziert nach dem Alter – RCT	95
4.3.1.3.2.6	Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien, stratifiziert nach dem Geschlecht – RCT.....	97
4.3.1.3.2.7	Gewichtsveränderung stratifiziert nach dem Alter – RCT	98
4.3.1.3.2.8	Gewichtsveränderung, stratifiziert nach dem Geschlecht – RCT....	99
4.3.1.3.2.9	Unerwünschte Ereignisse, stratifiziert nach dem Alter – RCT	100
4.3.1.3.2.10	Unerwünschte Ereignisse, stratifiziert nach dem Geschlecht – RCT	101
4.3.1.3.2.11	Kardiale Ereignisse stratifiziert nach dem Alter – RCT.....	102
4.3.1.3.2.12	Kardiale Ereignisse stratifiziert nach dem Geschlecht – RCT	103
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	103
4.3.2	Weitere Unterlagen.....	106
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	106
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	106
4.3.2.1.1.1	Studien des pharmazeutischen Unternehmers	106
4.3.2.1.1.2	Studien aus der bibliografischen Literaturrecherche	108
4.3.2.1.1.3	Studien aus der Suche in Studienregistern	112
4.3.2.1.1.4	Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel	113
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	116
4.3.2.1.2.1	Studiendesign und Studienpopulationen	116
4.3.2.1.2.2	Verzerrungspotenzial auf Studienebene	129
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	130
4.3.2.1.3.1	Ausmaß der Blutzuckerkontrolle (HbA1c) – RCT.....	132
4.3.2.1.3.1.1	Hypoglykämien – RCT.....	142
4.3.2.1.3.1.2	Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien	151
4.3.2.1.3.1.3	Gewichtsveränderung – RCT	152
4.3.2.1.3.1.4	Unerwünschte Ereignisse – RCT.....	162
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	170
4.3.2.2	Nicht randomisierte vergleichende Studien.....	170
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	170
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	170
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	171
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	171

4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	172
4.3.2.3	Weitere Untersuchungen.....	172
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	172
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	173
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	173
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	173
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	174
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen.....	174
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	177
4.4.1	Beurteilung der Aussagekraft der Nachweise	177
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	179
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	186
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	187
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	188
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	188
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	188
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	189
4.5.4	Verwendung von Surrogatendpunkten	189
4.6	Liste der eingeschlossenen Studien.....	190
4.7	Referenzliste.....	191
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		195
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		213
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund		223
Anhang 4-D : Liste der abgebrochenen Studien		233
Anhang 4-E : Liste der laufenden Studien.....		234
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....		235
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		287

Tabellenverzeichnis

	Seite
Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen	19
Tabelle 4-2: Zusammenfassende Darstellung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit	27
Tabelle 4-3: Prädefinierte Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen	33
Tabelle 4-4: Methodik zur gemeinsamen Betrachtung der Endpunkte Blutzuckerkontrolle (HbA1c) und Hypoglykämien	44
Tabelle 4-5: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	54
Tabelle 4-6: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel	55
Tabelle 4-7: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	58
Tabelle 4-8: Studienpool – RCT mit dem zu bewertenden Arzneimittel	59
Tabelle 4-9: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	60
Tabelle 4-10: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel	61
Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel	62
Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel	62
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel	62
Tabelle 4-14: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel	64
Tabelle 4-15: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	64
Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	65
Tabelle 4-17: Operationalisierung von Ausmaß der Blutzuckerkontrolle (HbA1c)	66
Tabelle 4-18: Bewertung des Verzerrungspotenzials für Ausmaß der Blutzuckerkontrolle (HbA1c) in RCT mit dem zu bewertenden Arzneimittel	66
Tabelle 4-19: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel	68

Tabelle 4-20: Ergebnisse für Studienabbrüche (aufgrund von unzureichender Blutzuckerkontrolle) oder die Einnahme der Notfallmedikation (aufgrund unzureichender Blutzuckerkontrolle verabreicht) aus RCT mit dem zu bewertenden Arzneimittel.....	70
Tabelle 4-21: Operationalisierung von Hypoglykämien.....	71
Tabelle 4-22: Bewertung des Verzerrungspotenzials für Hypoglykämien in RCT mit dem zu bewertenden Arzneimittel	72
Tabelle 4-23: Ergebnisse für patientenberichtete Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel	73
Tabelle 4-24: Ergebnisse für bestätigte Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel.....	73
Tabelle 4-25: Operationalisierung von gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien	75
Tabelle 4-26: Bewertung des Verzerrungspotenzials für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien in RCT mit dem zu bewertenden Arzneimittel.....	75
Tabelle 4-27: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 6,5% ohne Hypoglykämien erreicht haben.....	76
Tabelle 4-28: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 7% ohne Hypoglykämien erreicht haben (exklusive Patienten mit HbA1c <7% zu Baseline).....	76
Tabelle 4-29: Operationalisierung von Gewichtsveränderung.....	77
Tabelle 4-30: Bewertung des Verzerrungspotenzials für Gewichtsveränderung in RCT mit dem zu bewertenden Arzneimittel.....	78
Tabelle 4-31: Ergebnisse für Gewichtsveränderung (Änderung im Vergleich zum Ausgangswert) aus RCT mit dem zu bewertenden Arzneimittel.....	79
Tabelle 4-32: Operationalisierung von Unerwünschte Ereignisse.....	81
Tabelle 4-33: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel	81
Tabelle 4-34: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel.....	82
Tabelle 4-35: Operationalisierung von Kardiale Ereignisse	83
Tabelle 4-36: Bewertung des Verzerrungspotenzials für Kardiale Ereignisse in RCT mit dem zu bewertenden Arzneimittel.....	84
Tabelle 4-37: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel.....	84
Tabelle 4-38: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter	87
Tabelle 4-39: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter	88
Tabelle 4-40: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht.....	89

Tabelle 4-41: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht.....	90
Tabelle 4-42: Ergebnisse für patientenberichtete Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter	91
Tabelle 4-43: Ergebnisse für bestätigte Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter	92
Tabelle 4-44: Ergebnisse für patientenberichtete Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht	93
Tabelle 4-45: Ergebnisse für bestätigte Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht.....	94
Tabelle 4-46: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 6,5% ohne Hypoglykämien erreicht haben aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter.....	95
Tabelle 4-47: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 7% ohne Hypoglykämien erreicht haben (exklusive Patienten mit HbA1c <7% Ausgangswert) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter.....	96
Tabelle 4-48: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 6,5% ohne Hypoglykämien erreicht haben aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht	97
Tabelle 4-49: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 7% ohne Hypoglykämien erreicht haben (exklusive Patienten mit HbA1c < 7% Ausgangswert) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht	97
Tabelle 4-50: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter	98
Tabelle 4-51: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht.....	99
Tabelle 4-52: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter	100
Tabelle 4-53: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht.....	101
Tabelle 4-54: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach dem Alter	102
Tabelle 4-55: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach dem Geschlecht.....	103
Tabelle 4-56: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	107
Tabelle 4-57: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	108
Tabelle 4-58: Relevante Studien aus der Suche in Studienregistern – RCT mit der alternativen Vergleichstherapie.....	113
Tabelle 4-59: Studienpool – RCT für indirekte Vergleiche.....	115

Tabelle 4-60: Charakterisierung der eingeschlossenen Studien – RCT mit der zweckmäßigen bzw. alternativen Vergleichstherapie	117
Tabelle 4-61: Charakterisierung der Interventionen – RCT mit der zweckmäßigen bzw. alternativen Vergleichstherapie	123
Tabelle 4-62: Charakterisierung der Studienpopulationen – RCT mit der zweckmäßigen bzw. alternativen Vergleichstherapie	125
Tabelle 4-63: Charakterisierung der Studienpopulationen – RCT mit der zweckmäßigen bzw. alternativen Vergleichstherapie	126
Tabelle 4-64: Verzerrungspotenzial auf Studienebene – Studien für indirekte Vergleiche...	129
Tabelle 4-65: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	130
Tabelle 4-66: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	131
Tabelle 4-67: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) herangezogen wurden	132
Tabelle 4-68: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden	133
Tabelle 4-69: Operationalisierung von Ausmaß der Blutzuckerkontrolle HbA1c.....	134
Tabelle 4-70: Bewertung des Verzerrungspotenzials für Ausmaß der Blutzuckerkontrolle HbA1c in RCT für indirekte Vergleiche	135
Tabelle 4-71: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel	137
Tabelle 4-72: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) - indirekte Vergleiche	141
Tabelle 4-73: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) - indirekte Vergleiche	141
Tabelle 4-74: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff +Saxagliptin vs. Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) herangezogen wurden	142
Tabelle 4-75: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden	143
Tabelle 4-76: Operationalisierung von Hypoglykämien.....	144
Tabelle 4-77: Bewertung des Verzerrungspotenzials für Hypoglykämien in RCT für indirekte Vergleiche	145
Tabelle 4-78: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche.....	147
Tabelle 4-79: Ergebnisse für Hypoglykämien - indirekte Vergleiche	150
Tabelle 4-80: Ergebnisse für Hypoglykämien (patientenberichtet/bestätigt)- indirekte Vergleiche	151

Tabelle 4-81: Ergebnisse für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel – indirekter Vergleich	152
Tabelle 4-82: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) herangezogen wurden	153
Tabelle 4-83: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden	153
Tabelle 4-84: Operationalisierung von Gewichtsveränderung.....	154
Tabelle 4-85: Bewertung des Verzerrungspotenzials für Gewichtsveränderung in RCT für indirekte Vergleiche	155
Tabelle 4-86: Ergebnisse für Gewichtsveränderung aus RCT für indirekte Vergleiche.....	157
Tabelle 4-87: Ergebnisse für Gewichtsveränderung - indirekte Vergleiche	161
Tabelle 4-88: Ergebnisse für Gewichtsveränderung - indirekte Vergleiche	162
Tabelle 4-89: Ergebnisse für Gewichtsveränderung - indirekte Vergleiche	162
Tabelle 4-90: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden	163
Tabelle 4-91: Operationalisierung von Unerwünschte Ereignisse	164
Tabelle 4-92: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT für indirekte Vergleiche	165
Tabelle 4-93: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel.....	166
Tabelle 4-94: Ergebnisse für Gesamtrate unerwünschter Ereignisse –indirekter Vergleich .	168
Tabelle 4-95: Ergebnisse für schwerwiegende unerwünschte Ereignisse –indirekter Vergleich	169
Tabelle 4-96: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse – indirekter Vergleich.....	169
Tabelle 4-97: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	171
Tabelle 4-98: Operationalisierung von <Endpunkt xxx>.....	171
Tabelle 4-99: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	172
Tabelle 4-100: Operationalisierung von <Endpunkt xxx>.....	173
Tabelle 4-101: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	187
Tabelle 4-102: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	188
Tabelle 4-103: DIMDI Suchstrategie Embase 1947 bis heute: am 04. Februar 2013	197

Tabelle 4-104: DIMDI Suchstrategie MEDLINE 1950 bis heute: am 04. Februar 2013	198
Tabelle 4-105: DIMDI Suchstrategie Cochrane 1948 bis heute: am 04. Februar 2013.....	199
Tabelle 4-106: DIMDI Suchstrategie Embase 1947 bis heute: am 07. Februar 2013	200
Tabelle 4-107: DIMDI Suchstrategie MEDLINE 1950 bis heute: am 07. Februar 2013	202
Tabelle 4-108: DIMDI Suchstrategie Cochrane 1948 bis heute: am 07. Februar 2013.....	204
Tabelle 4-109: DIMDI Suchstrategie Embase 1947 bis heute: am 04. Februar 2013	206
Tabelle 4-110: DIMDI Suchstrategie MEDLINE 1950 bis heute: am 04. Februar 2013	208
Tabelle 4-111: DIMDI Suchstrategie Cochrane 1950 bis heute: am 04. Februar 2013.....	210
Tabelle 4-112: Im Volltext ausgeschlossene nicht relevante Publikationen für das zu bewertende Arzneimittel	223
Tabelle 4-113: Im Volltext ausgeschlossene nicht relevante Publikationen für die zweckmäßige Vergleichstherapie.....	224
Tabelle 4-114: Im Volltext ausgeschlossene nicht relevante Publikationen für die alternative Vergleichstherapie.....	232
Tabelle 4-115 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel.....	233
Tabelle 4-116 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	234
Tabelle 4-117 (Anhang): Studiendesign und -methodik für Studie CV181040.....	236
Tabelle 4-118 (Anhang): Studiendesign und -methodik für Studie Shank 1995	245
Tabelle 4-119 (Anhang): Studiendesign und -methodik für Studie von Birkeland 1994	251
Tabelle 4-120 (Anhang): Studiendesign und -methodik für Studie von Tovi 1998.....	254
Tabelle 4-121 (Anhang): Studiendesign und -methodik für Studie von Nathan 1988.....	258
Tabelle 4-122 (Anhang): Studiendesign und -methodik für die Turner 1998 (UKPDS 24)..	262
Tabelle 4-123 (Anhang): Studiendesign und -methodik für Studie Hermansen et al. 2007 ..	270
Tabelle 4-124 (Anhang): Studiendesign und -methodik für Studie Garber et al. 2008	276
Tabelle 4-125 (Anhang): Studiendesign und -methodik für Studie Lewin 2012.....	282
Tabelle 4-126 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie CV181040	288
Tabelle 4-127 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Shank 1995	306
Tabelle 4-128 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Birkeland 1994	319
Tabelle 4-129 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Tovi 1998	334
Tabelle 4-130 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Nathan 1988	348
Tabelle 4-131 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Turner 1998 (UKPDS 24)	364

Tabelle 4-132 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Hermansen K et al. 2007	380
Tabelle 4-133 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Garber AJ et al. 2008	398
Tabelle 4-134 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Lewin et al. 2012	418

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	57
Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit der zweckmäßigen Vergleichstherapie	110
Abbildung 3: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit der alternativen Vergleichstherapie	112
Abbildung 4: Meta-Analyse der eingeschlossenen Studien zum Effekt von Insulin vs. Sulfonylharnstoff auf den Endpunkt Ausmaß der Blutzuckerkontrolle.....	139
Abbildung 5: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt Ausmaß der Blutzuckerkontrolle	140
Abbildung 6: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt Hypoglykämien (patientenberichtet).....	150
Abbildung 7: Meta-Analyse der eingeschlossenen Studien zum Effekt von Insulin vs. Sulfonylharnstoff auf den Endpunkt Gewichtsveränderung	159
Abbildung 8: Meta-Analyse der eingeschlossenen Studien zum Effekt von Insulin vs. Sulfonylharnstoff auf den Endpunkt Gewichtsveränderung (vorbehandelte Patienten).....	160
Abbildung 9: Meta-Analyse der eingeschlossenen Studien zum Effekt von Insulin vs. Sulfonylharnstoff auf den Endpunkt Gewichtsveränderung (behandlungsnaive Patienten)..	160
Abbildung 10: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt Gewichtsveränderung	160
Abbildung 11: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt Gewichtsveränderung.....	161
Abbildung 12: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate unerwünschter Ereignisse“	167
Abbildung 13: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate schwerwiegender unerwünschter Ereignisse“	167
Abbildung 14: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Studienabbrüche aufgrund unerwünschter Ereignisse“	168
Abbildung 15: Flussdiagramm Studiendesign und-methodik für CV181040.....	244
Abbildung 16: Flussdiagramm Studiendesign und-methodik für Shank 1995	250
Abbildung 17: Flussdiagramm Studiendesign und-methodik für Birkenland 1994	253
Abbildung 18: Flussdiagramm Studiendesign und-methodik für Tovi 1998.....	257

Abbildung 19: Flussdiagramm Studiendesign und -methodik für die Studie Nathan et al., 1988 261

Abbildung 20: Flussdiagramm Studiendesign und-methodik für Turner 1998 (UKPDS 24) 269

Abbildung 21: Flussdiagramm Studiendesign und-methodik für Hermansen 2007 275

Abbildung 22: Flussdiagramm Studiendesign und-methodik für Garber et al. 2008 281

Abbildung 23: Flussdiagramm Studiendesign und-methodik für die Studie Lewin 2012..... 285

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AkdÄ	Arzneimittelkommission der deutschen Ärzteschaft
ANCOVA	Kovarianzanalyse-Modell
BMI	Body mass index
BMS/AZ	Bristol-Myers Squibb / AstraZeneca
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
DDG	Deutsche Diabetes-Gesellschaft
DIMDI	Deutsches Institut für Medizinische Dokumentation
DPP-4	Dipeptidyl-Peptidase-4
EG	Europäische Gemeinschaft
FEM	Modelle mit festen Effekten
FPG	Nüchternblutzucker
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
HbA1c	Glykiertes Hämoglobin
HR	Hazard Ratio
ITT	Intention to treat
ICTRP Search Portal	International Clinical Trials Registry Platform Search Portal
IVRS	interactive voice response system
LOCF	Last observation carried forward
MFPG	Mean Fasting Plasma Glucose
MFWBG	Mean Fasting Whole Blood Glucose
MTC	Mixed Treatment Comparison
MWD	Mittelwertdifferenz
OAD	Orales Antidiabetikum
OR	Odds Ratio
RCT	Randomized Controlled Trial
REM	Modell mit zufälligen Effekten
RR	Relatives Risiko
SD	Standardabweichung

SE	Standardfehler
SGB	Sozialgesetzbuch
SOC	Systemorganklasse
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
WHO	World Health Organization

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

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

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

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

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

4.1 Zusammenfassung der Inhalte von Modul 4

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

Fragestellung

Codierung B stellt den medizinischen Zusatznutzen von Saxagliptin in der Zweifachkombination mit dem oralen Antidiabetikum (OAD) Sulfonylharnstoff (Anwendungsgebiet Add-on zu Sulfonylharnstoff) gegenüber der zweckmäßigen Vergleichstherapie, der Kombinationsbehandlung mit Humaninsulin+Sulfonylharnstoff oder gegebenenfalls nur mit Humaninsulin, sowie gegenüber der alternativen Vergleichstherapie Sulfonylharnstoff und einem Dipeptidyl-Peptidase-4 (DPP-4)-Inhibitor, dar.

Somit ergeben sich zwei Fragestellungen:

Primäre Fragestellung:

Ist für Saxagliptin in der Indikation „Typ-2-Diabetes mellitus bei erwachsenen Patienten im Alter von 18 Jahren und älter als Ergänzung zu Diät und Bewegung, die mit einem Sulfonylharnstoff alleine nicht ausreichend kontrolliert sind, für die die Anwendung von Metformin ungeeignet erscheint und **bei denen eine Insulintherapie angezeigt ist**, basierend auf den Daten aus randomisierten kontrollierten Studien (RCT) ein medizinischer Zusatznutzen im Sinne des § 35a SGB V gegenüber der vom Gemeinsamen Bundesausschuss (G-BA) festgelegten zweckmäßigen Vergleichstherapie (Humaninsulin+Sulfonylharnstoff oder gegebenenfalls nur Humaninsulin) nachweisbar?

Sekundäre Fragestellung:

Ist für Saxagliptin in der Indikation „Typ-2-Diabetes mellitus bei erwachsenen Patienten im Alter von 18 Jahren und älter als Ergänzung zu Diät und Bewegung, die mit einem Sulfonylharnstoff alleine nicht ausreichend kontrolliert sind, für die die Anwendung von Metformin ungeeignet erscheint und **bei denen eine Insulintherapie noch nicht angezeigt ist**, basierend auf den Daten aus RCT ein medizinischer Zusatznutzen gegenüber der alternativen Vergleichstherapie (Sulfonylharnstoff plus einem anderen DPP-4) nachweisbar?

Der Bewertung des Zusatznutzens vorangestellt ist die Darstellung des medizinischen Nutzens von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Placebo mittels der Methodik der evidenzbasierten Medizin und mittels patientenrelevanter Endpunkte. Saxagliptin 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 Saxagliptin, 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

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

Darüber hinaus dienen Angaben zum HbA1c (glykiertes Hämoglobin)-Wert als Maß für die langfristige Blutzuckerkontrolle zur Interpretation der Ergebnisse hinsichtlich der patientenrelevanten Therapieziele, insbesondere bezüglich des Auftretens von Hypoglykämien. Hierzu soll zudem im Dossier als sogenannter Composite-Endpunkt, welcher post-hoc erhoben wurde, das Erreichen des HbA1c-Zielwerts (<6,5%/<7%) ohne Auftreten von Hypoglykämien betrachtet werden.

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).

Zur Identifizierung von relevanten Studien zu Saxagliptin wurde eine bibliografische Literaturrecherche durchgeführt, da die Zulassung weltweit vor über zwölf Monaten erteilt wurde.

Der Nachweis des **medizinischen Nutzens** von Sulfonylharnstoff+Saxagliptin im Vergleich zu Sulfonylharnstoff+Placebo basiert auf einer doppelblinden RCT von Bristol-Myers Squibb / AstraZeneca (BMS/AZ) (Studie CV181040). 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 Saxagliptin indirekte Vergleiche herangezogen.

Zur Beurteilung der klinischen Sicherheit und Wirksamkeit der ausgewählten zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) sowie der alternativen Vergleichstherapie (Sulfonylharnstoff+DPP-4-Inhibitor) im Vergleich mit dem zu bewertenden Arzneimittel (Sulfonylharnstoff+Saxagliptin) wurde eine bibliografische Literaturrecherche in den Datenbanken MEDLINE, EMBASE und Cochrane durchgeführt. Ziel dieser Literaturrecherche war es, RCT mit der zweckmäßigen Vergleichstherapie bzw. alternativen Vergleichstherapie zu identifizieren, bei denen die Kontrollgruppe mit einem Sulfonylharnstoff (mit oder ohne Placebo) behandelt wurde.

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

Ein-/Ausschlusskriterien für Studien

Die Ein- und Ausschlusskriterien der Studien zum Beleg des Nutzens und des Zusatznutzens der Behandlung mit Saxagliptin 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.	Patienten mit anderen Diabetes-Typen und behandlungsnaive Patienten Tierexperimentelle Studien
Intervention	Behandlung mit Saxagliptin (2,5 bzw. 5 mg einmal täglich)+Sulfonylharnstoff	
Vergleichstherapie	Sulfonylharnstoff+Insulin oder Insulin-Monotherapie ¹ bzw. DPP-4-Inhibitor ² +Sulfonylharnstoff (in zugelassener Dosierung, für die Bewertung des Zusatznutzens) oder Sulfonylharnstoff+Placebo (für die Bewertung des Nutzens)	Andere Kombinations- oder Monotherapien Insulin-Analoga
Endpunkte	Mindestens einer der oben genannten patientenrelevanten Endpunkte muss berücksichtigt werden (HbA1c nur in Kombination mit Hypoglykämien)	
Studientypen	RCT	Nicht-randomisierte Studien, Review-Artikel Case Reports
Studiendauer	≥24 Wochen ³	<24 Wochen
Publikationstyp	Vollpublikation oder Bericht	Keine Vollpublikation oder

	Einschlusskriterien	Ausschlusskriterien
	verfügbar, der den Kriterien des CONSORT Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht	Bericht verfügbar

RCT: Randomisierte kontrollierte Studie; HbA1c: Glykiertes Hämoglobin

¹ Beide vom G-BA vorgeschlagenen zweckmäßigen Vergleichstherapien sollen berücksichtigt werden

² Sitagliptin, Vildagliptin oder Linagliptin

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

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 und der alternativen Vergleichstherapie wurden wenn möglich anhand einer Meta-Analyse zusammengefasst. Die Meta-Analysen wurden mit der Software RevMan Version 5 durchgeführt und basieren auf Modellen mit zufälligen Effekten (REM).

Für dichotome Endpunkte wurde der Therapieeffekt in der Regel anhand des Odds Ratio (OR) zusammengefasst. Wenn in der Publikation oder dem Studienbericht keine OR berichtet wurden, so wurden diese inkl. exakter 95% Konfidenzintervalle unter Verwendung der RevMan-Software selbst berechnet. Im Fall, dass in einem Behandlungsarm kein Ereignis beobachtet wurde (Nullzelle), wurde für die Berechnung des OR eine Korrektur vorgenommen und 0,5 als fester Wert zu allen Zellen der Studienresultate hinzuaddiert. 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]). Für kontinuierliche Variablen wurden Mittelwertdifferenzen berechnet.

Zur Abschätzung der klinischen Relevanz wurden, falls keine validierten Irrelevanzgrenzen für den Gruppenunterschied vorliegen, Responderanalysen herangezogen. Bei Vorliegen eines validierten oder etablierten Responsekriteriums erfolgte eine Dichotomisierung im Sinne von Respondern/Non-Respondern anhand dieses Kriteriums. Danach erfolgte die Bewertung des Gruppenunterschieds anhand des OR (bzw. RR) für Response (bzw. Non-Response). Da bei Responderanalysen davon ausgegangen wird, dass die Responsedefinition eine Bewertung der klinischen Relevanz bereits auf individueller Ebene erlaubt, wurde ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen auch gleichzeitig als klinisch relevant gewertet. Die zusammenfassende Darstellung der Resultate der Einzelstudien sowie eine visuelle Einschätzung möglicher Heterogenität der Studienergebnisse erfolgte – im Falle einer Meta-Analyse – mit Hilfe eines Forest Plots, in dem die Therapieeffekte der Einzelstudien sowie ein zusammenfassender Effektschätzer mit dem entsprechenden Konfidenzintervall abgetragen wurden. Auch im Fall erheblicher Heterogenität werden grundsätzlich zusammenfassende Schätzer berechnet, die Interpretierbarkeit der Schätzer wurde allerdings kritisch diskutiert.

Bezüglich der Sensitivitätsanalysen wurden im Ergebnisteil zur Überprüfung der Robustheit der verwendeten Methodik und zum Ersetzen fehlender Werte bei stetigen Zielvariablen neben den primären Analysen (Last observation carried forward, LOCF-Methode) auch Repeated-Measures Analysen bzw. Observed Cases dargestellt.

Alle Ergebnisse wurden hinsichtlich potenzieller Effektmodifikatoren, d. h. klinischer Faktoren, die die Effekte beeinflussen, untersucht. Möglicherweise variierende Effekte zwischen verschiedenen Subgruppen wurden mittels Interaktionstests bewertet. 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+Saxagliptin) gegenüber der zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) bzw. gegenüber der alternativen

Vergleichstherapie (Sulfonylharnstoff+anderer 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+Saxagliptin versus Sulfonylharnstoff+Insulin oder Insulin-Monotherapie bzw. Sulfonylharnstoff+DPP-4-Inhibitor über den Brückenkompator Sulfonylharnstoff (+Placebo) durchgeführt.

Grundlage für diesen Vergleich waren 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 $\ln \theta_{BC} \pm 1.96 \cdot \sqrt{\text{Var}(\ln \theta_{BC})}$ (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 sowie Mortalität, stationäre Behandlungen, hyperosmolare Komata, gesundheitsbezogene Lebensqualität und 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 Saxagliptin empfohlene Hauptdosis von 5 mg.

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

Medizinischer Nutzen

Die **Ergebnisse zum Nutzen** des zu bewertenden Arzneimittels Saxagliptin in Kombination mit Sulfonylharnstoff basieren auf einer RCT der Phase III (CV181040), die mit der Kombination von Glibenclamid+Saxagliptin durchgeführt wurde. Da hier nur eine relevante RCT vorliegt, wurde keine Meta-Analyse durchgeführt.

Für das **Ausmaß der Blutzuckerkontrolle**, gemessen am HbA1c, konnte gezeigt werden, dass die Behandlung mit Sulfonylharnstoff+Saxagliptin sowohl nach 24 als auch nach 76 Wochen einen klinisch relevanten sowie statistisch signifikanten (Mittelwertsdifferenz (MWD) [95%-KI]: -0,72% [-0,88;-0,56] / -0,56% [-0,72;-0,39]) Unterschied im Vergleich zu Sulfonylharnstoff+Placebo 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 im Vergleich zur Sulfonylharnstoff+Saxagliptin-Gruppe nach 24 (16,6% versus 29,6%) als auch nach 76 Wochen (62,1% versus 74,9%) signifikant häufiger (OR [95%-KI]: 0,47 [0,31;0,72] / 0,55 [0,38;0,80]).

Bei der Anzahl von Patienten mit mindestens einer patientenberichteten oder bestätigten **Hypoglykämie** gab es weder nach 24 noch nach 76 Wochen statistisch signifikante Unterschiede zwischen den Gruppen. Auch für die Anzahl von Studienabbrüchen aufgrund einer Hypoglykämie gab es keine signifikanten Unterschiede.

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 damit ein wichtiges Therapieziel dar. In mehreren großen klinischen Studien konnte gezeigt werden, dass die normnahe Senkung des HbA1c-Wertes die Risiken für Myokardinfarkt und Tod sowie mikrovaskuläre Folgeerkrankungen (wie Nephropathien und Retinopathien) senkt.

Für den Endpunkt „Erreichen des HbA1c-Zielwerts <7% ohne Auftreten von Hypoglykämien“ gab es zu beiden Zeitpunkten einen signifikanten Unterschied zugunsten von Sulfonylharnstoff+Saxagliptin (OR [95%-KI] nach 24 bzw. 76 Wochen: 3,54 [1,28;9,80] bzw. 2,82 [1,16;6,87]).

In beiden Gruppen zeigte sich eine, wenn auch geringe, **Gewichtszunahme**, die in der Sulfonylharnstoff+Saxagliptin-Gruppe sowohl nach 24 als auch nach 76 Wochen etwas höher (0,8 kg und 1,0 kg) lag als in der Sulfonylharnstoff+Placebo-Gruppe (0,3 kg und 0,3 kg). Der Unterschied war für beide Zeitpunkte (LOCF) signifikant (MWD [95%-KI] nach 24 Wochen: 0,5 kg [0,1;0,8] und nach 76 Wochen: 0,70 kg [0,28;1,12]). Die Analyse mit den Observed-

Cases ergab nur nach 76 Wochen einen signifikanten Unterschied (MWD [95%-KI]: 1,80 kg [0,73;3,23]).

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.

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

Zusammenfassend kann man sagen, dass die Behandlung mit Saxagliptin bei Patienten, deren Blutzucker mit Sulfonylharnstoffen alleine nicht mehr ausreichend kontrolliert werden kann, sicher und wirksam ist. Die Kombination von Sulfonylharnstoff+Saxagliptin senkt den Blutzucker effektiv ohne die Hypoglykämie-Rate signifikant zu erhöhen. Die Gesamtrate an unerwünschten Ereignissen ist in beiden Behandlungsgruppen ähnlich.

Medizinischer Zusatznutzen

Die zweckmäßige Vergleichstherapie, die vom G-BA festgelegt wurde, stellt die Kombination aus

Die Ergebnisse zum **medizinischen** Sulfonylharnstoff und Humaninsulin oder gegebenenfalls die Behandlung nur mit Humaninsulin dar. Beide Vergleichstherapien wurden für die Bewertung des Zusatznutzens berücksichtigt. Auch wenn BMS/AZ die vom G-BA bestimmte zweckmäßige Vergleichstherapie akzeptiert, gilt zu berücksichtigen, dass es Patienten gibt, für die eine Insulintherapie noch nicht angezeigt ist. Für diese Patienten stellt die Kombinationsbehandlung aus Sulfonylharnstoff und einem anderen DPP-4-Inhibitor die alternative Vergleichstherapie dar. **Zusatznutzen** basieren auf einem adjustierten indirekten Vergleich von Sulfonylharnstoff+Saxagliptin (aus der Studie CV181040) mit Studien zur zweckmäßigen Vergleichstherapie sowie gegenüber der alternativen Vergleichstherapie mit Sulfonylharnstoff (+Placebo) als Brückenkomparator.

Die systematische Literaturrecherche identifizierte fünf relevante Studien für den indirekten Vergleich mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Insulin oder Insulin-Monotherapie (Birkeland 1994, Tovi 1998, Nathan 1999, Turner 1998 (UKPDS 24), Shank 1995). Alle Studien ermöglichten einen Vergleich gegen Insulin-Monotherapie. Die Studie von Shank 1995 erlaubte zudem einen Vergleich gegen Sulfonylharnstoff+Insulin. Diese Studie verwendete Glipizid als Sulfonylharnstoff. Für die alternative Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor wurden drei relevante Studien identifiziert (Hermansen 2007, Garber 2008 und Lewin 2012). In zwei Studien mit Sulfonylharnstoff+einem anderen DPP-4-Inhibitor wurde Glimperid als Sulfonylharnstoff benutzt (Hermansen 2007 und Garber 2008). Die Studie von Lewin 2012 verwendete diverse Sulfonylharnstoffe als Basistherapie.

Wenn es möglich war, wurden die einzelnen Studien meta-analysiert und ein Gesamtschätzer berechnet. Für die alternative Vergleichstherapie wurde sowohl ein Vergleich zur Gruppe Sulfonylharnstoff+andere DPP-4-Inhibitoren als auch zur Leitsubstanz Sulfonylharnstoff+Sitagliptin durchgeführt.

Der indirekte Vergleich mit der zweckmäßigen Vergleichstherapie für das **Ausmaß der Blutzuckerkontrolle (HbA1c)** konnte gegen Insulin-Monotherapie anhand der zwei Studien von Birkeland 1994 und Tovi 1998 durchgeführt werden. In der Studie von Shank 1995, die einen Vergleich mit Sulfonylharnstoff+Insulin ermöglicht hätte, standen nicht genügend Informationen für einen indirekten Vergleich zur Verfügung. Der Vergleich mit Sulfonylharnstoff+Saxagliptin gegenüber Insulin zeigte ein statistisch signifikantes Ergebnis zugunsten der Insulin-Behandlungsgruppe nach 50 Wochen (MWD [95%-KI]: 1,41% [0,70;2,13]). Da die HbA1c-senkenden Effekte von Antidiabetika u. a. stark vom Ausgangs-HbA1c abhängen, muss das Ergebnis aber hinterfragt werden, da der Ausgangs-HbA1c-Wert insbesondere in der Studie von Tovi 1998 höher war. Darüber hinaus zeigte der Vergleich mit der Studie von Nathan 1988, welche behandlungsnaive Patienten einschloss und deshalb separat betrachtet wurde, keinen statistisch signifikanten Unterschied mehr. Ein indirekter Vergleich für das Auftreten von **Hypoglykämien** war nur gegen Insulin-Monotherapie mit der Studie von Birkeland 1994 möglich. Es zeigte sich ein signifikantes Ergebnis nach 50 Wochen zugunsten von Sulfonylharnstoff+Saxagliptin gegenüber Insulin (OR [95%-KI]: 0,04 [0,00;0,88]).

Für den indirekten Vergleich mit der alternativen Vergleichstherapie Sulfonylharnstoff+anderer DPP-4-Inhibitor konnten alle drei Studien herangezogen werden. Der Vergleich ergab weder für den Endpunkt „Ausmaß der Blutzuckerkontrolle (HbA1c)“ noch für die Anzahl der Patienten mit mindestens einer Hypoglykämie einen statistisch signifikanten Unterschied zwischen Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+einem anderen DPP-4-Inhibitor als Gruppe bzw. Sulfonylharnstoff+Sitagliptin.

Gemessen an dem Ziel, eine adäquate Blutzuckerkontrolle bei gleichzeitiger Hypoglykämievermeidung zu erreichen, ergibt sich aus der gemeinsamen Betrachtung dieser Endpunkte kein Hinweis auf eine Über- oder Unterlegenheit weder für Sulfonylharnstoff+Saxagliptin gegenüber Insulin-Monotherapie noch für Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+DPP-4-Inhibitor.

Der indirekte Vergleich für den Endpunkt „**Gewichtsveränderung**“ war sowohl gegen Sulfonylharnstoff+Insulin sowie gegen Insulin-Monotherapie möglich. Der Vergleich von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Insulin nach 24 Wochen zeigte keinen statistisch signifikanten Unterschied.

Der Vergleich gegen Insulin-Monotherapie konnte mit Birkeland 1994 und Tovi 1998 durchgeführt werden. Er zeigte eine statistisch signifikante Gewichtsreduktion in der mit Sulfonylharnstoff+Saxagliptin behandelten Patientengruppe gegenüber den mit Insulin behandelten Patienten (MWD [95%-KI]: -4,18 kg [-6,29;-2,07]). Darüber hinaus wurde der

Vergleich mit den Studien von Nathan 1988 und Turner 1998, welche behandlungsnaive Patienten einschlossen, separat dargestellt. Für diesen Vergleich gab es keinen statistisch signifikanten Unterschied von Sulfonylharnstoff+Saxagliptin gegenüber Insulin. Bei der Betrachtung aller vier Studien ergab sich hingegen ebenfalls ein statistisch signifikanter Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin (MWD [95%-KI]: -2,67 kg [-4,48;-0,86]).

Der indirekte Vergleich mit der alternativen Vergleichstherapie konnte mit zwei der drei Studien durchgeführt werden und ergab keine statistisch signifikanten Unterschiede zwischen Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+einem anderen DPP-4-Inhibitor wie auch der Kombination Sulfonylharnstoff+Sitagliptin.

Gemessen an dem Ziel, eine Reduktion des Körpergewichts zu erreichen, ergibt sich gegenüber der zweckmäßigen Vergleichstherapie eine bedeutsame Vermeidung nicht schwerwiegender Nebenwirkungen und damit ein Anhaltspunkt für einen beträchtlichen Zusatznutzen für Sulfonylharnstoff+Saxagliptin.

Gegenüber Sulfonylharnstoff+DPP-4-Inhibitor ergibt sich keine Über- oder Unterlegenheit für Sulfonylharnstoff+Saxagliptin.

Ein indirekter Vergleich für den Endpunkt „**Unerwünschte Ereignisse**“ konnte nur für die alternative Vergleichstherapie durchgeführt werden. Hier konnten alle drei Studien berücksichtigt werden. Es ergaben sich weder für die Gesamtrate unerwünschter Ereignisse oder schwerwiegender unerwünschter Ereignisse noch für Studienabbrüche aufgrund von unerwünschten Ereignissen statistisch signifikante Unterschiede zwischen Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+anderen DPP-4-Inhibitoren als Gruppe. Der Vergleich gegen Sulfonylharnstoff und Sitagliptin als Einzelsubstanz zeigte in Bezug auf die schwerwiegenden unerwünschten Ereignisse und Studienabbrüche aufgrund unerwünschter Ereignisse keine Unterschiede. Allerdings konnte für die Gesamtrate unerwünschter Ereignisse ein statistisch signifikanter Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin detektiert werden. Dieser Unterschied wird aber in Frage gestellt, da der Endpunkt in der Studie von Hermansen 2007 sehr wahrscheinlich anders operationalisiert wurde als in der Studie CV181040.

Insgesamt ergibt sich für den Endpunkt „unerwünschter Ereignisse“ kein Hinweis auf einen Zusatznutzen gegenüber Sulfonylharnstoff+DPP-4-Inhibitoren.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Die Ergebnisse zum Zusatznutzen zeigen, dass es für die Endpunkte „Ausmaß der Blutzuckerkontrolle (HbA1c)“ und „Hypoglykämien“ keinen Anhaltspunkt auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen (Sulfonylharnstoff+Saxagliptin gegen zweckmäßige Vergleichstherapie) gibt.

Für den Endpunkt „Gewichtsveränderungen“ ergibt sich ein Anhaltspunkt auf einen beträchtlichen Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie.

Gegenüber der alternativen Vergleichstherapie zeigt sich für die Endpunkte „Ausmaß der Blutzuckerkontrolle (HbA1c)“, „Hypoglykämien“, „Gewichtsveränderungen“ und „unerwünschte Ereignisse“ kein Hinweis auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen (Sulfonylharnstoff+Saxagliptin gegen Sulfonylharnstoff+anderen DPP-4-Inhibitoren).

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

Endpunkt	Quantifizierung des Zusatznutzens	Wahrscheinlichkeit
Im Vergleich zur zweckmäßigen Vergleichstherapie		
Hypoglykämien unter Berücksichtigung des HbA1c	Kein Zusatznutzen	Anhaltspunkt
Gewichtsveränderung	Beträchtlicher Zusatznutzen	Anhaltspunkt
Im Vergleich zu Sulfonylharnstoff+anderen DPP-4-Inhibitoren		
Hypoglykämien unter Berücksichtigung des HbA1c	Kein Zusatznutzen	Hinweis
Gewichtsveränderung	Kein Zusatznutzen	Hinweis
Unerwünschte Ereignisse (einschließlich Hypoglykämien)	Kein Zusatznutzen	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. In der Nationalen VersorgungsLeitlinie (Konsultationsfassung) wird entsprechend bei einem Body-Mass-Index von 27-35 kg/m² eine Gewichtsabnahme von 5% empfohlen.

Insgesamt ergibt sich daher durch die Gewichtsreduktion im Vergleich zur zweckmäßigen Vergleichstherapie ein **Anhaltspunkt auf einen beträchtlichen Zusatznutzen**.

Im Vergleich zur alternativen Vergleichstherapie ergibt sich kein Hinweis auf einen Zusatznutzen.

Unabhängig von den Ergebnissen muss berücksichtigt werden, dass:

- die Behandlung mit Sulfonylharnstoffen oder Insulin, mit einem erhöhten Hypoglykämie-Risiko und mit Gewichtszunahmen assoziiert ist. Dies gilt dann auch

für die Kombination eines DPP-4-Inhibitors mit diesen Antidiabetika. Die Behandlung mit DPP-4-Inhibitoren allein wird hingegen in der Literatur als gewichtsneutral mit einem geringen Hypoglykämierisiko beschrieben.

- im Gegensatz zu einer Therapie mit OAD jede Insulintherapie von einer Blutglukoseselbstkontrolle begleitet sein sollte, was eine intensive Schulung voraussetzt. Je nach Art der Insulintherapie sind hierfür 2-4 Blutzuckermessungen pro Tag erforderlich.
- insbesondere eine konventionelle Insulintherapie einen negativen Einfluss auf die Flexibilität der Patienten haben kann. Um eine konventionelle Therapie erfolgreich durchzuführen, sind ein geregelter Tagesablauf mit festgesetzten Spritz- und Esszeiten erforderlich, sowie meistens auch Zwischenmahlzeiten zur Vermeidung von Hypoglykämien. Letzteres ist in Anbetracht der häufig bestehenden Gewichtsprobleme eher ungünstig.
- durch eine intensivierete Insulintherapie zwar eine deutlich größere Flexibilität im Alltag erreicht werden kann, dafür aber häufiger Insulin gespritzt und auch häufiger der Blutzucker gemessen werden muss.

Aus der Zusammenschau aller Ergebnisse lässt sich ableiten, dass Saxagliptin in Kombination mit einem Sulfonylharnstoff eine wirksame und gut verträgliche Alternative zur zweckmäßigen Vergleichstherapie darstellt, die im Vergleich zu dieser in bedeutsamem Maße zur Gewichtsreduktion beitragen kann.

4.2 Methodik

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

4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

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

Intervention

Bei der untersuchten Intervention handelt es sich um Saxagliptin (Onglyza®) als Hydrochlorid. Saxagliptin ist ein DPP-4-Inhibitor und ist bereits zugelassen in den Dosierungen 5 mg sowie 2,5 mg einmal täglich (bei Patienten mit Niereninsuffizienz).

Saxagliptin ist zugelassen zur Behandlung von erwachsenen Patienten ab 18 Jahren mit Typ-2-Diabetes mellitus zur Verbesserung der Blutzuckerkontrolle.

Saxagliptin ist zugelassen als Add-on Kombinationstherapie

- (1) in Kombination mit Metformin, wenn eine Metformin-Monotherapie, zusammen mit einer Diät und Bewegung, den Blutzucker nicht ausreichend kontrolliert (Codierung A).
- (2) in Kombination mit einem Sulfonylharnstoff bei Patienten, für die die Anwendung von Metformin ungeeignet erscheint, wenn eine Sulfonylharnstoff-Monotherapie, zusammen mit einer Diät und Bewegung, den Blutzucker nicht ausreichend kontrolliert (Codierung B).

- (3) in Kombination mit einem Thiazolidindion bei Patienten, für die die Anwendung eines Thiazolidindions geeignet erscheint, wenn eine Thiazolidindion-Monotherapie, zusammen mit einer Diät und Bewegung, den Blutzucker nicht ausreichend kontrolliert (Dossier nicht notwendig).
- (4) in Kombination mit Insulin (mit oder ohne Metformin), wenn diese Behandlung allein, zusammen mit einer Diät und Bewegung, den Blutzucker nicht ausreichend kontrolliert (Codierung C).
- (5) in Kombination mit Metformin und einem Sulfonylharnstoff, wenn diese Behandlung allein, zusammen mit einer Diät und Bewegung, den Blutzucker nicht ausreichend kontrolliert (Codierung D).

Das vorliegende Dossier stellt den Zusatznutzen in der Therapiesituation der Add-on Kombinationstherapie mit Sulfonylharnstoff dar (Codierung B).

Fragestellung

Die primäre Fragestellung dieser Nutzenbewertung adressiert den Zusatznutzen von Saxagliptin gegenüber der vom G-BA festgelegten zweckmäßigen Vergleichstherapie, der Kombinationsbehandlung mit Humaninsulin+Sulfonylharnstoff oder gegebenenfalls nur mit Humaninsulin. Zusätzlich wird der medizinische Zusatznutzen von Saxagliptin im Vergleich zu anderen DPP-4-Inhibitoren beurteilt. Somit ergeben sich zwei Fragestellungen.

Primäre Fragestellung:

Ist für Saxagliptin in der Indikation „Typ-2-Diabetes mellitus bei erwachsenen Patienten im Alter von 18 Jahren und älter als Ergänzung zu Diät und Bewegung, die mit einem Sulfonylharnstoff alleine nicht ausreichend kontrolliert sind, für die die Anwendung von Metformin ungeeignet erscheint und bei denen eine Insulintherapie angezeigt ist“, basierend auf den Daten aus RCT) ein medizinischer Zusatznutzen im Sinne des § 35a SGB V gegenüber der vom G-BA festgelegten zweckmäßigen Vergleichstherapie (Humaninsulin+Sulfonylharnstoff oder gegebenenfalls nur Humaninsulin) nachweisbar?

Sekundäre Fragestellung:

Ist für Saxagliptin in der Indikation „Typ-2-Diabetes mellitus bei erwachsenen Patienten im Alter von 18 Jahren und älter als Ergänzung zu Diät und Bewegung, die mit einem Sulfonylharnstoff alleine nicht ausreichend kontrolliert sind, für die die Anwendung von Metformin ungeeignet erscheint und bei denen eine Insulintherapie noch nicht angezeigt ist“, basierend auf den Daten aus RCT ein medizinischer Zusatznutzen gegenüber der alternativen Vergleichstherapie (Sulfonylharnstoff plus einem anderen DPP-4) nachweisbar?

Vergleichstherapie

Die definierte zweckmäßige Vergleichstherapie für dieses Anwendungsgebiet ist die Kombination aus Humaninsulin und einem Sulfonylharnstoff bzw. einer Therapie nur mit Humaninsulin.

Patientenpopulation

Im vorliegenden Anwendungsgebiet werden Studien mit erwachsenen Patienten mit manifestem Diabetes mellitus Typ 2 eingeschlossen, die unter der maximal verträglichen Dosis einer Sulfonylharnstoff-Monotherapie, zusammen mit einer Diät und Bewegung, einen nicht ausreichend kontrollierten Blutzucker (HbA1c) aufweisen.

Endpunkte

Die Bewertung wurde bezüglich patientenrelevanter Endpunkte vorgenommen. Detaillierte Informationen zu den einzelnen relevanten Endpunkten finden sich in Abschnitt 4.2.2.

Studientypen

RCT liefern für die Bewertung des Nutzens einer medizinischen Intervention die zuverlässigsten Ergebnisse und sind mit der geringsten Ergebnisunsicherheit behaftet. Die Bewertung des medizinischen Nutzens und Zusatznutzens erfolgt daher auf der Basis der Ergebnisse aus RCT.

Die einzelnen Komponenten dieser Fragestellungen (Patientenpopulation, Intervention, Vergleichstherapie, Endpunkte, Studientypen) werden in Abschnitt 4.2.2 näher spezifiziert.

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 Abschnitt 4.2.1 geeignet sind. Die Ein- und Ausschlusskriterien für Studien sind wie folgt definiert und begründet.

Population: Die Wahl der Selektionskriterien für die Studienpopulation ergibt sich aus der diesem Modul zugrunde liegenden Zielpopulation, für die Saxagliptin zugelassen ist (Bristol-Myers Squibb / AstraZeneca, 2013): erwachsene Patienten mit manifestem Diabetes mellitus Typ 2, für die die Anwendung von Metformin ungeeignet erscheint, die unter einer Sulfonylharnstoff-Monotherapie, zusammen mit einer Diät und Bewegung einen nicht ausreichend kontrollierten Blutzucker aufweisen.

Intervention: Behandlung mit Saxagliptin in der zugelassenen Wirkstärke (einmal täglich Saxagliptin 5 mg, sowie 2,5 mg bei Patienten mit Niereninsuffizienz) als Add-on Kombinationstherapie mit Sulfonylharnstoff.

Vergleichstherapie: Die zweckmäßige Vergleichstherapie entspricht der vom G-BA im Beratungsgespräch (Beratungsanforderung 2012-B-027) am 13.08.2012 (Klipper, 2012) festgelegten Therapie: die Kombinationsbehandlung mit Humaninsulin+Sulfonylharnstoff oder gegebenenfalls nur mit Humaninsulin. Für die Begründung der Wahl der zweckmäßigen Vergleichstherapie siehe Modul 3B, Abschnitt 3.1.2.

Wie bereits im Beratungsgespräch mit dem G-BA erläutert, sieht BMS/AZ bei Patienten, für die die zweckmäßige Vergleichstherapie, in diesem Fall eine Insulintherapie, noch nicht angezeigt ist, die Notwendigkeit eine alternative Vergleichstherapie zu definieren. Da der G-BA keine Empfehlung zur alternativen Vergleichstherapie abgegeben hat, sieht BMS/AZ andere DPP-4-Hemmer als die geeignete Vergleichstherapie in so einer Therapiesituation an. Dies entspricht sowohl der zugelassenen Indikation als auch der Positionierung der Substanzklasse in der Typ-2-Diabetesbehandlung durch den G-BA (G-BA, 2010; G-BA, 2012).

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 Blutzuckerkontrolle zur Interpretation der Ergebnisse hinsichtlich der patientenrelevanten Therapieziele, insbesondere bezüglich des Auftretens von Hypoglykämien. Näheres hierzu in Abschnitt 4.2.5.2. Hierzu soll zudem im Dossier als sogenannter Composite-Endpunkt das Erreichen des HbA1c-Zielwerts (<6,5%/<7%) ohne Auftreten von Hypoglykämien betrachtet werden.

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

Studiendauer: Entsprechend bisheriger Nutzenbewertungen von Arzneimitteln im Indikationsgebiet werden ausschließlich Studien mit einer Behandlungsdauer von mindestens 24 Wochen eingeschlossen, da die antidiabetische Therapie in der Regel dauerhaft gegeben werden muss und somit auch ein langfristiger Effekt der Intervention auf patientenrelevante Therapieziele (z. B. Gewichtsveränderungen) vom Interesse ist.

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.	Patienten mit anderen Diabetes-Typen und behandlungsnaive Patienten Tierexperimentelle Studien
Intervention	Behandlung mit Saxagliptin (2,5 bzw. 5 mg einmal täglich)+Sulfonylharnstoff	
Vergleichstherapie	Sulfonylharnstoff+Insulin oder Insulin-Monotherapie ¹ bzw. DPP-4-Inhibitor ² +Sulfonylharnstoff (in zugelassener Dosierung, für die Bewertung des Zusatznutzens) oder Sulfonylharnstoff+Placebo (für die Bewertung des	Andere Kombinations- oder Monotherapien Insulin-Analoga

	Einschlusskriterien	Ausschlusskriterien
	Nutzens)	
Endpunkte	Mindestens einer der oben genannten patientenrelevanten Endpunkte muss berücksichtigt werden (HbA1c nur in Kombination mit Hypoglykämien)	
Studientypen	RCT	Nicht-randomisierte Studien, Review-Artikel Case Reports
Studiendauer	≥24 Wochen ³	<24 Wochen
Publikationstyp	Vollpublikation, im clinicaltrials.gov Register publizierte Daten oder Bericht verfügbar, der den Kriterien des CONSORT Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht	Keine Vollpublikation oder Bericht verfügbar

RCT: Randomisierte kontrollierte Studie; HbA1c: Glykiertes Hämoglobin

¹ Beide vom G-BA vorgeschlagenen zweckmäßigen Vergleichstherapien sollen berücksichtigt werden

² Sitagliptin, Vildagliptin oder Linagliptin

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

4.2.3 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

Die Durchführung einer bibliografischen Literaturrecherche ist für das zu bewertende Arzneimittel Onglyza erforderlich, da die Zulassung weltweit vor über zwölf Monaten erteilt wurde (Zulassung in den USA seit 31.07.2009 und EU-Zulassung seit 01.10.2009).

Literaturrecherche zur Durchführung eines indirekten Vergleiches gegenüber der zweckmäßigen Vergleichstherapie

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

Literaturrecherche zur Durchführung eines indirekten Vergleiches gegenüber der alternativen Vergleichstherapie

Wie bereits erläutert, sieht BMS/AZ für die Patientengruppen, für die eine Insulintherapie noch nicht angezeigt ist, die Kombinationsbehandlung von Sulfonylharnstoff+DPP-4-Inhibitor als alternative Vergleichstherapie an. Da keine direkt vergleichende Studie vorliegt, werden auch hier indirekte Vergleiche herangezogen.

Dabei sollen Sitagliptin (Leitsubstanz) sowie Vildagliptin und Linagliptin als weitere in Deutschland zugelassene DPP-4-Inhibitoren berücksichtigt werden.

Beschreibung der Suchen

Zur Beurteilung der klinischen Sicherheit und Wirksamkeit der ausgewählten zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) bzw. alternativen Vergleichstherapie (Sulfonylharnstoff+DPP-4-Inhibitor) im Vergleich mit dem zu bewertenden Arzneimittel (Sulfonylharnstoff+Saxagliptin) wird eine bibliografische Literaturrecherche durchgeführt. Ziel dieser Literaturrecherche ist es, RCT mit der zweckmäßigen Vergleichstherapie bzw. alternativen Vergleichstherapie zu identifizieren, bei denen die Kontrollgruppe mit einem Sulfonylharnstoff mit oder ohne Placebo behandelt wurde, da Sulfonylharnstoff mit oder ohne Placebo als einziger Brückenkompator identifiziert wurde.

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 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) verwendet.

Die Ergebnisse der systematischen bibliografischen Literaturrecherchen sind in Abschnitt 4.3.1.1.2 sowie 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 Saxagliptin für die Bewertung berücksichtigt werden, wird eine Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien mit dem zu bewertenden Arzneimittel Saxagliptin durchgeführt. Die Suche erfolgt entsprechend der Vorgaben in dem Studienregister clinicaltrials.gov und über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO).

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

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

Studienregistersuche zur Durchführung eines indirekten Vergleiches mit Insulin ± Sulfonylharnstoff

Um sicherzustellen, dass alle verfügbaren Studiendaten zur Vergleichstherapie Sulfonylharnstoff+Insulin oder Insulin-Monotherapie (auch im Vergleich zur Sulfonylharnstoff mit/ohne Placebo) für die Bewertung berücksichtigt werden, wird eine weitere Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien mit Insulin und mit Sulfonylharnstoffen durchgeführt. Die Studienregistersuche für die

Vergleichstherapie wird in denselben Datenbanken wie für das zu bewertende Arzneimittel durchgeführt.

Die detaillierte Suchstrategie ist in Anhang 4-B dokumentiert.

Die identifizierten Studien werden auf Behandlung mit Insulin im Vergleich mit Sulfonylharnstoff (als Brückenkomparator) selektiert.

Studienregistersuche zur Durchführung eines indirekten Vergleiches mit einer Kombinationstherapie aus Sulfonylharnstoffen und DPP-4-Inhibitoren

Um sicherzustellen, dass alle verfügbaren Studiendaten zur alternativen Vergleichstherapie für die Bewertung berücksichtigt werden, wird eine weitere Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien mit Sulfonylharnstoffen und DPP-4-Inhibitoren durchgeführt. Die Studienregistersuche für die alternative Vergleichstherapie wird in denselben Datenbanken wie für das zu bewertende Arzneimittel durchgeführt.

Die detaillierte Suchstrategie ist in Anhang 4-B dokumentiert.

Die identifizierten Studien werden auf Behandlung mit Sulfonylharnstoff+DPP-4-Inhibitor im Vergleich mit Sulfonylharnstoff (als Brückenkomparator) selektiert.

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 bibliographischen Literaturrecherche und der Studienregistersuche identifizierten Studien für das zu bewertende Arzneimittel werden unter Anwendung der in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien nach Titel und Abstracts unabhängig von zwei Personen selektiert. Diskrepanzen werden durch Diskussion aufgelöst. Die selektierten Studien werden mit dem schon bekannten Studienpool abgeglichen und neue, für die jeweiligen Fragestellungen relevante Studien wurden dem Studienpool hinzugefügt.

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-G):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

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

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

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

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

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

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

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

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

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

Bewertung des Verzerrungspotenzials

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

Verzerrungsaspekte der Ergebnisse auf Studienebene:

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

Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

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

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

Abschließend wird unter Berücksichtigung der oben genannten Aspekte das Verzerrungspotenzial auf Ebene der Endpunkte als „hoch“ oder „niedrig“ eingestuft. Von einem ‚niedrigen‘ Verzerrungspotenzial wird ausgegangen, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Eine relevante Verzerrung liegt vor, falls sich die Aussagen und Schlussfolgerungen bezüglich der Ergebnisse bei Behebung des verzerrenden Aspektes verändern würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ wird in der Bewertung berücksichtigt. In diesem Fall werden dann Sensitivitätsanalysen durchgeführt. Die Klassifizierung dient der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

Die Vorgehensweise bei der Bewertung des Verzerrungspotenzials für jede Studie und jeden Endpunkt (Operationalisierung) richtet 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 gegebenen Fragestellungen relevanten Studien werden Detailinformationen zu Design und Methodik extrahiert und im Abschnitt 4.3.1.2 und 4.3.2.1.2 beschrieben. Die Beschreibung erfolgt gemäß den Punkten 2b bis 14 des CONSORT-Statements (siehe Anhang 4-F). Der Patientenfluss wird gemäß CONSORT-Flow-Chart dargestellt.

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.

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.

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

Zur Charakterisierung der in die Nutzenbewertung eingeschlossenen Studien hinsichtlich der untersuchten Patientenpopulation werden sowohl demografische als auch krankheitsspezifische Daten für die Einzelstudien zusammengefasst. Folgende Faktoren werden berücksichtigt: Alter, Geschlecht, Gewicht, Body-Mass-Index (BMI), HbA1c vor Randomisierung, Krankheitsdauer und Anzahl der Studienabbrecher. Stetige Variablen werden anhand des Mittelwertes und der Standardabweichung (SD) beschrieben, kategoriale 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. Näheres zur Art der Adjustierung kann den originalen Studienberichten entnommen werden. Zusätzlich dazu wird für dichotome Zielvariablen das OR als das vom G-BA präferierte Effektmaß berechnet. Im Fall, dass in einem Behandlungsarm kein Ereignis beobachtet wurde (Nullzelle), wurde für die Berechnung des OR eine Korrektur vorgenommen und 0,5 als fester Wert zu allen Zellen der Studienresultate hinzuaddiert. Alle Effektmaße werden mit den jeweiligen Intervallschätzern angegeben. Wenn nicht in den Publikationen angegeben, wurde der Standardfehler für die Differenz von Mittelwerten (A minus B) mittels der Formel $SE_{Dif} = \sqrt{SE_A^2 + SE_B^2}$ berechnet. Für die Umrechnung von Streuungsmaßen normalverteilter, kontinuierlicher Variablen wurden übliche Zusammenhänge verwendet, z. B. Standardfehler=Standardabweichung/ \sqrt{n} .

Die folgenden patientenrelevanten Zielgrößen sollen berücksichtigt werden:

- Hypoglykämien
- Gewichtsveränderungen
- Unerwünschte Ereignisse
- Folgekomplikationen und Mortalität
 - Kardiovaskuläre Morbidität und Mortalität
 - Zerebrovaskuläre Morbidität und Mortalität
 - Gefäßbedingte nichtkardiale und nichtzerebrale Morbidität und Mortalität

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

In Bezug auf die Verbesserung des Gesundheitszustandes werden Hypoglykämien nur unter der Berücksichtigung der Blutzuckerkontrolle (HbA1c) betrachtet. Das liegt darin begründet, dass bei der blutzuckersenkenden Behandlung von Patienten mit Diabetes mellitus Typ 2 die Vermeidung von Hypoglykämien ein wichtiges Therapieziel ist (IQWiG, 2009b), 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 Hypoglykämie-Rate 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 gemeinsamen Betrachtung der Endpunkte Blutzuckerkontrolle (HbA1c) und Hypoglykämien

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

Kein Unterschied: kein Vor- bzw. Nachteil für eine der beiden Behandlungsoptionen.
Nachteil: Statistisch signifikanter und klinisch relevanter Unterschied zu Ungunsten des zu bewertenden Arzneimittels; Als klinisch relevant wird ein HbA1c Unterschied von >0,35% (Obergrenze des KI >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, 2009a). 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% (Untergrenze des KI >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 (sprich >0,175%) diskutiert. 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.

¹ Für 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.

² Auch 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 zum aktiven Komparator (Zusatznutzen) soll vor allem eine Verbesserung der Therapieerträglichkeit in Bezug auf die Hypoglykämie-Rate bewirkt werden. Der Effekt auf den Parameter HbA1c wird daher als nachrangig gegenüber dem Effekt auf die Hypoglykämie-Rate angesehen.

Hierzu soll zudem im Dossier als sogenannter Composite-Endpunkt das Erreichen des HbA1c-Zielwerts (<6,5%/<7%) ohne Auftreten von Hypoglykämien betrachtet werden. In den deutschen Leitlinien wird ein HbA1c-Zielwert von <6,5% empfohlen, welcher aber nur angestrebt werden soll, wenn dabei auch das Auftreten bedeutender Nebenwirkungen wie Hypoglykämien vermieden werden kann. In den Leitlinien der DDG und der

Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ) wird darauf hingewiesen, dass die Therapieziele individuell auf den Patienten abgestimmt werden sollen und dabei der allgemeine Gesundheitszustand, Begleiterkrankungen, Alter sowie das individuelle Hypoglykämierisiko in Betracht gezogen werden müssen (AkdÄ, 2009; Matthaei et al., 2009). Es hat sich gezeigt, dass nicht alle derzeit in Deutschland verfügbaren Antidiabetika dafür geeignet sind, dieses Therapieziel zu erreichen (AkdÄ, 2009; Matthaei et al., 2009). Die Betrachtung des kombinierten Endpunktes hat gegenüber der Betrachtung der Hypoglykämiehäufigkeit in Relation zur HbA1c-Senkung im Gruppenvergleich den Vorteil, dass so patientenindividuell das Erreichen des oben beschriebenen Therapieziels bewertet werden kann.

Zur Analyse der Verträglichkeit und Sicherheit werden die Gesamtrate unerwünschter Ereignisse, Gesamtrate schwerwiegender unerwünschter Ereignisse, Therapieabbrüche aufgrund unerwünschter Ereignisse und kardiale Ereignisse (im Rahmen der Erhebung der unerwünschten Ereignisse in der Studie) betrachtet.

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

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

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 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 von Heterogenitätstest oder I^2 -Maß), erfolgt die quantitative Zusammenfassung der Studienresultate mittels einer Meta-Analyse.

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

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.

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

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

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

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

Die quantitative Untersuchung der statistischen Heterogenität erfolgt mit Hilfe eines statistischen Heterogenitätstests (Q-Statistik, p-Wert) und der Berechnung der I^2 -Statistik (Higgins et al., 2003). Um statistische Heterogenität festzustellen wird ein Signifikanzlevel von 0,2 verwendet. Für die I^2 -Statistik gelten die im Cochrane Handbuch (Higgins und Green, 2008) beschriebenen Grenzwerte für Heterogenität (>50% indiziert erhebliche Heterogenität). Ist die statistische Heterogenität nicht bedeutsam, d. h. $p > 0,2$ (Heterogenitätstest), so wird ein zusammenfassender Schätzer berechnet. Auch im Fall erheblicher Heterogenität werden grundsätzlich zusammenfassende Schätzer berechnet (Skipka und Bender, 2010), die Interpretierbarkeit der Schätzer wird allerdings kritisch diskutiert und mögliche Ursachen der Heterogenität werden mittels Sensitivitäts- und Subgruppenanalysen untersucht.

In folgenden Fällen werden keine Meta-Analysen durchgeführt:

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

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

4.2.5.4 Sensitivitätsanalysen

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

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

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

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

Folgende Sensitivitätsanalysen sind im Ergebnisteil adressiert: Zur Überprüfung der Robustheit der verwendeten Methodik und zum Ersetzen fehlender Werte bei stetigen Zielvariablen werden neben den primären Analysen (Last observation carried forward, LOCF Methode) auch Repeated-Measures Analysen bzw. Observed Cases 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-Regressionsanalysen 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 untersucht. Möglicherweise variierende Effekte zwischen verschiedenen Subgruppen werden mittels Stratifikation und eines 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. 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; ≥ 65 Jahre)
- Geschlecht (männlich; weiblich)

Für weitere in den Studienprotokollen der Einzelstudien geplante und berichtete Subgruppenanalysen werden die Daten stratifiziert dargestellt und die p-Werte von Interaktionstests berichtet.

Für die Studie CV181040 wurden die Daten stratifiziert dargestellt und Interaktionstests (falls nicht anders angegeben) für die subgruppenbildenden Faktoren für alle Endpunkte, für die nicht eine Interaktionstestung per Protokoll schon geplant war, post hoc durchgeführt.

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

4.2.5.6 Indirekte Vergleiche

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

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

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

„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.

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.*

Für die primäre Fragestellung liegt im betreffenden Anwendungsgebiet, Add-on Kombination mit Sulfonylharnstoff, keine direkt vergleichende Head-to-Head-Studie vor. Deshalb wurde ein adjustierter, indirekter Vergleich zur Bewertung des medizinischen Zusatznutzens herangezogen. Zusätzlich wurde ein adjustierter, indirekter Vergleich gegen andere in Deutschland zugelassene DPP-4-Inhibitoren (Sitagliptin, Vildagliptin und Linagliptin) betrachtet. Für den Vergleich zwischen Saxagliptin und Sitagliptin wird ebenfalls ein adjustierter indirekter Vergleich in die Nutzenbewertung eingeschlossen.

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. Lüthmann, 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.

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 wurde mittels eines frequentistischen Ansatzes durchgeführt. Er beruht auf einem einfachen Netzwerk und soll für den Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin oder Insulin-Monotherapie bzw. die alternative Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor durchgeführt.

Grundlage für diesen Vergleich sind 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

$$\ln \theta_{BC} \pm 1.96 * \sqrt{\text{Var}(\ln \theta_{BC})} \quad (3)$$

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, war die Untersuchung der Heterogenität der eingeschlossenen Studien mittels statistischer Tests sowie gebräuchlicher Heterogenitätsmaße vorgesehen. Des Weiteren wurden die eingeschlossenen Studien hinsichtlich wichtiger Eigenschaften (Ein- und Ausschlusskriterien; Patientencharakteristika etc.) charakterisiert und auf das Vorliegen von wesentlichen Unterschieden untersucht, die die Ergebnisse des indirekten Vergleiches verzerren könnten. Bei substantieller klinischer Heterogenität, die zu einer starken Verzerrung der Ergebnisse des indirekten Vergleiches führen kann, können entsprechende Studien aus dem indirekten Vergleich begründet ausgeschlossen werden.

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
CV181040	ja	abgeschlossen	Behandlungsphase: 24 Wochen Verlängerungsphase: 52 Wochen	Glibenclamid (7,5 mg/Tag) +Saxagliptin (2,5 und 5 mg/Tag) Glibenclamid (hochtitriert auf 10 mg/Tag) +Placebo
D1680C00007	ja	abgeschlossen	Behandlungsphase: 12 Wochen Verlängerungsphase: 40 Wochen	Vorbehandlung +Saxagliptin (2,5 mg/Tag) Vorbehandlung +Placebo
D1680C00003 (SAVOR-TIMI 53)	nein	laufend	4 Jahre*	Vorbehandlung +Saxagliptin (2,5 und 5 mg/Tag) Vorbehandlung +Placebo
CV181142	nein	laufend	6 Wochen	Vorbehandlung +Saxagliptin (5 mg/Tag) Vorbehandlung +Placebo

* Studiendauer orientiert sich am Auftreten der Ereignisse

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 ist zum 21.1.2013 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
D1680C00007	Die mit Sulfonylharnstoffen vorbehandelte Teilpopulation enthält zu wenige Patienten.
D1680C00003 (SAVOR-TIMI 53)	Die Studie läuft noch. Es liegen keine Ergebnisse vor.*
CV181142	Die Studie läuft noch. Es liegen keine Ergebnisse vor. Außerdem beträgt die Studiendauer nur 6 Wochen.

*Ergebnisse werden für 2013 erwartet

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.

Die umfassende bibliographische Literaturrecherche vom 04.02.2013 nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel Saxagliptin erzielte insgesamt 209 Treffer. Nach Ausschluss der Duplikate (n=17) wurden die verbleibenden 192 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den in Abschnitt 4.2.2 definierten Kriterien gesichtet. Die Studien wurden von zwei Personen unabhängig voneinander nach Relevanz selektiert. Basierend auf den Informationen des Titels und des Abstracts wurden 183 Publikationen begründet durch Ausschlusskriterien als nicht

relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Für die verbleibenden neun Publikationen, deren Titel und Abstract die Einschlusskriterien erfüllen, wurden im Anschluss die Volltexte hinsichtlich dieser Kriterien überprüft. Im Ergebnis entsprachen sieben nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C, Tabelle 4-112, Ausschlussgründe: es handelte sich nicht um eine randomisierte klinische Studie).

Somit ergab die bibliographische Literaturrecherche zwei relevante Publikationen. Bei den identifizierten Studien handelt es sich um die Publikationen von (Chacra et al., 2009; Chacra et al., 2011).

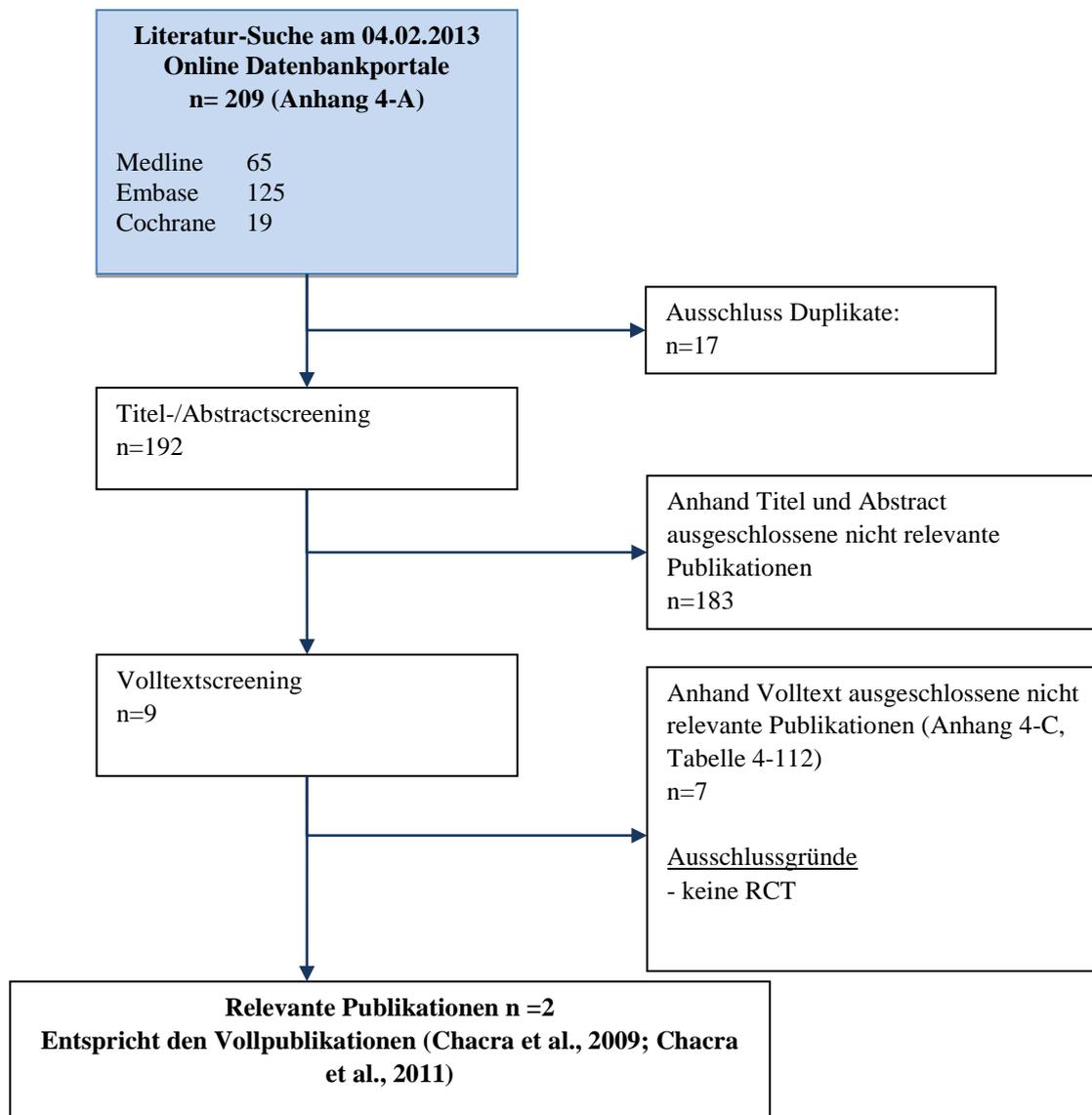


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

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)
CV181040	clinicaltrials.gov [NCT00313313] (ClinicalTrials.gov, 2013b)	ja	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.			

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

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

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-8: 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						
CV1810 40	ja	ja	nein	ja (Bristol-Myers Squibb, 2008)	ja [NCT00313313] (ClinicalTrials.gov, 2013b)	ja (Chacra et al., 2009; Chacra et al., 2011)
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
--						
<p>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.</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-F zu hinterlegen.

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

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
CV181040	RCT, doppelblind, parallel, placebokontrolliert, multizentrisch	Erwachsene Patienten mit Diabetes mellitus Typ-2 und einer submaximalen Dosis von einem Sulfonylharnstoff für mindestens 2 Monate, deren Blutzuckerwerte nicht ausreichend kontrolliert werden können (HbA1c ≥7,5% und ≤10,0%)	Glibenclamid+ Saxagliptin 2,5 mg (n=248) Glibenclamid+ Saxagliptin 5 mg (n=253) Glibenclamid+ Placebo (n=267)	Placebo-Lead-in- Phase: 4 Wochen Kurzzeitphase von 24 Wochen Langzeitphase von 52 Wochen	International 04/2006 – 09/2008	Primäres Ziel: Veränderung des HbA1c-Wertes (ausgehend vom Ausgangswert) Weitere patientenrelevante Endpunkte: Hypoglykämien Anteil an Probanden, die eine HbA1c-Wert <7,0% ohne Hypoglykämien erreicht haben Veränderung des Gewichtes Unerwünschte Ereignisse Kardiale Ereignisse

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

Studie	Sulfonylharnstoff + Saxagliptin	Sulfonylharnstoff + Placebo	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
CV181040	Glibenclamid 2,5 mg (zusätzlich zu der offenen Gabe von Glibenclamid 7,5 mg) + Saxagliptin 2,5 mg oder 5 mg	Glibenclamid 2,5 mg (zusätzlich zu der offenen Gabe von Glibenclamid 7,5 mg) + Placebo	<p>Lead-in-Phase: Einfachblinde Placebo-Gabe und (offene Gabe) von Glibenclamid 7,5 mg (Patienten mussten ihre aktuelle Einnahme von Sulfonylharnstoffen stoppen)</p> <p>Behandlung doppelblind (Double-dummy-Methode): Saxagliptin 2,5 / 5 mg+Glibenclamid versus Placebo+Glibenclamid (nach Woche 2 oder 4 konnte hochtitriert werden zu 15 mg)</p> <p>Kriterien für Hochtitrierung in Woche 2 und 4: MFPG (Mean Fasting Plasma Glucose) ≥ 100 mg/dL (5,5 mmol/L) oder MFWBG (Mean Fasting Whole Blood Glucose) ≥ 95 mg/dL (5,3 mmol/L)</p> <p>Im Fall einer Hypoglykämie konnte die offene Gabe von Glibenclamid auf 5 mg heruntertitriert werden.</p> <p>Patienten, die bestimmte glykämische Kriterien während der 24-wöchigen Behandlungsphase nicht erreichten, wurden mit bis zu 2500 mg Metformin zusätzlich als Notfallmedikation behandelt, danach darf Glibenclamid nicht mehr ab-/auftitriert werden.</p> <p>Patienten mussten aufgrund unzureichender Wirksamkeit die Studie abbrechen und konnten in die Langzeitphase zusätzlich mit open-label Metformin eintreten.</p> <p>Kriterien für Notfallmedikation: in Woche 4 und 6: MFPG > 240 mg/dL oder MFWBG > 221 mg/dL bzw. in Woche 8 MFPG > 220 mg/dL oder MFWBG > 203 mg/dL bzw. in Woche 12, 16, 20 und 24 MFPG > 200 mg/dL oder MFWBG > 185 mg/dL</p> <p>Langzeitstudie: 52 Wochen</p> <p>Kriterien für Notfallmedikation: in Woche 30, 37 und 50 der HbA1c $> 8,0\%$ war, oder in Woche 63 der HbA1c-Wert $> 7,5\%$</p>

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

Studie Gruppe	N	Alter (SD)	Geschlecht w/m (%)	BMI (kg/m ²) Mittelwert (SD)	Gewicht (kg) Mittelwert (SD)
CV181040					
Glibenclamid+Saxagliptin 2,5 mg	248	55,36 (9,58)	54,4 / 45,6	29,10 (4,50)	75,19 (14,44)
Glibenclamid+Saxagliptin 5 mg	253	54,85 (9,96)	56,5 / 43,5	29,19 (4,55)	76,22 (17,64)
Glibenclamid+Placebo	267	55,06 (10,69)	53,9 / 46,1	28,83 (4,74)	75,56 (17,35)

SD: Standardabweichung

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

Studie Gruppe	N	Baseline HbA1c (%) Mittelwert (SD)	Diabetes- dauer (Jahre) Mittelwert (SD)	Glibenclamid-Dosis (mg) Mittelwert (SD)	
CV181040				Während 24 Wochen:	Während 76 Wochen:
Glibenclamid+Saxagliptin 2,5 mg	248	8,4 (0,9)	7,1 (5,9)	7,4 (0,5)	7,3 (0,9)
Glibenclamid+Saxagliptin 5 mg	253	8,5 (0,9)	6,8 (5,8)	7,4 (0,6)	7,3 (0,8)
Glibenclamid+Placebo	267	8,4 (0,9)	6,8 (5,7)	14,6 (1,3)	15,7 (2,9)

SD: Standardabweichung

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

Studie Gruppe	N	Anzahl der Studienabbrecher n (%)		Patienten mit Notfallmedikation n (%)	
CV181040		Nach 24 Wochen:*	Nach 76 Wochen:	Nach 24 Wochen	Nach 76 Wochen:
Glibenclamid+Saxagliptin 2,5 mg	248	56 (22,6)	63 (25,4)	42 (16,9)	149 (60,1)
Glibenclamid+Saxagliptin 5 mg	253	58 (22,9)	66 (26,1)	41(16,2)	156 (61,7)
Glibenclamid+Placebo	267	91 (34,1)	82 (30,7)	78 (29,2)	179 (67,0)

*Patienten mit Notfallmedikation waren in der 24-Wochen-Analyse eine Teilmenge der Studienabbrecher. Patienten, die wegen mangelnder glykämischer Kontrolle die Notfallmedikation erhielten, konnten jedoch an der 76-Wochen-Verlängerungsphase teilnehmen.

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

In der Studie CV181040 wurden erwachsene Männer und Frauen mit Diabetes mellitus Typ 2 eingeschlossen, die einen HbA1c-Wert zwischen 7,5% und 10,0% hatten, der mit einer submaximalen Dosis an Sulfonylharnstoff in einem Zeitraum von mindestens zwei Monaten nur unzureichend kontrolliert werden konnte. Alle Studienteilnehmer waren mit einem Sulfonylharnstoff vorbehandelt. Die Patienten wurden mit 2,5 oder 5 mg Saxagliptin und Glibenclamid oder mit Placebo und Glibenclamid behandelt, wobei eine Anpassung der Glibenclamideinnahme vollzogen werden konnte. Die demographischen und krankheitsspezifischen Charakteristika zwischen den Behandlungsgruppen Saxagliptin+Glibenclamid (2,5 bzw. 5 mg) und Placebo+Glibenclamid sind sich ähnlich. Das mittlere Alter betrug 55,36 bzw. 54,85 und 55,06 Jahre, der Anteil der Frauen lag bei 54,4 bzw. 56,5 und 53,9%. Auch für die mittleren BMI-Werte (29,10 kg/m², 29,19 kg/m² und 28,83 kg/m²), das mittlere Gewicht (75,19 kg, 76,22 kg und 75,56 kg), den HbA1c-Wert (8,4, 8,5 und 8,4%) und die mittlere Krankheitsdauer (7,1, 6,8 und 6,8 Jahre) konnten keine wesentlichen Unterschiede zwischen den Behandlungsarmen festgestellt werden. Der Anteil der Studienabbrecher war in den Saxagliptin-Behandlungsgruppen (22,6% und 22,9% nach 24 Wochen und 25,4% und 26,1% nach 76 Wochen) niedriger als in der Placebo-Gruppe (34,1% nach 24 Wochen und 30,7% nach 76 Wochen). Die niedrigere Rate der Studienabbrecher nach 76 im Vergleich zu 24 Wochen erklärt sich durch die Tatsache, dass in der 24-Wochen-Analyse alle Patienten, die die Notfallmedikation erhielten, als Studienabbrecher gewertet wurden. Diese Patienten konnten jedoch trotzdem an der Verlängerungsphase teilnehmen und wurden deshalb in der 76-Wochen-Analyse nicht als Studienabbrecher gewertet. Der Anteil der Patienten, die die Notallmedikation erhielten, betrug in den Saxagliptin-Behandlungsgruppen (16,9% und 16,2% nach 24 Wochen und 60,1% und 61,7% nach 76 Wochen). In der Placebo-Gruppe (29,2% nach 24 Wochen und 67,0% nach 76 Wochen).

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

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

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
CV181040	ja	ja	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der Studie wurde eine Randomisierung mit adäquat generierter Randomisierungssequenz durchgeführt. Die Verdeckung der Gruppenzuteilung (mit Zuteilungsverhältnis 1:1:1) erfolgte zentral und unabhängig durch ein interactive voice response system (IVRS). Die Patienten und Behandler waren über die gesamte Studiendauer hinweg (76 Wochen) verblindet. Die Verblindung wurde auch bei notwendiger Titrierung des Sulfonylharnstoffs in Abhängigkeit von Laborwerten mittels einer Double-dummy-Technik aufrechterhalten und konnte im Fall eines medizinischen Notfalls oder Schwangerschaft vom Prüfarzt aufgehoben werden. Während der gesamten Studie konnte eine Notfallmedikation (Metformin) aufgrund unzureichender glykämischer Kontrolle gegeben werden (präspezifizierte Kriterien bzgl. des Nüchternblutzuckers). Die Erhebung und Auswertung der Ergebnisse zu den einzelnen definierten Endpunkten erfolgte planmäßig und wie vorab definiert. Es ergeben sich keine Anhaltspunkte für ein Verzerrungspotenzial der Studie.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

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

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

Studie	Hypoglykämien	Gewichtsveränderung	Unerwünschte Ereignisse	Folgekomplikationen	Gesamt mortalität
CV181040	ja	ja	ja	ja*	nein**

*Kardiale Ereignisse werden zur Beurteilung der patientenrelevanten Endpunkte „kardiale Morbidität und Mortalität“ verwendet. Eine gepoolte Analyse von weiteren Folgekomplikationen wurde ergänzend hinzugezogen (siehe Abschnitt 4.3.2.3). Für andere Folgekomplikationen liegen keine Daten vor; ** Mortalität war in der Studie CV181040 nicht als Endpunkt vorgesehen und wurde daher nicht adäquat im Rahmen einer Überlebenszeitanalyse untersucht. Insgesamt gab es einen Todesfall in der Glibenclamid+Saxagliptin-Gruppe und vier Todesfälle in der Glibenclamid+Placebo-Gruppe nach 74 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)
CV181040	nein	nein	nein	nein	ja

Nicht berichtete Zielgrößen

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

Es werden im Folgenden nur der Vergleich zwischen den Gruppen Glibenclamid+Saxagliptin 5 mg und Glibenclamid+Placebo beschrieben, da dies der zugelassenen Dosis für die Studienpopulation entspricht. Die Patienten in der Glibenclamid+Saxagliptin 2,5 mg-Gruppe werden dargestellt, es wird aber nicht weiter auf diese Gruppe eingegangen.

4.3.1.3.1 Hypoglykämien unter Berücksichtigung des HbA1c-Wertes – RCT

In den folgenden Abschnitten 4.3.1.3.1.1 und 4.3.1.3.1.2 wird zunächst das Ausmaß der Blutzuckerkontrolle (gemessen am HbA1c) und dann das Auftreten von Hypoglykämien getrennt voneinander operationalisiert und beschrieben. Im Anschluss werden diese Endpunkte gemeinsam betrachtet (Abschnitt 4.3.1.3.1.1). Dafür wird als Composite-Endpunkt das Erreichen des HbA1c-Zielwerts (<6,5%/<7%) ohne Auftreten von Hypoglykämien betrachtet (AstraZeneca / Bristol-Myers Squibb, 2012).

4.3.1.3.1.1 Ausmaß der Blutzuckerkontrolle (HbA1c) – RCT

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

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

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung

- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

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

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

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

Studie	Operationalisierung
CV181040	Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 24/76 (falls der Wert nicht erhoben wurde, wurde der letztmögliche vor Woche 24/76 (LOCF) betrachtet) Anzahl der Patienten, die die Studie aufgrund von unzureichender Wirksamkeit abgebrochen haben oder ein Notfallmedikament bekommen haben.

LOCF: Last observation carried forward

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

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

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

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
CV181040	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Ausmaß der Blutzuckerkontrolle (HbA1c) wurde in dieser Studie anhand der Differenz der HbA1c-Werte in Woche 24 und Woche 76 relativ zum Wert zu Studienbeginn beurteilt. Der Vergleich der Behandlungsgruppen erfolgte auf der Basis der Differenz der adjustierten

Mittelwerte mit einem Kovarianzanalyse-Modell (ANCOVA). Die Adjustierung erfolgte bezüglich des Ausgangswertes. Die Analyse basiert auf dem „Randomized Subjects Set“, der alle Patienten einschließt, die mindestens eine Dosis der Studienmedikation bekommen haben und eine Messung danach hatten. Fehlende Werte wurden durch die LOCF-Methode ersetzt. Bei Patienten mit Notfallmedikation wurde der letzte Wert vor ihrer Notfallmedikation verwendet. In der Sulfonylharnstoff+Placebo-Gruppe gab es sowohl nach 24 als auch nach 76 Wochen mehr Studienabbrecher und mehr Patienten, die eine Notfallmedikation erhielten als in der Sulfonylharnstoff+Saxagliptin-Gruppe (nach 76 Wochen haben 60,1% der Sulfonylharnstoff+Saxagliptin und 67% der Sulfonylharnstoff+Placebo-Gruppe die Studie abgebrochen). Daher wurde zusätzlich noch eine Analyse mit einem Repeated-Measures Modell vorgenommen. Eine statistische Auswertung der Ergebnisse fand wie geplant statt.

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

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

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

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+Saxagliptin				Sulfonylharnstoff+Placebo				
LOCF									
24 Wochen	2,5 mg: 248	8,36 (0,057)	7,84 (0,074)	-0,53 (0,059)	264	8,44 (0,055)	8,52 (0,077)	0,08 (0,057)	-0,61 (0,082) [-0,77;-0,45]
	5 mg: 250	8,48 (0,056)	7,84 (0,074)	-0,64 (0,059)					-0,72 (0,082) [-0,88;-0,56]
76 Wochen	2,5 mg: 246	8,36 (0,057)	8,27 (0,074)	-0,11 (0,060)	264	8,44 (0,055)	8,79 (0,067)	0,36 (0,055)	-0,47 (0,084) [-0,63;-0,31]
	5 mg: 250	8,48 (0,056)	8,27 (0,073)	-0,20 (0,060)					-0,56 (0,083) [-0,72;-0,39]
Repeated-Measures Analyse									
24 Wochen	2,5 mg: nicht berichtet	8,36 (0,057)	nicht berichtet	-0,56 (0,057)	nicht berichtet	8,44 (0,055)	nicht berichtet	0,07 (0,058)	-0,63 (0,082) [-0,79;-0,47]
	5 mg : nicht berichtet	8,49 (0,057)	nicht berichtet	-0,67 (0,057)					-0,75 (0,082) [-0,91;-0,59]°
76 Wochen	2,5 mg: nicht berichtet	8,36 (0,057)	nicht berichtet	0,11 (0,082)	nicht berichtet	8,44 (0,055)	nicht berichtet	0,69 (0,114)	-0,59 (0,140) [-0,86;-0,31]
	5 mg :nicht berichtet	8,49 (0,057)	nicht berichtet	0,03 (0,085)					-0,67 (0,142) [-0,94;-0,39]°

* adjustiert nach Ausgangswert; SD: Standardabweichung; SE: Standardfehler; KI: Konfidenzintervall

Die placebokontrollierte Studie zeigt den Vorteil von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Placebo bei der Senkung des Blutzucker-Wertes (HbA1c). In Woche 24 / 76 wird eine Reduzierung des HbA1c-Wertes um 0,64%/0,20% bei der Sulfonylharnstoff+Saxagliptin-Gruppe und eine Erhöhung um 0,08%/0,36% bei Sulfonylharnstoff+Placebo beobachtet. Der Unterschied zwischen den zwei Behandlungsgruppen gegenüber Placebo ist signifikant sowohl nach 24 als auch nach 76 Wochen. Die Repeated-Measures Analyse bestätigt die Robustheit der Ergebnisse.

Als unterstützende Wirksamkeitsmessung wurden die Studienabbrüche (aufgrund unzureichender Blutzuckerkontrolle) oder die Einnahme der Notfallmedikation (aufgrund unzureichender Blutzuckerkontrolle verabreicht) erhoben.

Tabelle 4-20: Ergebnisse für Studienabbrüche (aufgrund von unzureichender Blutzuckerkontrolle) oder die Einnahme der Notfallmedikation (aufgrund unzureichender Blutzuckerkontrolle verabreicht) aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle, n (%)
CV181040		
Nach 24 Wochen		
Sulfonylharnstoff+Saxagliptin 2,5 mg	248	45 (18,1)
Sulfonylharnstoff+Saxagliptin 5 mg	253	42 (16,6)
Sulfonylharnstoff+Placebo	267	79 (29,6)
2,5 mg vs. Placebo OR [95%-KI]		0,53 [0,35;0,80]
5 mg vs. Placebo OR [95%-KI]		0,47 [0,31;0,72]
Nach 76 Wochen		
Sulfonylharnstoff+Saxagliptin 2,5 mg	248	157 (63,3)
Sulfonylharnstoff+Saxagliptin 5 mg	253	157 (62,1)
Sulfonylharnstoff+Placebo	267	200 (74,9)
2,5 mg vs. Placebo OR [95%-KI]		0,58 [0,40;0,84]
5 mg vs. Placebo OR [95%-KI]		0,55 [0,38;0,80]

KI: Konfidenzintervall

Die Behandlung mit Sulfonylharnstoff+Saxagliptin war bezogen auf die HbA1c-Reduktion in beiden Dosisgruppen nach 24 Wochen und nach 76 Wochen signifikant besser als die Behandlung mit Sulfonylharnstoff+Placebo. 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+Saxagliptin-Gruppen.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorlag.

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

4.3.1.3.1.2 Hypoglykämien – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst

werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

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

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

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

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

Tabelle 4-21: Operationalisierung von Hypoglykämien

Studie	Operationalisierung
CV181040	Anzahl von Patienten mit mindestens einer Hypoglykämie innerhalb der 24/76 Wochen der Behandlungsphase (patientenberichtete Hypoglykämien) Anzahl von Patienten mit mindestens einer bestätigten symptomatischen Hypoglykämie innerhalb der 24/76 Wochen (Blutzuckerwert ≤ 50 mg/dL) (bestätigte Hypoglykämien)

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

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

Tabelle 4-22: 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
CV181040	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Hypoglykämien wurden im Rahmen der Erhebung der unerwünschten Ereignisse auf einer speziellen Case Report Form (CRF)-Seite erhoben. Es wurden sowohl bestätigte (symptomatisch, Blutzuckerwert ≤ 50 mg/dL) als auch patientenberichtete Hypoglykämien (anhand einer prädefinierten Liste von Preferred Terms, die eine Diagnose einer Hypoglykämie darstellt) berichtet. Der Endpunkt wurde mittels des „Treated Subjects Data Set“ ausgewertet, d.h. alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben. In der Kurzzeitphase wurden für die Auswertung der Hypoglykämien die Daten der Patienten vor der Notfallmedikation verwendet. Zusätzlich wurden auch die bestätigten Hypoglykämien mit den Patienten, die auch eine Notfallmedikation bekommen haben, ausgewertet. Auch in der Langzeitphase wurden die Daten der primären Analyse einschließlich der Patienten mit Notfallmedikation ausgewertet. Zusätzlich wurde auch die Analyse der Daten der Patienten ohne Notfallmediaktion vorgenommen.

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 4-23: Ergebnisse für patientenberichtete Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel

Studie CV181040	Zu bewertendes Arzneimittel N / n (%)	Vergleichstherapie N / n (%)		Odds Ratio [95%-KI]
24 Wochen	Sulfonylharnstoff+Saxagliptin N/n (%)	Sulfonylharnstoff+Placebo N/n (%)		
Patienten mit mind. einer Hypoglykämie*	2,5 mg: 248 33 (13,3)	267	27 (10,1)	1,36 [0,79;2,34]
	5 mg: 253 37 (14,6)			1,52 [0,90;2,58]
76 Wochen	Sulfonylharnstoff+Saxagliptin N/n (%)	Sulfonylharnstoff+Placebo N/n (%)		
Patienten mit mind einer Hypoglykämie*	2,5 mg: 248 40 (16,1)	267	32 (12,0)	1,41 [0,86;2,33]
	5 mg: 253 46 (18,2)			1,63 [1,00;2,66]
Patienten mit mind einer Hypoglykämie**	2,5 mg: 248 60 (24,2)	267	55 (20,6)	1,23 [0,81;1,86]
	5 mg: 253 58 (22,9)			1,15 [0,76;1,74]

* Daten ohne Patienten mit Notfallmedikation

** Daten einschließlich der Patienten mit Notfallmedikation

KI: Konfidenzintervall

Tabelle 4-24: Ergebnisse für bestätigte Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel

Studie CV181040	Zu bewertendes Arzneimittel N/n (%)	Vergleichstherapie N/n (%)		Odds Ratio [95%-KI]
24 Wochen	Sulfonylharnstoff+Saxagliptin N/n (%)	Sulfonylharnstoff+Placebo N/n (%)		
Patienten mit mind einer Hypoglykämie*	2,5 mg: 248 6 (2,4)	267	2 (0,7)	3,29 [0,66;16,43]
	5 mg: 253 2 (0,8)			1,06 [0,15;7,55]
Patienten mit mind einer Hypoglykämie**	2,5 mg: 248 6 (2,4)	267	2 (0,7)	3,29 [0,66;16,43]
	5 mg: 253 2 (0,8)			1,06 [0,15;7,55]
76 Wochen	Sulfonylharnstoff+Saxagliptin N/n (%)	Sulfonylharnstoff+Placebo N/n (%)		
Patienten mit mind einer Hypoglykämie *	2,5 mg: 248 8 (3,2)	267	5 (1,9)	1,75 [0,56;5,41]
	5 mg: 253 4 (1,6)			0,84 [0,22;3,17]
Patienten mit mind einer Hypoglykämie **	2,5 mg: 248 13 (5,2)	267	9 (3,4)	1,59 [0,67;3,78]
	5 mg: 253 8 (3,2)			0,94 [0,36;2,47]

* Daten ohne Patienten mit Notfallmedikation

** Daten einschließlich der Patienten mit Notfallmedikation

KI: Konfidenzintervall

Nach 24 Wochen berichteten in der Sulfonylharnstoff+Saxagliptin-Gruppe 14,6% der Patienten über Hypoglykämien, während in der Sulfonylharnstoff+Placebo-Gruppe 10,1% der Patienten patientenberichtete Hypoglykämie hatten. Nach 76 Wochen berichteten 22,9% der

mit Sulfonylharnstoff+Saxagliptin behandelten Patienten eine Hypoglykämie, während in der Sulfonylharnstoff+Placebo-Gruppe 20,6% der Patienten eine Hypoglykämie berichteten.

Bestätigte Hypoglykämien (Blutzuckerwert ≤ 50 mg/dL) kamen nach 24 Wochen in der Sulfonylharnstoff+Saxagliptin-Gruppe bei 0,8 % der Patienten vor. Nach 76 Wochen traten bei 3,2% der Sulfonylharnstoff+Saxagliptin-Patienten Hypoglykämien auf. In der Sulfonylharnstoff+Placebo-Gruppe hatten 0,7% der Patienten nach 24 Wochen und 3,4% der Patienten nach 76 Wochen bestätigte Hypoglykämien.

Sowohl bei den patientenberichteten als auch bei den bestätigten Hypoglykämien zeigten die Odds Ratio nach 24 und 76 Wochen keinen signifikanten Unterschied.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorlag.

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 Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

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

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

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

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

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

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

Tabelle 4-25: Operationalisierung von gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

Studie	Operationalisierung
CV181040	Anhand der Häufigkeiten von Hypoglykämien unter Berücksichtigung der Blutzuckersenkung Anhand von Responder-Analysen: Anzahl der Patienten („Responder“), die einen Ausgangswert des HbA1c $\geq 7,0\%$ aufweisen und einen HbA1c von unter 7,0% ohne Hypoglykämien bis Woche 24 und 76 erreicht haben (LOCF) Anzahl der Patienten („Responder“), die einen HbA1c von unter 6,5% ohne Hypoglykämien bis Woche 24/76 erreicht haben (LOCF)

LOCF: Last observation carried forward

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

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

Tabelle 4-26: Bewertung des Verzerrungspotenzials für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
CV181040	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

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

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

Das Auftreten von Hypoglykämien kann nur unter Berücksichtigung der Blutzuckersenkung sinnvoll beurteilt werden. Deswegen wird im Folgenden die post hoc berechnete Anzahl der Patienten, die die glykämische Response (HbA1c <6,5%/<7%) erreichten, ohne dass eine Hypoglykämie aufgetreten ist, dargestellt (AstraZeneca / Bristol-Myers Squibb, 2012).

Tabelle 4-27: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 6,5% ohne Hypoglykämien erreicht haben

Studie	Zu bewertendes Arzneimittel N/n (%)		Vergleichstherapie N/n (%)		Odds Ratio [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo		
24 Wochen	5 mg: 250	16 (6,4)	264	5 (1,9)	3,54 [1,28;9,82]
76 Wochen	5 mg: 251	6 (2,4)	264	3 (1,1)	2,13 [0,53;8,61]

KI: Konfidenzintervall

Nach 24 Wochen hatten signifikant mehr Patienten in der Sulfonylharnstoff+Saxagliptin-Gruppe einen HbA1c-Wert unter 6,5%, ohne dass eine Hypoglykämie aufgetreten ist. Nach 76 Wochen war der Unterschied zwischen den Behandlungsoptionen nicht mehr signifikant.

Tabelle 4-28: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 7% ohne Hypoglykämien erreicht haben (exklusive Patienten mit HbA1c <7% zu Baseline)

Studie	Zu bewertendes Arzneimittel N/n (%)		Vergleichstherapie N/n (%)		Odds Ratio [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo		
24 Wochen	5 mg: 247	43 (17,4)	258	14 (5,4)	3,54 [1,28;9,80]
76 Wochen	5 mg: 247	18 (7,3)	258	7 (2,7)	2,82 [1,16;6,87]

KI: Konfidenzintervall

Die Patienten der Sulfonylharnstoff+Saxagliptin-Gruppe erreichten zu beiden Zeitpunkten signifikant häufiger einen HbA1c-Wert von unter 7% ohne dass eine Hypoglykämie aufgetreten ist als Patienten in der Sulfonylharnstoff+Placebo-Gruppe.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorlag.

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 4-29: Operationalisierung von Gewichtsveränderung

Studie	Operationalisierung
CV181040	Änderung des Körpergewichts in kg von Studienbeginn bis Woche 24/76 (falls der Wert nicht erhoben wurde, wurde der letztmögliche vor Woche 24/76 (LOCF) betrachtet). Zusätzlich wurde ein Repeated-Measures angewendet.

LOCF= Last Observation Carried Forward

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

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

Tabelle 4-30: 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
CV181040	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Eine Gewichtsveränderung war in der Studie ein Wirksamkeitsendpunkt und wurde anhand der Differenz des Körpergewichts in Woche 24 und Woche 76 relativ zum Wert zu Studienbeginn beurteilt. Die Adjustierung erfolgte nur in Woche 24 bezüglich des Ausgangswertes. Die Analyse basiert auf dem „Randomized Subjects Set“, der alle Patienten, die mindestens eine Dosis der Studienmedikation bekommen haben und eine Messung danach hatten, einschließt. Fehlende Werte wurden durch die LOCF-Methode ersetzt. Bei Patienten mit Notfallmedikation wurde der letzte Wert vor ihrer Notfallmedikation verwendet. Es gab mehr Studienabbrecher sowie mehr Patienten mit Notfallmedikation in der Sulfonylharnstoff+Placebo-Gruppe als in der Sulfonylharnstoff+Saxagliptin-Gruppe sowohl nach 24 als auch nach 76 Wochen (nach 76 Wochen haben 60,1% der Sulfonylharnstoff+Saxagliptin und 67% der Sulfonylharnstoff+Placebo-Gruppe die Studie abgebrochen). Daher wurde zusätzlich eine Analyse mit beobachteten Werten vorgenommen, um die Robustheit der Ergebnisse zu bestätigen. Eine statistische Auswertung der Ergebnisse fand wie geplant statt.

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 4-31: Ergebnisse für Gewichtsveränderung (Änderung im Vergleich zum Ausgangswert) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertendes Arzneimittel				Vergleichstherapie				Gruppenunterschied der adjustierten Mittelwerte in kg (SE) [95%-KI]
	N	Ausgangswert Mittelwert (SE)	Endwert Mittelwert (SE)	Veränderung (in kg) Mittelwert* (SE)	N	Ausgangswert Mittelwert (SE)	Endwert Mittelwert (SE)	Veränderung (in kg) Mittelwert* (SE)	
CV181040	Sulfonylharnstoff+Saxagliptin				Sulfonylharnstoff+Placebo				
LOCF									
24 Wochen	2,5 mg: 247	75,1 (0,92)	75,9 (0,93)	0,7 (0,14)	265	75,6 (1,07)	76,0 (1,09)	0,3 (0,13)	0,4 (0,19) [0,0;0,8]
	5 mg: 253	76,2 (1,11)	77,0 (1,10)	0,8 (0,13)					0,5 (0,19) [0,1;0,8]
76 Wochen	2,5 mg: 247	75,1 (0,92)	75,9 (0,93)	0,8 (0,15)	265	75,6 (1,07)	75,9 (1,09)	0,3 (0,15)	0,50 (0,2121) [0,08;0,92]°
	5 mg: 253	76,2 (1,11)	77,2 (1,12)	1,0 (0,15)					0,70 (0,2121) [0,28;1,12]°
Observed Cases									
24 Wochen	2,5 mg: 191	74,9 (1,06)	75,7 (1,08)	0,8 (0,16)	178	74,2 (1,11)	74,7 (1,15)	0,5 (0,18)	0,30 (0,2408) [-0,17;0,77]°
	5 mg: 196	75,9 (1,18)	76,8 (1,16)	0,9 (0,16)					0,40 (0,2408) [-0,07;0,87]°
76 Wochen	2,5 mg: 61	78,3 (2,14)	78,6 (2,10)	0,3 (0,42)	27	75,8 (2,31)	75,5 (2,23)	-0,4 (0,55)	0,70 (0,6920) [-0,66;2,06]°
	5 mg: 57	78,6 (2,55)	80,0 (2,62)	1,4 (0,48)					1,80 (0,7300) [0,73;3,23]°

* adjustierter Mittelwert nach Ausgangswert; SD: Standardabweichung; SE: Standardfehler; KI: Konfidenzintervall; °eigene Berechnung

Die Studie CV181040 zeigte nach 24 sowie nach 76 Wochen eine statistisch signifikant höhere Gewichtszunahme in der Sulfonylharnstoff+Saxagliptin-Gruppe als in der Kontrollgruppe, wobei die Gewichtszunahme in beiden Behandlungsoptionen gering war (bis 1,0 kg nach 76 Wochen in der Sulfonylharnstoff+Saxagliptin-Gruppe und 0,3 kg nach 76 Wochen in der Sulfonylharnstoff+Placebo-Gruppe).

Bei der Analyse der beobachteten Werte lag nach 24 Wochen kein signifikantes Ergebnis mehr vor. Allerdings konnte nach 76 Wochen ein signifikanter Unterschied zwischen der Gruppe mit Sulfonylharnstoff+5 mg Saxagliptin und der Gruppe mit Sulfonylharnstoff+Placebo zugunsten von Sulfonylharnstoff+Placebo beobachtet werden.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorlag.

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

4.3.1.3.1.3 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 4-32: Operationalisierung von Unerwünschte Ereignisse

Studie	Operationalisierung ^a
CV181040	Gesamtrate unerwünschter Ereignisse Schwerwiegende unerwünschte Ereignisse Studienabbruch aufgrund unerwünschter Ereignisse

^a Hinsichtlich der unerwünschten Ereignisse und der schwerwiegenden unerwünschten Ereignisse wurden die in der Studie vorhandenen Definitionen verwendet

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

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

Tabelle 4-33: 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
CV181040	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt „unerwünschte Ereignisse“ in der Studie ist niedrig. Die Überwachung und Erfassung unerwünschter Ereignisse erfolgte in der Studie gemäß Good Clinical Practice (GCP) Standards. Die Endpunkterheber waren zu jedem Zeitpunkt der Erhebung verblindet. Der Endpunkt wurde bei behandelten Patienten, die mindestens eine Dosis der Studienmedikation bekommen haben, erhoben (Treated Subjects Data Set). Eine statistische Auswertung der Ergebnisse fand wie geplant statt. In der Kurzzeitphase wurden die unerwünschten Ereignisse mit den Daten vor Erhalt der Notfallmedikation ausgewertet. Sensitivitätsanalysen wurden mit den Daten einschließlich der

Patienten mit Notfallmedikation durchgeführt. In der Langzeitphase erfolgte für die primäre Analyse die Auswertung der Daten einschließlich der Patienten mit Notfallmedikation.

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 4-34: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate unerwünschter Ereignisse, n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
CV181040¹				
Nach 24 Wochen ²				
Sulfonylharnstoff+Saxagliptin 2,5 mg	248	186 (75,0)	4 (1,6)	3 (1,2)
Sulfonylharnstoff+Saxagliptin 5 mg	253	183 (72,3)	6 (2,4)	8 (3,2)
Sulfonylharnstoff +Placebo	267	205 (76,8)	6 (2,2)	4 (1,5)
2,5 mg: OR [95%-KI]		0,91 [0,61;1,36]	0,71 [0,2;2,56]	0,81 [0,18;3,63]
5 mg: OR [95%-KI]		0,79 [0,53;1,17]	1,06 [0,34;3,32]	2,15 [0,64;7,22]
Nach 24 Wochen ³				
Sulfonylharnstoff+Saxagliptin 2,5 mg	248	190 (76,6)	4 (1,6)	4 (1,6)
Sulfonylharnstoff+Saxagliptin 5 mg	253	185 (73,1)	6 (2,4)	7 (2,8)
Sulfonylharnstoff +Placebo	267	212 (79,4)	7 (2,6)	4 (1,5)
2,5 mg: OR [95%-KI]		0,85 [0,56;1,29]	0,61 [0,18;2,11]	1,91 [0,55;6,60]
5 mg: OR [95%-KI]		0,71 [0,47;1,06]	0,90 [0,30;2,72]	1,06 [0,26;4,27]
Nach 76 Wochen ³				
Sulfonylharnstoff+Saxagliptin 2,5 mg	248	212 (85,5)	18 (7,3)	15 (6,0)
Sulfonylharnstoff+Saxagliptin 5 mg	253	215 (85,0)	12 (4,6)	16 (6,3)
Sulfonylharnstoff +Placebo	267	237 (88,8)	23 (8,6)	12 (4,5)
2,5 mg: OR [95%-KI]		0,75 [0,44;1,25]	0,83 [0,44;1,58]	1,34 [0,61;2,92]
5 mg: OR [95%-KI]		0,72 [0,43;1,20]	0,53 [0,26;1,09]	1,43 [0,66;3,10]

¹ Bei der Betrachtung der unerwünschten Ereignisse sind die Hypoglykämien eingeschlossen

² Daten ohne Patienten mit Notfallmedikation

³ Daten einschließlich der Patienten mit Notfallmedikation

KI: Konfidenzintervall

In der Studie CV181040 unterscheidet sich der Anteil der Patienten mit unerwünschten Ereignissen, schwerwiegenden unerwünschten Ereignissen und Studienabbrüchen aufgrund unerwünschter Ereignisse statistisch nicht zwischen den jeweiligen Behandlungsgruppen.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorlag.

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.

4.3.1.3.1.4 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 4-35: Operationalisierung von Kardiale Ereignisse

Studie	Operationalisierung
CV181040	Anzahl der Patienten mit mindestens einem kardialen Ereignis innerhalb von 24 bzw. 76 Wochen. Die Erhebung der kardialen Ereignisse erfolgte im Rahmen der Erhebung der unerwünschten Ereignisse (SOC)

SOC: Systemorganklasse

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

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

Tabelle 4-36: 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
CV181040	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) kardialer 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 verblindet. Der Endpunkt wurde bei behandelten Patienten, die mindestens eine Studienmedikation bekommen hatten, erhoben.

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 4-37: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
CV181040		
Nach 24 Wochen		
Sulfonylharnstoff+Saxagliptin 2,5 mg	248	5 (2,0)
Sulfonylharnstoff+Saxagliptin 5 mg	253	10 (4,0)
Sulfonylharnstoff+Placebo	267	10 (3,7)
2,5 mg: OR [95%-KI]		0,53 [0,18;1,57]
5 mg: OR [95%-KI]		1,06 [0,43;2,59]
Nach 76 Wochen		
Sulfonylharnstoff+Saxagliptin 2,5 mg	248	14 (5,6)
Sulfonylharnstoff+Saxagliptin 5 mg	253	15 (5,9)
Sulfonylharnstoff +Placebo	267	15 (5,6)
2,5 mg: OR [95%-KI]		1,01 [0,47;2,13]
5 mg: OR [95%-KI]		1,06 [0,51;2,21]

KI: Konfidenzintervall

In der Studie CV181040 wurde nach 24 sowie nach 76 Wochen kein signifikanter Unterschied in den Behandlungsgruppen detektiert.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie vorlag.

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.

4.3.1.3.2 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 alle Endpunkte werden Subgruppenanalysen für die folgenden Merkmale dargestellt (AstraZeneca / Bristol-Myers Squibb, 2012):

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

Die Wahl der Alters-Cuts beruht darauf, dass die DDG den älteren Patienten als in der Regel älter als 65 Jahre definiert (Hader C et al., 2004). Das entspricht annähernd den oben dargestellten a priori definierten Alters-Cuts. Dadurch können Subgruppenanalysen für Erwachsene (<65 Jahre) und ältere (≥ 65 Jahre) Patienten dargestellt werden.

Diabetes mellitus kann gemäß der DDG in die folgenden Stadien eingeteilt werden (Kerner et al., 2004; AkdÄ, 2009):

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

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

Die nachfolgenden Subgruppen-Analysen basieren auf der LOCF-Methode. Nach 24 Wochen wurden die Analysen mit den Daten ohne Patienten mit Notfallmedikation vorgenommen, nach 76 Wochen wurden die Daten inklusive der Patienten, die eine Notfallmedikation erhielten, in die Analyse mit eingeschlossen. Bei den Sicherheitsendpunkten (Hypoglykämien, unerwünschte Ereignisse und kardiale Ereignisse) wurden zu beiden Zeitpunkten die Daten inklusive der Patienten mit Notfallmedikation berücksichtigt. Des Weiteren wurde in den Subgruppen nur die Saxagliptin 5 mg (nicht mehr die 2,5 mg) berücksichtigt.

Dadurch, dass nur eine Studie für die Nutzenbewertung des zu bewertenden Arzneimittel zur Verfügung steht, kann keine Meta-Analyse durchgeführt werden

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

Tabelle 4-38: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter

Studie	Zu bewertendes Arzneimittel		Vergleichstherapie			Gruppenunterschied Mittelwert % (SE) [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo			
24 Wochen	N / Mittelwert ¹ / SE		N / Mittelwert ¹ / SE			
Alter <65 Jahre	5 mg: 209	-0,62 0,064	212	0,07 0,064		-0,69 (0,091) [-0,87;-0,51]
Alter ≥65 Jahre	5 mg: 41	-0,74 0,145	52	0,13 0,130		-0,87 (0,195) [-1,47;-0,06]
Interaktions- test ²	5 mg vs. Placebo					p=0,5930
76 Wochen						
Alter <65 Jahre	5 mg: 209	-0,14 0,065	212	0,38 0,065		-0,52 (0,092) [-0,70;-0,34]
Alter ≥65 Jahre	5 mg: 41	-0,49 0,147	52	0,26 0,131		-0,75 (0,197) [-1,14;-0,37]
Interaktions- test ²	5 mg vs. Placebo					p=0,5347

SE: Standardfehler; ¹ Änderung des absoluten HbA1c-Wertes (in %) im Vergleich zum Ausgangswert; ² Interaktionstest basierend auf dem ANCOVA-Modell; KI: Konfidenzintervall

Tabelle 4-39: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter

Studie	N	Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle, N/n (%)
CV181040		
Nach 24 Wochen		
Alter <65 Jahre		
Sulfonylharnstoff+Saxagliptin	211	37 (17,5)
Sulfonylharnstoff+Placebo	215	70 (32,6)
OR [95%-KI]		0,44 [0,28;0,69]
Alter ≥65 Jahre		
Sulfonylharnstoff+Saxagliptin	42	5 (11,9)
Sulfonylharnstoff+Placebo	52	9 (17,3)
OR [95%-KI]		0,65 [0,20;2,10]
Interaktionstest*	5 mg vs. Placebo	I ² =0%; p=0,55
Nach 76 Wochen		
Alter <65 Jahre		
Sulfonylharnstoff+Saxagliptin	211	138 (65,4)
Sulfonylharnstoff+Placebo	215	164 (76,3)
OR [95%-KI]		0,59 [0,38;0,90]
Alter ≥65 Jahre		
Sulfonylharnstoff+Saxagliptin	42	19 (45,2)
Sulfonylharnstoff+Placebo	52	36 (69,2)
OR [95%-KI]		0,37 [0,16;0,86]
Interaktionstest*	5 mg vs. Placebo	I ² =0%; p=0,33

* eigene Berechnung; KI: Konfidenzintervall

In der Studie CV181040 ergab der Interaktionstest nach 24 und nach 76 Wochen keinen Hinweis auf eine Effektmodifikation durch das Alter der Patienten ($p=0,5930$ und $p=0,5347$). In beiden Altersklassen war die Reduktion des HbA1c-Wertes nach Woche 24 und 76 in der Sulfonylharnstoff+Saxagliptin-Gruppe signifikant höher als in der Sulfonylharnstoff+Placebo-Gruppe. Anhand der zusätzlichen Operationalisierung (Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle) konnte keine Interaktion festgestellt werden.

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

Tabelle 4-40: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht

Studie	Zu bewertendes Arzneimittel		Vergleichstherapie			Gruppenunterschied Mittelwert % (SE) [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo			
24 Wochen	N / Mittelwert¹ / SE		N / Mittelwert¹ / SE			
männlich	5 mg: 108	-0,87 0,088	121	-0,14 0,083		-0,72 (0,121) [-0,96;-0,48]
weiblich	5 mg: 142	-0,47 0,077	143	0,27 0,076		-0,74 (0,108) [-0,95;-0,53]
Interaktions- test	5 mg vs. Placebo					p=0,9560
76 Wochen						
männlich	5 mg: 108	-0,32 0,09	121	0,13 0,085		-0,45 (0,124) [-0,69;-0,21]
weiblich	5 mg: 142	-0,10 0,078	143	0,55 0,078		-0,65 (0,111) [-0,87;-0,44]
Interaktions- test	5 mg vs. Placebo					p=0,0580

SE: Standardfehler; ¹ adjustierter Mittelwert (Änderung des absoluten HbA1c-Wertes (in %) im Vergleich zum Ausgangswert); KI: Konfidenzintervall

Tabelle 4-41: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht

Studie	N	Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle, n (%)
CV181040		
Nach 24 Wochen		
männlich		
Sulfonylharnstoff+Saxagliptin	110	12 (10,9)
Sulfonylharnstoff+Placebo	123	30 (24,4)
OR [95%-KI]		0,38 [0,18;0,79]
weiblich		
Sulfonylharnstoff+Saxagliptin	143	30 (21,0)
Sulfonylharnstoff+Placebo	144	49 (34,0)
OR [95%-KI]		0,51 [0,30;0,87]
Interaktionstest*	5 mg vs. Placebo	I ² =0%; p=0,51
Nach 76 Wochen		
männlich		
Sulfonylharnstoff+Saxagliptin	110	60 (54,5)
Sulfonylharnstoff+Placebo	123	81 (65,9)
OR [95%-KI]		0,62 [0,37;1,06]
weiblich		
Sulfonylharnstoff+Saxagliptin	143	97 (67,8)
Sulfonylharnstoff+Placebo	144	119 (82,6)
OR [95%-KI]		0,44 [0,28;0,77]
Interaktionstest*	5 mg vs. Placebo	I ² =0%; p=0,39

* eigene Berechnung; KI: Konfidenzintervall

Die Interaktionstests für die Studie CV181040 nach 24 Wochen ergaben keine Hinweise auf eine Effektmodifikation durch das Geschlecht der Patienten (p=0,956). Jedoch gab es nach 76 Wochen eine signifikante Interaktion durch das Geschlecht (p=0,058). In beiden Geschlechtsgruppen gab es eine signifikante Reduktion des HbA1c-Wertes zwischen den Behandlungsgruppen zu Gunsten der Sulfonylharnstoff+Saxagliptin-Gruppe. Da die Interaktion quantitativ war, ergab sich keine Änderung in der Gesamtinterpretation des Ergebnisses. Auch anhand der Operationalisierungen Studienabbrüche oder Patienten mit Notfallmedikation aufgrund unzureichender glykämischer Kontrolle konnte keine Interaktion festgestellt werden.

4.3.1.3.2.3 Hypoglykämien, stratifiziert nach dem Alter – RCT

Tabelle 4-42: Ergebnisse für patientenberichtete Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter

Studie	Zu bewertendes Arzneimittel N/n (%)		Vergleichstherapie N/n (%)		Odds Ratio [95%-KI] p-Wert
CV181040	Sulfonylharnstoff+ Saxagliptin		Sulfonylharnstoff+ Placebo		
24 Wochen	N/n (%)		N/n (%)		
Alter <65 Jahre	5 mg: 211	32 (15,2)	215	21 (9,8)	1,65 [0,92;2,97]
Alter ≥65 Jahre	5 mg: 42	5 (11,9)	52	6 (11,5)	1,04 [0,29;3,66]
Interaktions- test*	5 mg vs. Placebo	I ² =0%; p=0,51			
76 Wochen					
Alter <65 Jahre	5 mg: 211	47 (22,3)	215	41 (19,1)	1,22 [0,76;1,95]
Alter ≥65 Jahre	5 mg: 42	11 (26,2)	52	14 (26,9)	0,96 [0,38;2,42]
Interaktions- test*	5 mg vs. Placebo	I ² =0%; p=0,66			

* eigene Berechnung; KI: Konfidenzintervall

Tabelle 4-43: Ergebnisse für bestätigte Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter

Studie	Zu bewertendes Arzneimittel N/n (%)		Vergleichstherapie N/n (%)		Odds Ratio [95%-KI] p-Wert
CV181040	Sulfonylharnstoff+ Saxagliptin		Sulfonylharnstoff+ Placebo		
24 Wochen	N/n (%)		N/n (%)		
Alter <65 Jahre	5 mg: 211	2 (0,9)	215	1 (0,5)	2,05 [0,18;22,76]
Alter ≥65 Jahre	5 mg: 42	0 (0,0)	52	1 (1,9)	0,4 [0,02;10,17]
Interaktionstest*	5 mg vs. Placebo	I ² =0%; p=0,43			
76 Wochen					
Alter <65 Jahre	5 mg: 211	7 (3,3)	215	7 (3,3)	1,02 [0,36;2,85]
Alter ≥65 Jahre	5 mg: 42	1 (2,4)	52	2 (3,8)	0,62 [0,06;6,59]
Interaktionstest*	5 mg vs. Placebo	I ² =0%; p=0,71			

* eigene Berechnung; KI: Konfidenzintervall

Die Interaktionstests für den Endpunkt patientenberichtete Hypoglykämien ergaben sich für die Studie CV181040 weder nach 24 noch nach 76 Wochen Hinweise auf eine Effektmodifikation durch das Alter der Patienten (I²=0%; p=0,51 und I²=0%; p=0,66). Zudem gab es unabhängig vom Alter keinen signifikanten Unterschied beim Auftreten von bestätigten Hypoglykämien zwischen den Behandlungsgruppen (I²=0%; p=0,43 und I²=0%, p=0,71).

4.3.1.3.2.4 Hypoglykämien, stratifiziert nach dem Geschlecht – RCT

Tabelle 4-44: Ergebnisse für patientenberichtete Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht

	Zu bewertendes Arzneimittel N/n (%)		Vergleichstherapie N/n (%)		Odds Ratio [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+ Placebo		
24 Wochen	N/n (%)		N/n (%)		
männlich	5 mg: 110	15 (13,6)	123	14 (11,4)	1,23 [0,56;2,68]
weiblich	5 mg: 143	22 (15,4)	144	13 (9,0)	1,83 [0,88;3,8]
Interaktionstest*	5 mg vs. Placebo	I ² =0%; p=0,46			
76 Wochen					
männlich	5 mg: 110	20 (18,2)	123	25 (20,3)	0,87 [0,45;1,68]
weiblich	5 mg: 143	38 (26,6)	144	30 (20,8)	1,36 [0,79;2,36]
Interaktionstest*	5 mg vs. Placebo	I ² =5,6%; p=0,3			

* eigene Berechnung; KI: Konfidenzintervall

Tabelle 4-45: Ergebnisse für bestätigte Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht

Studie	Zu bewertendes Arzneimittel N/n (%)		Vergleichstherapie N/n (%)		Odds Ratio [95%-KI] p-Wert
CV181040	Sulfonylharnstoff+ Saxagliptin		Sulfonylharnstoff+ Placebo		
24 Wochen	N/n (%)		N/n (%)		
männlich	5 mg: 110	1 (0,9)	123	2 (1,6)	0,56 [0,05;6,21]
weiblich	5 mg: 143	1 (0,7)	144	0 (0,0)	3,04 [0,12;75,30]
Interaktions- test*	5 mg vs. Placebo	I ² =0%; p=0,41			
76 Wochen					
männlich	5 mg: 110	3 (2,7)	123	5 (4,1)	0,66 [0,15;2,84]
weiblich	5 mg: 143	5 (3,5)	144	4 (2,8)	1,27 [0,33;4,82]
Interaktions- test*	5 mg vs. Placebo	I ² =0%; p=0,52			

* eigene Berechnung; KI: Konfidenzintervall

In der Studie CV181040 haben Männer und Frauen in beiden Behandlungsgruppen in etwa das gleich große Risiko für eine Hypoglykämie (patientenberichtet und bestätigt). Es liegt somit kein Hinweis auf eine Effektmodifikation durch das Geschlecht vor.

4.3.1.3.2.5 Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien, stratifiziert nach dem Alter – RCT

Tabelle 4-46: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 6,5% ohne Hypoglykämien erreicht haben aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter

Studie	Zu bewertendes Arzneimittel N/n (%)		Vergleichstherapie N/n (%)		Odds Ratio [95%-KI] p-Wert
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo		
24 Wochen	N/n (%)		N/n (%)		
Alter <65 Jahre	5 mg: 209	12 (5,7)	212	3 (1,4)	4,24 [1,18;15,26]
Alter ≥65 Jahre	5 mg: 41	4 (9,5)	52	2 (3,8)	2,70 [0,47;15,55]
Interaktions-test*	5 mg vs. Placebo				I ² =0%; p=0,68
76 Wochen					
Alter <65 Jahre	5 mg: 210	3 (1,4)	212	2 (0,9)	1,52 [0,25;9,20]
Alter ≥65 Jahre	5 mg: 41	3 (7,1)	52	1 (1,9)	4,03 [0,40;40,23]
Interaktions-test*	5 mg vs. Placebo				I ² =0%; p=0,51

* eigene Berechnung; KI: Konfidenzintervall

Tabelle 4-47: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 7% ohne Hypoglykämien erreicht haben (exklusive Patienten mit HbA1c <7% Ausgangswert) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter

Studie	Zu bewertendes Arzneimittel N/n (%)		Vergleichstherapie N/n (%)		Odds Ratio [95%-KI] p-Wert
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo		
24 Wochen	N/n (%)		N/n (%)		
Alter <65 Jahre	5 mg: 206	35 (17,0)	209	11 (5,3)	3,68 [1,82;7,48]
Alter ≥65 Jahre	5 mg: 41	8 (19,5)	49	3 (6,1)	3,72 [0,92;15,08]
Interaktions- test*	5 mg vs. Placebo				I ² =0%; p=0,99
76 Wochen					
Alter <65 Jahre	5 mg: 206	12 (5,8)	209	3 (1,4)	4,25 [1,18;15,28]
Alter ≥65 Jahre	5 mg: 41	6 (14,6)	49	4 (8,2)	1,93 [0,50;7,37]
Interaktions- test*	5 mg vs. Placebo				I ² =0%; p=0,40

* eigene Berechnung; KI: Konfidenzintervall

Die Interaktionstests für die Studie CV181040 ergaben weder nach 24 noch nach 76 Wochen Hinweise auf eine Effektmodifikation durch das Alter der Patienten bezüglich der gemeinsamen Betrachtung des HbA1c-Wertes und Hypoglykämien.

4.3.1.3.2.6 Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien, stratifiziert nach dem Geschlecht – RCT

Tabelle 4-48: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 6,5% ohne Hypoglykämien erreicht haben aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht

Studie	Zu bewertendes Arzneimittel		Vergleichstherapie		Odds Ratio [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo		
24 Wochen	N/n (%)		N/n (%)		
männlich	5 mg: 108	10 (9,3)	121	4 (3,3)	2,98 [0,91;9,78]
weiblich	5 mg: 143	6 (4,2)	143	1 (0,7)	6,22 [0,74;52,33]
Interaktions- test*	5 mg vs Placebo				I ² =0%; p=0,56
76 Wochen					
männlich	5 mg: 109	5 (4,6)	121	2 (1,7)	2,86 [0,54;15,06]
weiblich	5 mg: 142	1 (0,7)	143	1 (0,7)	1,01 [0,06;16,37]
Interaktions- test*	5 mg vs Placebo				I ² =0%; p=0,53

* eigene Berechnung; KI: Konfidenzintervall

Tabelle 4-49: Ergebnisse für Anzahl der Patienten, die einen HbA1c-Wert von unter 7% ohne Hypoglykämien erreicht haben (exklusive Patienten mit HbA1c < 7% Ausgangswert) aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht

Studie	Zu bewertendes Arzneimittel		Vergleichstherapie		Odds Ratio [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo		
24 Wochen	N/n (%)		N/n (%)		
männlich	5 mg: 107	29 (27,1)	117	9 (7,7)	4,46 [2,00;9,95]
weiblich	5 mg: 140	14 (10,0)	141	5 (3,5)	3,02 [1,06;8,63]
Interaktions- test*	5 mg vs Placebo				I ² =0%; p=0,56
76 Wochen					
männlich	5 mg: 107	12 (11,2)	117	5 (4,3)	2,83 [0,96;8,32]
weiblich	5 mg: 140	6 (4,3)	141	2 (1,4)	3,11 [0,62;15,69]
Interaktions- test*	5 mg vs Placebo				I ² =0%; p=0,92

* eigene Berechnung; KI: Konfidenzintervall

Die Interaktionstests für die Studie CV181040 ergaben keine Hinweise auf eine Effektmodifikation durch das Geschlecht der Patienten bezüglich der gemeinsamen Betrachtung des HbA1c-Wertes und Hypoglykämien.

4.3.1.3.2.7 Gewichtsveränderung stratifiziert nach dem Alter – RCT

Tabelle 4-50: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter

Studie	Zu bewertendes Arzneimittel		Vergleichstherapie			Gruppenunterschied Mittelwert % (SE) [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo			
24 Wochen	N / Mittelwert¹ / SE		N / Mittelwert¹ / SE			
Alter <65 Jahre	5 mg: 211	0,97 0,148	213	0,41 0,146		0,56 (0,208) [0,15;0,96]
Alter ≥65 Jahre	5 mg: 42	0,09 0,332	52	0,08 0,297		0,01 (0,444) [-0,86;0,88]
Interaktions- test²	5 mg vs. Placebo					p=0,3494
76 Wochen						
Alter <65 Jahre	5 mg: 211	1,0 0,166	213	0,43 0,165		0,57 (0,234) [0,11;1,03]
Alter ≥65 Jahre	5 mg: 42	0,84 0,373	52	-0,31 0,334		1,15 (0,500) [0,17;2,13]
Interaktions- test²	5 mg vs. Placebo					p=0,0806

SE: Standardfehler; ¹Änderung des Körpergewichts in kg im Vergleich zum Ausgangswert; ²Interaktionstest basierend auf dem ANCOVA-Modell; KI: Konfidenzintervall

Der Interaktionstest für die Gewichtsveränderung nach 24 Wochen ergab keine Hinweise auf eine Effektmodifikation durch das Alter der Patienten (p=0,3494). Jedoch gab es nach 76 Wochen eine signifikante Interaktion (p=0,0806); jedoch gab es in keiner Subgruppe ein statistisch signifikantes Ergebnis. Daher gibt es auch keine Veränderung an der Interpretation der Gesamtpopulation des Ergebnisses.

4.3.1.3.2.8 Gewichtsveränderung, stratifiziert nach dem Geschlecht – RCT

Tabelle 4-51: Ergebnisse für Gewichtsveränderung aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht

Studie	Zu bewertendes Arzneimittel		Vergleichstherapie			Gruppenunterschied Mittelwert % (SE) [95%-KI]
CV181040	Sulfonylharnstoff+Saxagliptin		Sulfonylharnstoff+Placebo			
24 Wochen	N / Mittelwert ¹ / SE		N / Mittelwert ¹ / SE			
männlich	5 mg: 110	0,76 0,208	122	0,5 0,196		0,26 (0,282) [-0,29;0,82]
weiblich	5 mg: 143	0,87 0,181	143	0,22 0,181		0,65 (0,254) [0,15;1,15]
Interaktions- test²	5 mg vs. Placebo					p=0,5534
76 Wochen						
männlich	5 mg: 110	1,02 0,234	122	0,38 0,221		0,64 (0,317) [0,02;1,27]
weiblich	5 mg: 143	0,94 0,204	143	0,21 0,203		0,73 (0,285) [0,17;1,29]
Interaktions- test²	5 mg vs. Placebo					p=0,4474

SE: Standardfehler, ¹ Änderung des Körpergewichts in kg im Vergleich zum Ausgangswert; ² Interaktionstest basierend auf dem ANCOVA-Modell; KI: Konfidenzintervall

Auf Basis der Studie CV181040 zeigte sich kein Hinweis (p=0,5534 nach 24 Wochen und p=0,4474 nach 76 Wochen) auf eine Modifikation des Behandlungseffektes von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Placebo durch den Faktor Geschlecht in Hinsicht auf die Veränderung des Körpergewichts.

4.3.1.3.2.9 Unerwünschte Ereignisse, stratifiziert nach dem Alter – RCT

Tabelle 4-52: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Alter

Studie	N	Gesamtrate unerwünschter Ereignisse, n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
CV181040				
24 Wochen				
Alter <65 Jahre				
Sulfonylharnstoff+Saxagliptin	211	153 (72,5)	5 (2,4)	4 (1,9)
Sulfonylharnstoff+Placebo	215	160 (74,4)	3 (1,4)	3 (1,4)
OR [95%-KI]		0,91 [0,59;1,39]	1,72 [0,4;7,27]	1,37 [0,3;6,18]
Alter ≥65 Jahre				
Sulfonylharnstoff+Saxagliptin	42	30 (71,4)	1 (2,4)	4 (9,5)
Sulfonylharnstoff+Placebo	52	45 (86,5)	3 (5,8)	1 (1,9)
OR [95%-KI]		0,39 [0,14;1,10]	0,4 [0,04;3,98]	5,37 [49,99]
Interaktionstest*	5 mg vs. Placebo	I ² =54% ; p=0,14	I ² =9,9% ; p=0,29	I ² =0% ; p=0,32
76 Wochen				
Alter <65 Jahre				
Sulfonylharnstoff+Saxagliptin	211	179 (84,8)	9 (4,3)	8 (3,8)
Sulfonylharnstoff+Placebo	215	189 (87,9)	16 (7,4)	10 (4,7)
OR [95%-KI]		0,77 [0,44;1,34]	0,55 [0,24;1,28]	0,81 [0,31;2,09]
Alter ≥65 Jahre				
Sulfonylharnstoff+Saxagliptin	42	36 (85,7)	3 (7,1)	8 (19,0)
Sulfonylharnstoff+Placebo	52	48 (92,3)	7 (13,5)	2 (3,8)
OR [95%-KI]		0,50 [0,13;1,90]	0,49 [0,12;2,04]	5,88 [1,18;29,42]
Interaktionstest*	5 mg vs. Placebo	I ² =0% ; p=0,56	I ² =0% ; p=0,89	I ² =76,9% ; p=0,04

* eigene Berechnung; KI: Konfidenzintervall

Die Interaktionstests für die Studie CV181040 zeigten zum Niveau 0,2 statistisch signifikante Interaktionen (p=0,14; I²=54%) für den Faktor Alter nach 24 Wochen hinsichtlich dem Auftreten von unerwünschten Ereignissen, jedoch sind diese Effekte gleichgerichtet (kein Unterschied zwischen den Behandlungsoptionen bei der älteren Subpopulation und kein Unterschied bei der jüngeren Subpopulation). Nach 76 Wochen war keine Interaktion nachweisbar (p=0,56 I²=0%). Interaktionen bezüglich schwerwiegender unerwünschter Ereignisse zeigten sich nicht. In Bezug auf die Studienabbrecher aufgrund unerwünschter Ereignisse gab es nach 76 Wochen eine signifikante Interaktion zwischen den Altersgruppen. Bei der Gesamtrate unerwünschter Ereignisse sowie der Rate der schwerwiegenden unerwünschten Ereignisse gab es zu keinem Zeitpunkt in den beiden Altersgruppen signifikante Unterschiede zwischen den Behandlungsgruppen.

4.3.1.3.2.10 Unerwünschte Ereignisse, stratifiziert nach dem Geschlecht – RCT

Tabelle 4-53: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach Geschlecht

Studie	N	Gesamtrate unerwünschter Ereignisse, n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
CV181040				
24 Wochen				
männlich				
Sulfonylharnstoff+Saxagliptin	110	78 (70,9)	3 (2,7)	2 (1,8)
Sulfonylharnstoff+Placebo	123	91 (74,0)	4 (3,3)	2 (1,6)
OR [95%-KI]		0,86 [0,48;1,52]	0,83 [0,18;3,81]	1,12 [0,16;8,09]
weiblich				
Sulfonylharnstoff+Saxagliptin	143	105 (73,4)	3 (2,1)	6 (4,2)
Sulfonylharnstoff+Placebo	144	114 (79,2)	2 (1,4)	2 (1,4)
OR [95%-KI]		0,73 [0,42;1,26]	1,52 [0,25;9,24]	3,11 [0,62;15,67]
Interaktionstest*	5 mg vs. Placebo	I ² =0% ; p=0,68	I ² =0% ; p=0,62	I ² =0% ; p=0,43
76 Wochen				
männlich				
Sulfonylharnstoff+Saxagliptin	110	92 (83,6)	7 (6,4)	4 (3,6)
Sulfonylharnstoff+Placebo	123	103 (83,7)	13 (10,6)	6 (4,9)
OR [95%-KI]		0,99 [0,49;1,99]	0,58 [0,22;1,5]	0,74 [0,2;2,68]
weiblich				
Sulfonylharnstoff+Saxagliptin	143	123 (86,0)	5 (3,5)	12 (8,4)
Sulfonylharnstoff+Placebo	144	134 (93,1)	10 (6,9)	6 (4,2)
OR [95%-KI]		0,46 [0,21;1,02]	0,49 [0,16;1,46]	2,11 [0,77;5,78]
Interaktionstest*	5 mg vs. Placebo	I ² =51% ; p=0,18	I ² =0% ; p=0,82	I ² =36,8% ; p=0,21

* eigene Berechnung; KI: Konfidenzintervall

Es ergab sich ein Hinweis auf eine Effektmodifikation durch das Geschlecht in Bezug auf die Gesamtrate der unerwünschten Ereignisse. Jedoch waren die Effekte in beiden Subpopulationen gleichgerichtet, so dass diese Interaktion keinen Einfluss auf das Ergebnis der Gesamtpopulation hat. Weitere Interaktionen in Bezug auf die schwerwiegenden unerwünschten Ereignisse oder auf Studienabbrüche aufgrund unerwünschter Ereignisse fanden sich nicht.

4.3.1.3.2.11 Kardiale Ereignisse stratifiziert nach dem Alter – RCT

Tabelle 4-54: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach dem Alter

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
CV181040		
Nach 24 Wochen		
Alter <65 Jahre		
Sulfonylharnstoff +Saxagliptin	211	9 (4,3)
Sulfonylharnstoff+Placebo	215	4 (1,9)
OR [95%-KI]		2,33 [0,71;7,68]
Alter ≥65 Jahre		
Sulfonylharnstoff +Saxagliptin	42	1 (2,4)
Sulfonylharnstoff+Placebo	52	6 (11,5)
OR [95%-KI]		0,19 [0,02;1,62]
Interaktionstest*	5 mg vs. Placebo	I ² =75,1%; p=0,05
Nach 76 Wochen		
Alter <65 Jahre		
Sulfonylharnstoff +Saxagliptin	211	12 (5,7)
Sulfonylharnstoff+Placebo	215	8 (3,7)
OR [95%-KI]		1,55 [0,62;3,86]
Alter ≥65 Jahre		
Sulfonylharnstoff +Saxagliptin	42	3 (7,1)
Sulfonylharnstoff+Placebo	52	7 (13,5)
OR [95%-KI]		0,49 [0,12;2,04]
Interaktionstest*	5 mg vs. Placebo	I ² =42,8%; p=0,16

* eigene Berechnung; KI: Konfidenzintervall

Sowohl nach 24 als auch nach 76 Wochen ergab sich ein Hinweis auf eine Effektmodifikation durch das Alter der Patienten. Aufgrund der sehr niedrigen Ereigniszahlen ergibt sich aus der Interaktion kein Einfluss auf die Gesamtinterpretation der Population.

4.3.1.3.2.12 Kardiale Ereignisse stratifiziert nach dem Geschlecht – RCT

Tabelle 4-55: Ergebnisse für Kardiale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel stratifiziert nach dem Geschlecht

Studie	N	Gesamtrate kardiale Ereignisse, n (%)
CV181040		
24 Wochen		
männlich		
Sulfonylharnstoff+Saxagliptin	110	6 (5,5)
Sulfonylharnstoff+Placebo	123	4 (3,3)
OR [95%-KI]		1,72 [0,47;6,25]
weiblich		
Sulfonylharnstoff+Saxagliptin	143	4 (2,8)
Sulfonylharnstoff+Placebo	144	6 (4,2)
OR [95%-KI]		0,66 [0,18;2,40]
Interaktionstest*	5 mg vs. Placebo	I ² =4,6%; p=0,31
76 Wochen		
männlich		
Sulfonylharnstoff+Saxagliptin	110	9 (8,2)
Sulfonylharnstoff+Placebo	123	7 (5,7)
OR [95%-KI]		1,48 [0,53;4,11]
weiblich		
Sulfonylharnstoff+Saxagliptin	143	6 (4,2)
Sulfonylharnstoff+Placebo	144	8 (5,6)
OR [95%-KI]		0,74 [0,25;2,20]
Interaktionstest*	5 mg vs. Placebo	I ² =0%; p=0,37

* eigene Berechnung; KI: Konfidenzintervall

Es ergab sich kein Hinweis auf eine Effektmodifikation durch das Geschlecht in Bezug auf die Gesamtrate der kardialen Ereignisse.

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

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

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

Die in die Nutzenbewertung eingeschlossene placebokontrollierte RCT CV181040 war eine Zulassungsstudie der Phase III für Saxagliptin. Im Folgenden sollen die Ergebnisse für die 5 mg-Gruppe nochmal zusammenfassend betrachtet werden:

Bezüglich des **Ausmaßes der Blutzuckerkontrolle** (HbA1c) ergaben sich sowohl nach 24 als auch nach 76 Wochen (LOCF) statistisch signifikante Unterschiede zwischen Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+Placebo (MWD [95%-KI] zugunsten von Sulfonylharnstoff+Saxagliptin nach 24 Wochen: -0,72% [-0,88;-0,56] und nach 76 Wochen -0,56% [-0,72;-0,39]). Die Analyse mit dem Repeated-Measures Modell bestätigt die Robustheit der Ergebnisse. Die unterstützende Wirksamkeitsmessung (Anteil von Patienten, die die Studie aufgrund von unzureichender Blutzuckerkontrolle abgebrochen haben oder eine Notfallmedikation, die aufgrund unzureichender Blutzuckerkontrolle verabreicht wurde, bekommen haben) ergab in der Sulfonylharnstoff+Saxagliptin-Gruppen gegenüber der Sulfonylharnstoff+Placebo-Gruppe nach 24 und 76 Wochen signifikant geringere Raten (OR [95%-KI] nach 24 Wochen: 0,47 [0,31;0,72] und nach 76 Wochen 0,55 [0,38;0,80]).

Für den Endpunkt „Ausmaß der Blutzuckerkontrolle“ ergaben sich hinsichtlich der in Subgruppenanalysen untersuchten Faktoren Alter und Geschlecht keine Hinweise auf Effektmodifikation, alle Interaktionstests zeigten entweder p-Werte $>0,2$ oder konnten aufgrund der Konstellation des Befundes (unter Berücksichtigung von Konsistenz; Robustheit der beobachteten Interaktion; Plausibilität; zugrunde liegende Gruppengröße) als nicht fazitrelevant identifiziert werden.

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

Bezüglich der Anzahl von Patienten mit mindestens einer Hypoglykämie (patientenberichtet und durch Blutzuckerwert ≤ 50 mg/dL bestätigt) gab es weder nach 24 Wochen (patientenberichtet: OR [95%-KI]: 1,52 [0,90;2,58], bestätigt: 1,06 [0,15;7,55]) noch nach 76 Wochen (patientenberichtet: OR [95%-KI]: 1,15 [0,76;1,74], bestätigt: 0,94 [0,36;2,47]) statistisch signifikante Unterschiede zwischen den Gruppen.

Hinsichtlich der in Subgruppenanalysen untersuchten Faktoren Alter und Geschlecht ergaben sich keine Hinweise auf Effektmodifikation, alle Interaktionstests zeigten p-Werte $>0,2$.

Beim kombinierten Endpunkt „Erreichen des HbA1c-Zielwerts ($<6,5\%/<7\%$) ohne Auftreten von Hypoglykämien“ ergaben sich folgende Ergebnisse: nach 24 Wochen hatten signifikant mehr Patienten in der Sulfonylharnstoff+Saxagliptin-Gruppe einen HbA1c-Wert unter 6,5%, ohne dass eine Hypoglykämie aufgetreten ist (OR [95%-KI]: 3,54 [1,28;9,82]). Nach 76 Wochen war dieser Unterschied nicht mehr signifikant (OR [95%-KI]: 2,13 [0,53;8,61]). Wird glykämische Kontrolle als HbA1c $<7\%$ definiert, gab es zu beiden Zeitpunkten einen signifikanten Unterschied zugunsten von Sulfonylharnstoff+Saxagliptin (OR [95%-KI]: 3,54 [1,28;9,80] / 2,82 [1,16;6,87]).

Bezüglich der **Gewichtsveränderung** wurde beobachtet, dass die Patienten in beiden Gruppen an Gewicht zunahmen. Die Gewichtszunahme war in beiden Gruppen gering, in der Sulfonylharnstoff+Saxagliptin-Gruppe sowohl nach 24 als auch nach 76 Wochen aber etwas höher (0,8 kg und 1,0 kg) als in der Sulfonylharnstoff+Placebo-Gruppe (0,3 kg und 0,3 kg). Dieser Unterschied war für beide Zeitpunkte (LOCF) signifikant (MWD [95%-KI] nach 24 Wochen: 0,5 kg [0,1;0,8] und nach 76 Wochen: 0,70 kg [0,28;1,12]). Die Analyse mit den Observed-Cases zeigte hingegen lediglich nach 76 Wochen einen signifikanten Unterschied (MWD [95%-KI]: 1,80 kg [0,73;3,23]).

Für die Gewichtveränderung ergab sich kein Hinweis auf eine Effektmodifikation durch das Geschlecht. Bezüglich des Alters gab es jedoch nach 76 Wochen eine signifikante Interaktion ($p=0,0806$). Die Interaktion hat jedoch keinen Einfluß auf die Interpretation der Gesamtpopulation. Daraus ergibt sich keine Änderung des aus der Bewertung der Gesamtpopulation abgeleiteten Fazits.

Zusammenfassend zeigte sich im Vergleich Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+Placebo keine statistisch signifikant erhöhte **Gesamtrate unerwünschter Ereignisse** (OR [95%-KI]: 0,71 [0,47;1,06] nach 24 Wochen und 0,72 [0,43;1,20] nach 76 Wochen); keine signifikant erhöhte **Gesamtrate der schwerwiegenden unerwünschten Ereignisse** (OR [95%-KI]: 0,90 [0,30;2,72] nach 24 Wochen und 0,53 [0,26;1,09] nach 76 Wochen) und keine statistisch signifikant erhöhte **Abbruchrate aufgrund unerwünschter Ereignisse** (OR [95%-KI]: 1,06 [0,26;4,27] nach 24 Wochen und 1,43 [0,66;3,10] nach 76 Wochen).

Hinweise auf Effektmodifikationen durch das Alter zeigten sich für unerwünschte Ereignisse nach 24 Wochen ($p=0,14$) und für Studienabbrüche aufgrund unerwünschter Ereignisse nach 76 Wochen ($p=0,04$). Interaktionen bezüglich schwerwiegender unerwünschter Ereignisse zeigten sich nicht. Für das Geschlecht ergab sich lediglich für die Gesamtrate der unerwünschten Ereignisse nach 76 Wochen eine Effektmodifikation ($p=0,18$). Keine dieser Effektmodifikationen hat eine Auswirkung auf das abgeleitete Fazit (Begründung siehe Abschnitt 4.3.1.3.2.9 und 4.3.1.3.2.10).

Beim Auftreten von **kardialen Ereignissen** ergaben sich keine signifikanten Unterschiede zwischen den beiden Gruppen (OR [95%-KI] nach 24 Wochen: 1,06 [0,43;2,59] und nach 76 Wochen: 1,06 [0,51;2,21]).

Hinweise auf Effektmodifikationen durch das Alter zeigten sich für kardiale Ereignisse sowohl nach 24 Wochen ($p=0,05$) als auch nach 76 Wochen ($p=0,16$). Die Anzahl der Ereignisse ist sehr klein und der Effekt ist in jeder Subpopulation gleichgerichtet, so dass die Interaktion keinen Einfluss auf die Gesamtinterpretation der Population hatte. Bezüglich des Geschlechts zeigten sich keine Effektmodifikationen.

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

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

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

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

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

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 4-56: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
CV181040	ja	abgeschlossen	Behandlungsphase: 24 Wochen Verlängerungsphase: 12 Monate	Glibenclamid (7,5 mg/Tag) +Saxagliptin (2,5 und 5 mg/Tag) Glibenclamid (hochtitriert auf 10 mg/Tag) +Placebo
D1680C00007	ja	abgeschlossen	Behandlungsphase: 12 Wochen Verlängerungsphase: 40 Wochen	Vorbehandlung +Saxagliptin (2,5 mg/Tag) Vorbehandlung +Placebo
D1680C00003 (SAVOR-TIMI 53)	nein	laufend	4 Jahre	Vorbehandlung +Saxagliptin (2,5 und 5 mg/Tag) Vorbehandlung +Placebo
CV181142	nein	laufend	6 Wochen	Vorbehandlung +Saxagliptin (5 mg/Tag) Vorbehandlung +Placebo

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 ist zum 21.1.2013 abgebildet.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-56 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-57: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
D1680C00007	Die mit Sulfonylharnstoffen vorbehandelte Teilpopulation enthält zu wenige Patienten.
D1680C00003 (SAVOR-TIMI 53)	Die Studie läuft noch. Es liegen keine Ergebnisse vor*.
CV181142	Die Studie läuft noch. Es liegen keine Ergebnisse vor. Außerdem beträgt die Studiendauer nur sechs Wochen.

* Ergebnisse werden für 2013 erwartet

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 mit der zweckmäßigen Vergleichstherapie

Die umfassende bibliographische Literaturrecherche vom 07.02.2013 nach randomisierten kontrollierten Studien mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Insulin oder Insulin-Monotherapie erzielte insgesamt 4295 Treffer. Nach Ausschluss der Duplikate (n=1380) wurden die verbleibenden 2915 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den in Abschnitt 4.2.2 definierten Kriterien gesichtet. Die Publikationen wurden von zwei Personen unabhängig voneinander nach Relevanz selektiert.

Basierend auf den Informationen des Titels und des Abstracts wurden 2833 Publikationen begründet durch Ausschlusskriterien als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Für die verbleibenden 82 Publikationen, deren Titel und Abstract die Einschlusskriterien erfüllen, wurden im Anschluss die Volltexte hinsichtlich dieser Kriterien überprüft. Im Ergebnis entsprachen 77 nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C, Tabelle 4-113). Ausschlussgründe: falsche Intervention; falsche Vergleichstherapie; keine RCT; zu geringe Studiendauer; keine patientenrelevanten Endpunkte; falsche Patientenpopulation; Doppelpublikation; Zeitpunkt stimmt nicht überein, weshalb kein indirekter Vergleich möglich war.

Somit ergab die bibliografische Literaturrecherche fünf relevante Publikationen. Bei den identifizierten Studien handelt es sich um die Publikationen von (Nathan et al., 1988; Birkeland et al., 1994; Shank et al., 1995; Tovi et al., 1998; Turner et al., 1998).

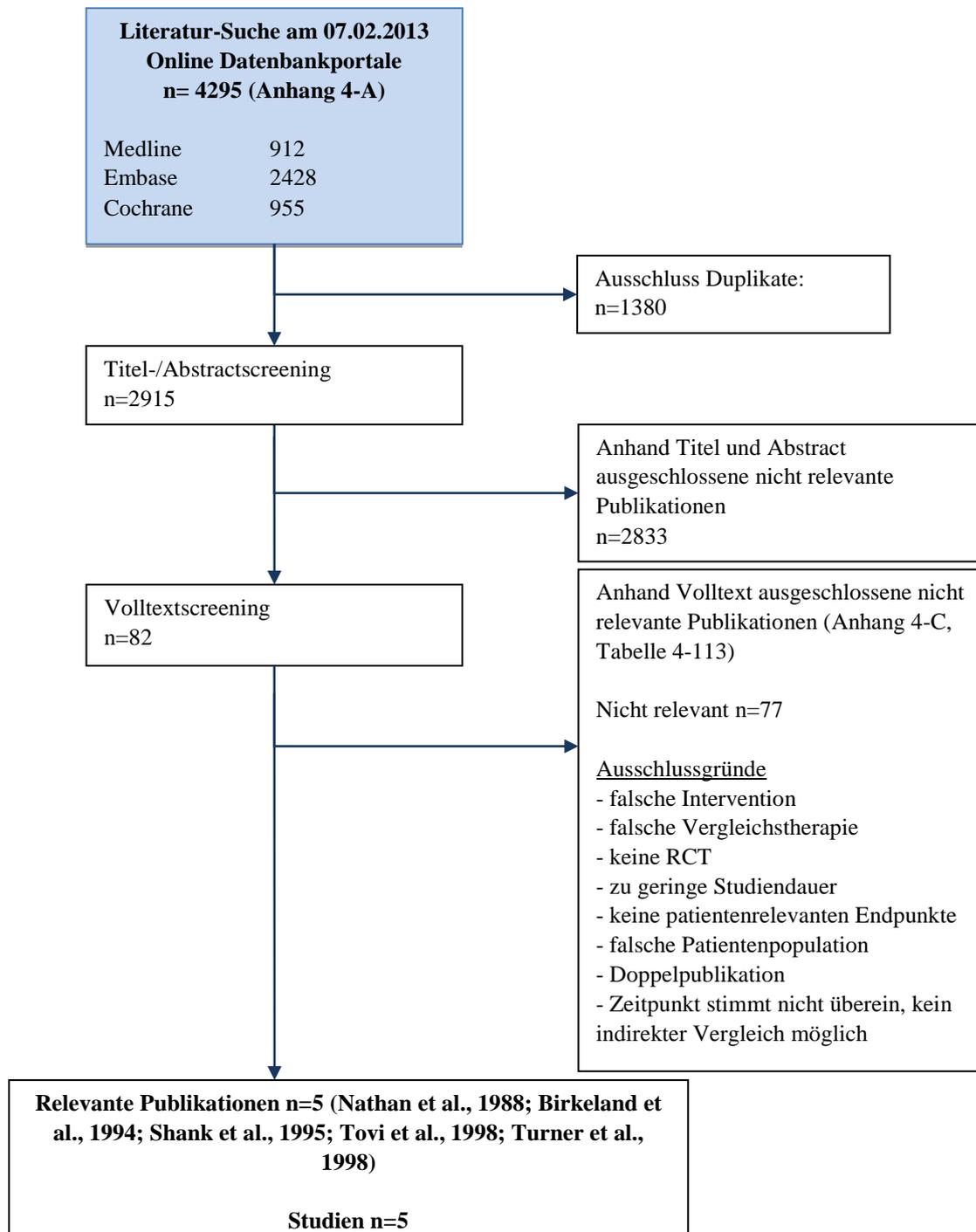


Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit der zweckmäßigen Vergleichstherapie

Indirekter Vergleich mit der alternativen Vergleichstherapie

Die umfassende bibliographische Literaturrecherche vom 04.02.2013 nach randomisierten, kontrollierten Studien mit der alternativen Vergleichstherapie (Sitagliptin, Vildagliptin und Linagliptin) in Add-on Kombination mit Metformin erzielte insgesamt 1160 Treffer. Nach Ausschluss der Duplikate (n=164) wurden die verbleibenden 996 Publikationen auf Grundlage des Titels und – sofern verfügbar – Abstracts nach den in Abschnitt 4.2.2 definierten Kriterien gesichtet. Die Studien wurden von zwei Personen unabhängig voneinander nach Relevanz selektiert. Basierend auf den Informationen des Titels und des Abstracts wurden 991 Publikationen begründet durch Ausschlusskriterien als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Für die verbleibenden fünf Publikationen, deren Titel und Abstract die Einschlusskriterien erfüllen, wurden im Anschluss die Volltexte hinsichtlich dieser Kriterien überprüft. Im Ergebnis entsprachen zwei nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C, Tabelle 4-114). Ausschlussgründe: falscher Kontrollarm, zu geringe Behandlungsdauer).

Somit ergab die bibliografische Literaturrecherche drei relevante Publikationen. Bei den identifizierten Studien handelt es sich um die Publikationen von (Hermansen et al., 2007; Garber et al., 2008; Lewin et al., 2012).

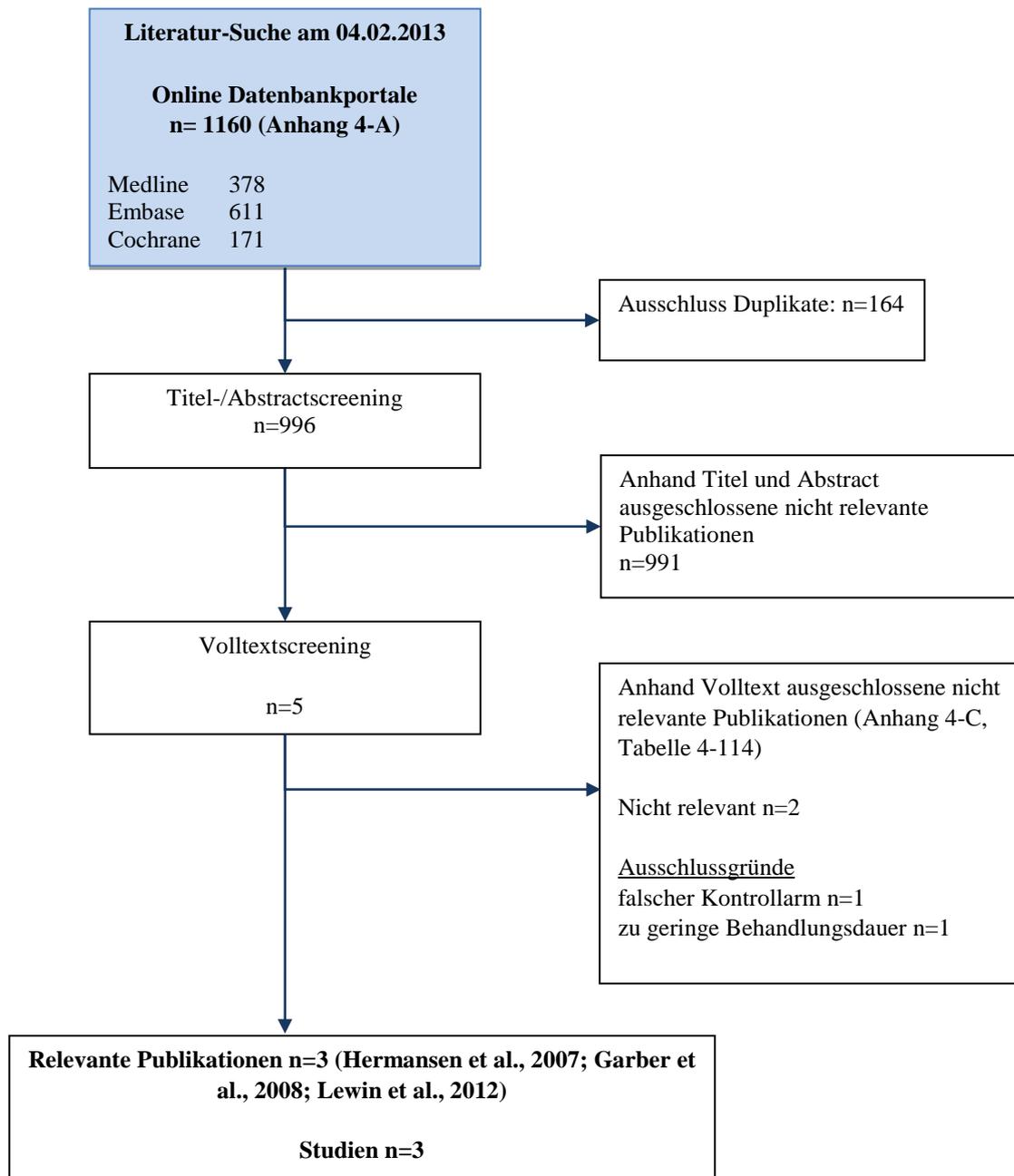


Abbildung 3: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit der alternativen Vergleichstherapie

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 4-56) 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 mit dem zu bewertenden Arzneimittel sind in Tabelle 4-7 aufgeführt.

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

Die für den indirekten Vergleich verwendeten Studien mit der alternativen Vergleichstherapie werden in Tabelle 4-58 genannt.

Tabelle 4-58: 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)
Garber 2008 (CLAF237A2305)	clinicaltrials.gov [NCT00099944] (ClinicalTrials.gov, 2013a) WHO ICTRP [NCT00099944] (ICTRP Search Portal, 2013a)	nein	ja
Lewin 2012	clinicaltrials.gov [NCT00819091] (ClinicalTrials.gov, 2013e)	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 mit dem zu bewertenden Arzneimittel

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 Vergleichs-

therapie. 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 4-59: Studienpool – RCT für indirekte Vergleiche

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, zu bewertendes Arzneimittel						
CV181040	ja	ja	nein	ja (Bristol-Myers Squibb, 2008)	ja [NCT00313313] (ClinicalTrials.gov, 2013b)	ja (Chacra et al., 2009; Chacra et al., 2011)
placebokontrolliert, zweckmäßige Vergleichstherapie(n)						
Birkeland 1994	nein	nein	ja	nein	nein	ja (Birkeland et al., 1994)
Nathan 1988	nein	nein	ja	nein	nein	ja (Nathan et al., 1988)
Shank 1995	nein	nein	ja	nein	nein	ja (Shank et al., 1995)
Tovi 1998	nein	nein	ja	nein	nein	ja (Tovi et al., 1998)
Turner 1998	nein	nein	ja	nein	nein	ja (Turner et al., 1998)
placebokontrolliert, alternative Vergleichstherapie(n)						
Garber 2008 (CLAF237 A2305)	nein	nein	ja	nein	[NCT00099944] (ClinicalTrials.gov, 2013a) [NCT00099944] (ICTRP Search Portal, 2013a)	ja (Garber et al., 2008)
Hermansen 2007	nein	nein	ja	nein	nein	ja (Hermansen et al., 2007)
Lewin 2012	nein	nein	ja	nein	[NCT00819091] (ClinicalTrials.gov, 2013e)	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-F zu hinterlegen.

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

Tabelle 4-60: Charakterisierung der eingeschlossenen Studien – RCT mit der zweckmäßigen bzw. alternativen Vergleichstherapie

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
Studie CV181040 siehe Abschnitt 4.3.1.2						
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin						
Shank 1995	RCT, doppelblind, placebokontrolliert, mit offener Phase	Nicht-insulinabhängiger Typ-2-Diabetes Patienten im Alter von 33 bis 70 Jahre	3 Gruppen: 1) nur Glipizid (n=10) 2) NPH Insulin beim Zubettgehen und Placebo (n=10) 3) NPH Insulin beim Zubettgehen und Glipizid (n=10)	Phase 1: Alle Patienten erhielten Glipizid (offen), 2 Monate Phase 2: Randomisierung zu eine der drei Behandlungsgruppen: -Glipizid -Insulin (+Placebo) -Insulin+Glipizid (Insulin-Gabe wurde nicht verblindet) 3 Monate Phase 3: bei den Patientengruppen mit Insulin-Gabe wurde die Insulin-Dosis hochtitriert 3 Monate Phase 4: alle drei	Keine Angabe	Primärer Endpunkt und weitere patientenrelevante Endpunkte: Gewichtszunahme, Blutdruck und Albumin- und Proteinausscheidung im Urin Bestimmung von Glukose, Insulin und der Konzentration der freien Fettsäuren (FFA) aus Plasma HbA1c Nüchtern Lipid-Profil im Serum (Gesamt-Cholesterin, HDL-Cholesterin, LDL-Cholesterin, Triglyceride) Während euglykämischen Insulin Clamp (am Ende aller 3 Phasen): - Gesamt-Glukoseaufnahme - Glukose-Oxidation - nicht-oxidative Glukoseabgabe - HPG-Abfall während der Insulininfusion Basalwerte und Werte während euglykämischen Insulin Clamp (am Ende aller 3 Phasen):

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

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	<RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	<relevante Charakteristika, z. B. Schweregrad>		Behandlungsgruppen wechselten ihre Medikation zu Insulin+Glipizid 6 Monate		- Plasma FFA Konzentration (Plasma und während euglykämischen Insulin Clamp) - Lipidoxidation (Plasma und während euglykämischen Insulin Clamp)
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie						
Birkeland 1994	RCT, prospektiv, parallel, aktiv-kontrolliert, Langzeit, offen, monozentrisch	Patienten mit Diabetes mellitus Typ 2 und Einhaltung einer Diät und/oder zusätzlicher Glibenclamid-Einnahme	Insulin (8 IU um 8 und 22 Uhr, Hochtitration je nach Bedarf möglich) (n=18) vs. Glibenclamid (durchschnittlich max. 10,5 mg/Tag) (n=16)	Run in: 3 Monate Behandlung: 12 Monate	Diabeteszentrum Aker, Norwegen, Zeitraum nicht berichtet	Primärer Endpunkt und weitere patientenrelevante Endpunkte: Langzeiteffekt von Insulin gegen Sulfonylharnstoff bei Diabetes mellitus Typ-2 Patienten Blutzuckerkontrolle, Herz-Kreislauf Faktoren, Insulinresistenz, Mikroalbuminurie, Lp(a) Lipoprotein, Triglyceride (TG), gesamt- und HDL-Cholesterin, diastolischer und systolischer Blutdruck
Tovi 1998	RCT, monozentrisch, parallel, aktivkontrolliert, offen	Ältere Patienten mit Diabetes mellitus Typ 2 mit zweimaliger erfolgloser oraler Antidiabetikatherapie, ohne Dekompensation, schlechter	Insulin 0,53±0,27 Einheiten/kg (n=19) vs. Sulfonylharnstoff (Glibenclamid 7-10 mg oder Glipizid 10-15 mg) (n=16)	12 Monate	Schweden, Zeitraum nicht berichtet	Primärer Endpunkt und weitere patientenrelevante Endpunkte: Effekt der verbesserten Stoffwechselkontrolle auf das Wohlbefinden und die Symptome von älteren Patienten mit Typ-2 Diabetes mellitus und

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
		Blutzuckerkontrolle trotz hoher Sulfonylharnstoffdosis				Insulintherapie Entdecken von Symptomen/Anzeichen, die das Unwohlsein durch hohen Bluthochdruck und andere nicht altersbedingte Symptome verursachen und für den Beginn einer notwendigen Insulintherapie sprechen (Vergleich mit einer gesunden Kontrollgruppe) Veränderung des HbA1c-Wertes über HPLC Technik, Blutzuckerwert, Gewicht und BMI Mikroalbuminurie Blutdruck, Achilles Reflex, Fußgeschwüre, Retinopathie Messungen jeweils zu Beginn, nach 6 und 12 Monaten Fragebogen zur Lebensqualität („Goteburg quality-of-life instrument“) alle 3 Monate (Vergleich mit einer gesunden Kontrollgruppe)
Nathan 1999	RCT, monozentrisch, doppel-blind, doppel-dummy, placebo-kontrolliert,	Patienten mit Typ-2-Diabetes, nicht Insulinabhängig und bisher keine gute Blutzuckerkontrolle	Insulin (15-30 Einheiten/Tag) (n=15) vs (Glibenclamid 2,5-10 mg/zweimal	Run-in: 1 Monat Behandlung: 9 Monate	Diabetesklinik Massachusetts, Zeitraum nicht berichtet	Primärer Endpunkt und weitere patientenrelevante Endpunkte: Effekt, Risiko und Vorteile einer Insulin- gegen Glibenclamid-Therapie bei Patienten mit

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

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	parallel, zweiarmig	durch Diät allein erreicht	täglich) (n=16) Anpassung der Medikation war möglich			Diabetes mellitus Typ 2 Nüchtern Blutzuckerwert <6,4 mmol/L ohne Hypoglykämie, gut kontrolliert Unterschied erfolgloser Therapie mit Insulin oder oraler Antidiabetika Messungen jeden Monat (bis 9 Monate)
Turner 1998 (UKPDS 24)	RCT, prospektiv, multizentrisch	Patienten im Alter von 25-65 Jahre mit neu diagnostiziertem Typ-2-Diabetes mellitus	Primary diet failure Group: Insulin (n=178) Sulfonylharnstoff (Chlorpropamid oder Glibenclamid) (n=231) Metformin (n=49)	Run-in: 3 Monate Behandlung: 6 Jahre	15 Studienzentren in Großbritannien, 1977-1991	Primärer Endpunkt und weitere patientenrelevante Endpunkte: Ansprechen nach einem Jahr und nach 6 Jahren: mediane Nüchternplasmaglukose, mediane HbA1c-Konzentration, Körpergewicht (Mittelwert) Veränderung des Körpergewichts Plasmainsulinspiegel über 6 Jahre hinweg, β-Zellfunktion und Insulinsensitivität nach einem Jahr Jährliche Rate der hypoglykämischen Episoden
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	RCT, doppelblind, parallel, aktivkontrolliert, multizentrisch,	Patienten im Alter von 18 bis 75 mit Diabetes mellitus Typ 2, deren	Stratum 1: Glimепirid+Placebo (n=106)	Umstellungsphase (wenn erforderlich): bis zu 14 Wochen	Weltweit; 2005-2007	Primärer Endpunkt: HbA1c-Änderung in Woche 24 im Vergleich zum Ausgangswert

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

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	international	Blutzuckerwert mit Glimepirid oder Glimepirid+ Metformin nicht ausreichend kontrolliert werden kann (HbA1c $\geq 7,5\%$ und $\leq 10,5\%$)	Glimepirid+ Sitagliptin (n=106) Stratum 2: Glimepirid+ Metformin+Placebo (n=113) Glimepirid+ Metformin+ Sitagliptin (n=116)	Placebo-Run-in-Phase: 2 Wochen Behandlungsphase: 24 Wochen		Weitere patientenrelevante 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+ Vildagliptin 50 mg/Tag (n=132) Glimepirid+Vildagliptin 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 Weitere patientenrelevante Endpunkte: Gewichtveränderung in Woche 24 im Vergleich zum Ausgangswert Anzahl bestätigter Hypoglykämien Unerwünschte Ereignisse
Lewin 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%	Sulfonylharnstoff+ Linagliptin (n=161) Sulfonylharnstoff+ Placebo (n=84)	Wash-out Phase: 4 Wochen (Absetzen anderer oraler Antidiabetika) Placebo-Run-in-Phase: 2 Wochen Behandlungsphase: 18 Wochen	Weltweit; 2008-2010	Primärer Endpunkt: HbA1c-Änderung in Woche 18 im Vergleich zum Ausgangswert Weitere patientenrelevante Endpunkte: Anzahl Hypoglykämien Unerwünschte Ereignisse

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
		bis 10,0%)				

Tabelle 4-61: Charakterisierung der Interventionen – RCT mit der zweckmäßigen bzw. alternativen Vergleichstherapie

Studie	Insulin	SU	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run- in-Phase etc.</i>
Studie CV101840 siehe Abschnitt 4.3.1.2			
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin			
Shank 1995	(Glipizid+) Insulin	Glipizid	Phase 1: Alle Patienten erhielten Glipizid (offen) Phase 2: Randomisierung zu einer der drei Behandlungsgruppen: Glipizid, Insulin (+Placebo), Insulin+Glipizid Phase 3: bei den Patientengruppen mit Insulin- Gabe wurde die Insulin-Dosis hochtitriert Phase 4: alle drei Behandlungsgruppen wechselten ihre Medikation zu Insulin+Glipizid
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie			
Birkeland 1994	Insulin	Glibenclamid	Run-in 3 Monate: nur Diät oder zusätzlich Glibenclamid-Einnahme Insulin (durchschnittlich 0,51 Einheiten b.i.d., 0,06-0,91 Einheiten) um 8 und 22 Uhr vs. Glibenclamid (durchschnittlich 7,88 mg anfangs und 8,88 mg am Ende, 1,75-10,5 mg) Anpassung während der Studie war möglich 2 Patienten erhielten multiples Dosierungsschema (normales Insulin 30 min vor dem Essen und intermediate-acting Insulin um 22 Uhr)
Tovi 1998	Insulin	Sulfonylharnstoff (Glibenclamid oder Glipizid)	Insulin 0,53±0,27 Einheiten/kg vs. Sulfonylharnstoff (Glibenclamid 7-10 mg oder Glipizide 10-15 mg) Anpassung bis Blutzuckerwert bei 6-12 mmol/L
Nathan 1999	Insulin	Placebo (Glibenclamid)	Run-in 1 Monat: Diät Nur falls Nüchternblutzuckerspiegel <7,8 mmol/L darf Studie fortgeführt werden Behandlung: Insulin (15-30 Einheiten/Tag) vs. Placebo (Glibenclamid 2,5-10 Einheiten/Tag zweimal täglich) Anpassung der Medikation war möglich Double-dummy-Medikation
Turner 1998 (UKPDS 24)	Insulin	Chlorpropamid oder Glibenclamid	Run-in 3 Monate: nur Diät, Untersuchungen einmal monatlich
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor			
Hermansen 2007*	Glimepirid 4-8 mg/Tag + Sitagliptin	Glimepirid 4-8 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;

Studie	Insulin	SU	<i>ggf. weitere Spalten mit Behandlungscharakteristika</i> z. B. Vorbehandlung, Behandlung in der Run- in-Phase etc.
	100 mg/Tag		<p>Stabilisierungsphase von bis zu 10 Wochen) Danach erfolgte eine 2-wöchige Placebo-Run-in-Phase.</p> <p>Behandlung (Stratum 1): 24 Wochen entweder Glimepirid oder Glimepirid+Sitagliptin.</p> <p>Eine Dosisveränderung der Hintergrundtherapie (Glimepirid) war nicht vorgesehen; Rescue-Therapie mit Pioglitazon war unter vordefinierten Bedingungen möglich.</p>
Garber 2008*	Glimepirid 4 mg/Tag + Vildagliptin 50 mg/Tag	Glimepirid 4 mg/Tag + Placebo	<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 100 mg, oder Placebo)</p> <p>Behandlung: 24 Wochen Vildagliptin 50 mg oder 100 mg 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>
Lewin 2012*	Sulfonylharnstoff in stabiler Tagesdosis + Linagliptin 5 mg/Tag	Sulfonylharnstoff in stabiler Tagesdosis + Placebo	<p>Vorbehandlung: In einer wash-out Phase von 4 Wochen wurde der Patient nur mit Sulfonylharnstoffen behandelt, danach für 2 Wochen zusätzlich mit Placebo (Run-in 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 Rescue-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.

Tabelle 4-62: Charakterisierung der Studienpopulationen – RCT mit der zweckmäßigen bzw. alternativen Vergleichstherapie

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD)	Geschlecht w/m (%)	BMI (kg/m ²) Mittelwert (SD)	Gewicht (kg) Mittelwert (SD)
Studie CV181040 siehe Abschnitt 4.3.1.2					
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin					
Shank 1995					
Glipizid+Insulin	10	Gemeinsam:	Nicht	Gemeinsam:	87,3 (5,5)
Insulin	10	53 (2)	berichtet	30,5 (0,6)	74,9 (2,6)
Glipizid	10				79,2 (3,6)
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie					
Birkeland 1994					
Insulin	18	Gemeinsam:	55,6 / 44,5	26,4 (2,97)	75,3 (13,1)
Glibenclamid	16	59,2 (6,0)	18,8 / 81,3	26,3 (4,0)	77,1 (14,4)
Tovi 1998					
Insulin	19	73,4 (5,0)	57,9 / 42,1	26,3 (4,6)	72,9 (14,1)
Sulfonylharnstoff (Glibenclamid oder Glipizid)	16	77,3 (6,2)	56,3 / 43,8	28,3 (4,3)	78,3 (15,3)
Nathan 1999					
Insulin	15	53,5 (7,1)	46,7 / 53,3	28,6 (5,1)	80,6 (14)
Glibenclamid	16	50,3 (5,9)	43,8 / 56,3	30,3 (5,7)	89,3 (15)
Turner 1998 (UKPDS 24)	458	50 (10)	47/53	26,1 (6,0)	72 (16)
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin 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
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
Lewin2012					
Sulfonylharnstoff+ Linagliptin 5 mg	161	57,2 (9,8) ^b	52,2 / 47,8	28,40 (5,02) ^b	74,5 (17,0) ^b
Sulfonylharnstoff+Placebo	84	56,2 (10,2) ^b	38,1 / 61,9	28,21 (5,12) ^b	76,1 (17,0) ^b

^aStandardfehler (SE); ^bStandardabweichung (SD)

Tabelle 4-63: Charakterisierung der Studienpopulationen – RCT mit der zweckmäßigen bzw. alternativen Vergleichstherapie

Studie Gruppe	N	Baseline HbA1c (%) Mittelwert (SD)	Diabetesdauer (Jahre) Mittelwert (SD)	Anzahl der Studienabbrecher n (%)
Studie CV181040 siehe Abschnitt 4.3.1.2				
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin				
Shank 1995	30	Nur als Graphik	Nicht berichtet	6 (20)
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie				
Birkeland 1994				
Insulin	18	8,8 (1,27)	Gemeinsam:	0 (0,0)
Glibenclamid	16	8,6 (1,2)	7,6 (3,0)	2 (12,5)
Tovi 1998				
Insulin	19	9,3 (1,4)	11,5 (6,9)	3 (15,8)
Sulfonylharnstoffs	16	9,1 (1,2)	11,1 (6,7)	2 (12,5)
Nathan 1988				
Insulin	15	10,3 (1,5)	3,7 (3,9)	0 (0,0)
Glibenclamid	16	10,5 (2,4)	5,9 (10,8)	0 (0,0)
Turner 1998 (UKPDS 24)	458	Gemeinsam: 10,5 [9,4;12,1] ^c	Nicht berichtet	Gemeinsam: 102 (18,2)
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin 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	14 (13,2)
Glimepirid+Placebo	106	8,43 (0,80) ^b	8,0 (6,5) ^b	23 (21,7)
Garber AJ et al. 2008				
Glimepirid+Vildagliptin 50 mg	132	8,5 (0,9) ^b	6,9 (5,2) ^b	19 (14,4)
Glimepirid+Placebo	144	8,5 (1,0) ^b	7,8 (5,8) ^b	36 (25,0)
Lewin 2012				
Sulfonylharnstoff+Linagliptin 5 mg	161	8,6 (0,72) ^b	Nicht berichtet	10 (6,2)
Sulfonylharnstoff+Placebo	84	8,61 (0,85) ^b	Nicht berichtet	7 (8,3)

^bStandardabweichung (SD)^cMedian [Range]

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

Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin oder Insulin-Monotherapie:

Studie CV181040: siehe Abschnitt 4.3.1.2

Sulfonylharnstoff+Insulin gegen Sulfonylharnstoff

Die Studie von Shank 1995 ist eine doppelblinde, placebokontrollierte Studie mit vier Phasen. In Phase 1 (zwei Monate) wurden die Patienten mit Glipizid vorbehandelt um sicher zu stellen, dass eine Sulfonylharnstoff-Monotherapie für die Behandlung des Typ-2-Diabetes mellitus nicht ausreichend ist. In Phase 2 (drei Monate) wurden die Patienten zu drei Behandlungsgruppen randomisiert: Insulin+Placebo, Sulfonylharnstoff+Insulin oder Sulfonylharnstoff+Placebo. Während Phase 2 durfte die Insulindosis nicht erhöht werden. In der dreimonatigen Phase 3 wurde die Insulindosis dann hochtitriert. In der offenen Phase 4 erhielt die Sulfonylharnstoff-Gruppe für weitere sechs Monate zusätzlich Insulin, die Insulingruppe erhielt zusätzlich Sulfonylharnstoff.

Insulin gegen Sulfonylharnstoff

Bei der Studie von Birkeland 1994 handelt es sich um eine aktivkontrollierte, offene, parallele, monozentrische, randomisierte klinische Studie aus Norwegen. In der Studie wurden Typ-2-Diabetes-Patienten im Alter von $59,2 \pm 1,0$ Jahren und einer Diabetesdauer von $7,6 \pm 0,5$ Jahren, deren Blutzucker mit Diät, Bewegung und Sulfonylharnstoff-Monotherapie nicht ausreichend kontrolliert werden kann, untersucht. Primäres Ziel dieser Studie war es, die Wirksamkeit in Bezug auf die Blutzuckerkontrolle, der Insulinresistenz und des kardialen Risikos von Sulfonylharnstoff mit Insulin zu vergleichen. In der zur Glibenclamid-Behandlung randomisierten Gruppe war der Anteil der Frauen deutlich niedriger (18,75%).

Bei der Studie von Tovi 1998 handelt es sich um eine randomisierte, kontrollierte, parallele, monozentrische, offene Studie aus Schweden. Bei der Patientenpopulation handelte es sich um Typ-2-Diabetes-Patienten, welche durch eine Sulfonylharnstoff-Monotherapie ihren Blutzucker nicht ausreichend kontrollieren konnten. Bei den Patienten handelte es sich um ältere Personen ($74,2 \pm 14,2$ Jahre) mit einer durchschnittlichen Diabetesdauer von $11,5 \pm 6,4$ Jahren. Das primäre Ziel dieser Studie war die Untersuchung der Lebensqualität von älteren Typ-2-Diabetes-Patienten durch verbesserte glykämische Kontrolle.

Die Studie von Nathan 1999 ist eine randomisierte, placebo-kontrollierte, doppelblinde, klinische Studie, in der die Wirksamkeit, das Risiko und der Vorteil einer Insulin-Behandlung gegenüber einer Behandlung mit Sulfonylharnstoff untersucht wurde. Bei der Patientenpopulation handelt es sich um nicht insulinabhängige Patienten mit Diabetes mellitus Typ 2, im Alter von 30-70 Jahren, deren Blutzucker durch Diät alleine nicht ausreichend kontrolliert ist (behandlungsnaive Patienten).

Die Studie von Turner 1998 (UKPDS 24) ist eine multizentrische, prospektive, randomisierte kontrollierte Interventionsstudie aus Großbritannien mit Patienten mit neu diagnostiziertem Typ 2 Diabetes mellitus. Patienten, deren Erkrankungen durch die während der dreimonatigen Run-in-Periode eingehaltene Diät allein nicht ausreichend kontrolliert werden konnten, wurden mit Sulfonylharnstoff, Insulin oder Metformin behandelt. Die Patienten (Primary diet failure Group) wurden über 6 Jahre hinweg untersucht und das Ansprechen auf die Therapie untereinander und mit einer Hauptrandomisierungsgruppe (Main randomisation Group) verglichen.

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

Studie CV181040: siehe Abschnitt 4.3.1.2

In der Studie von Hermansen 2007 handelt es sich um eine randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte, 24-wöchige Studie, welche die Wirksamkeit und Sicherheit von Sitagliptin als Add-on Therapie zu Glimepirid (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 Glimepirid oder Glimepirid+Metformin nicht ausreichend kontrolliert werden konnte ($HbA1c \geq 7,5\%$ und $\leq 10,5\%$). Nach einer Einstellungsphase auf die Hintergrundmedikation (Glimepirid-Monotherapie 4-8 mg/Tag oder Glimepirid 4-8 mg/Tag+Metformin 1500-3000 mg/Tag) von insgesamt bis zu 14 Wochen, erfolgte eine 2-wöchige 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 Glimepirid (mit oder ohne Metformin). Die Studiendauer nach Randomisierung betrug 24 Wochen. Die Auswertung der Studie erfolgte in zwei Strata (Vorbehandlung mit Glimepirid-Monotherapie oder Glimepirid+Metformin). Zur Möglichkeit einer Dosisänderung der Hintergrundtherapie (Glimepirid), 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.

Bei der Studie von Garber 2008 handelt es 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 4 mg/Tag mit Placebo verglichen hat. Eingeschlossen wurden Patienten im Alter von 18 bis 80 mit Diabetes mellitus Typ 2, deren Blutzuckerwert mit einer Sulfonylharnstoff-Monotherapie nicht ausreichend kontrolliert werden konnte ($HbA1c$ 7,5% bis 11,0%). Alle Patienten wurden auf Glimepirid 4 mg/Tag eingestellt. Nach vier Wochen stabiler Therapie und Erfüllen der Ein- und Ausschlusskriterien erfolgte die Randomisierung in die Behandlungsgruppen (Vildagliptin 50 mg oder 100 mg, 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 empfohlene Dosierung bei einer Zweifach-Kombinationstherapie mit einem Sulfonylharnstoff in Deutschland ist.

Lewin 2012: Es handelt sich um eine randomisierte, doppelblinde, parallele, multizentrische, aktivkontrollierte, 18-wöchige Studie, welche die Wirksamkeit und Sicherheit von Linagliptin 5 mg/Tag als Add-on Therapie zu einer Sulfonylharnstofftherapie mit Placebo verglichen hat. Es gab bei dieser Studie keine Vorschriften hinsichtlich der Sulfonylharnstoff-Einzelsubstanz.

In einer Wash-out-Phase von vier Wochen wurden Patienten, die zusätzlich mit anderen oralen Antidiabetika behandelt wurden, auf eine Sulfonylharnstoff-Monotherapie umgestellt; danach erfolgte für alle Patienten eine 2-wöchige Placebo-Run-in-Phase. Das jeweilige Sulfonylharnstoff-Präparat wurde für die gesamte Dauer der Studie (inklusive der 4-wöchigen Wash-out-Phase und der 2-wöchigen 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.

4.3.2.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

Tabelle 4-64: 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 CV101840 siehe Abschnitt 4.3.1.2.2							
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin							
Shank 1995	ja	nein	nein**	nein**	unklar*	nein	hoch
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie							
Birkeland 1994	unklar*	nein	nein	nein	unklar*	nein	hoch
Tovi 1994	unklar*	nein	nein	nein	unklar*	nein	hoch
Nathan 1988	ja	ja	ja	ja	unklar*	nein	niedrig
Turner (UKPDS)	unklar*	nein	nein	nein	unklar*	nein	hoch
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor							
Hermansen 2007	ja	ja	ja	ja	unklar*	nein	niedrig
Garber 2008	ja	ja	ja	ja	unklar*	nein	niedrig
Lewin 2012	ja	ja	ja	ja	unklar*	nein	niedrig

* es liegt nur die Publikation der Studie vor (kein Studienbericht); ** Insulin Behandlung war nicht verblindet

Begründen Sie für jede Studie die abschließende Einschätzung.

Für alle zu beurteilenden Studien zur zweckmäßigen Vergleichstherapie gilt, dass das Verzerrungspotenzial auf Studienebene mit Ausnahme von Nathan 1988 als „hoch“ eingestuft werden kann, da die Studien mit der zweckmäßigen Vergleichstherapie zwar alle randomisiert durchgeführt wurden, jedoch nicht verblindet gewesen sind.

Für die DPP-4-Studien erfolgte die Gruppenzuteilung jeweils verdeckt, sowohl Patient als auch Behandler waren verblindet (doppelblind) und es gab auch keine sonstigen das Verzerrungspotenzial beeinflussenden Faktoren. Daher wird hier ein niedriges Verzerrungspotential angenommen.

Für alle Studien war eine ergebnisgesteuerte Berichterstattung nicht abschließend zu beurteilen (auszuschließen), da keine Studienberichte 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 4-65: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	Hypoglykämien	Gewichtsveränderung	Unerwünschte Ereignisse	Folgekomplikationen	Gesamtmortalität
CV181040	ja	ja	ja	nein	nein
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin					
Shank 1995	nein	ja	nein	nein	nein
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie					
Birkeland 1994	ja	ja	nein	nein	nein
Tovi 1998	ja	ja	nein	nein	nein
Nathan 1999	ja	ja	nein	nein	nein
Turner 1998 (UKPDS 24)	ja	ja	nein	nein	nein
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor					
Hermansen 2007	ja	ja	ja	nein	nein
Garber 2008	ja	ja	ja	nein	nein
Lewin 2012	ja	ja	ja	nein	nein

Tabelle 4-66: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	Hyperosmolare Komata	Stationäre Behandlungen	Gesundheitsbezogene Lebensqualität	Therapiezufriedenheit	Ausmaß der Blutzuckersenkung (HbA1c)
CV181040	nein	nein	nein	nein	ja
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin					
Shank 1995	nein	nein	ja	ja	ja
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie					
Birkeland 1994	nein	nein	nein	nein	ja
Tovi 1998	nein	nein	ja	ja	ja
Nathan 1999	nein	nein	ja	ja	ja
Turner 1998 (UKPDS 24)	nein	nein	nein	nein	ja
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor					
Hermansen 2007	nein	nein	nein	nein	ja
Garber 2008	nein	nein	nein	nein	ja
Lewin 2012	nein	nein	nein	nein	ja

Im Folgenden soll der indirekte Vergleich mit der zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) und der alternativen Vergleichstherapie (Sulfonylharnstoff+DPP-4-Inhibitor) dargestellt werden.

- Für die zweckmäßige Vergleichstherapie Sulfonylharnstoff+Insulin wird ein indirekter Vergleich (eine Studie von Shank 1995) nach 24 Wochen durchgeführt.
- Für den indirekten Vergleich gegen Insulin-Monotherapie nach 50 Wochen wurden fünf Studien identifiziert (Birkeland 1994, Tovi 1998, Nathan 1988, Turner 1998 und Shank 1995) und wenn möglich gepoolt. Da die Studien von Nathan und Turner behandlungsnaive Patienten einschließen, sollen diese Studien nochmal separat als Sensitivitätsanalyse dargestellt werden. Die Studie von Shank 1995 berichtet nur nach 24 Wochen, so dass die Daten dargestellt, jedoch nicht in die Meta-Analyse eingeschlossen werden.
- Für die alternative Vergleichstherapie sollen zunächst alle drei Studien mit den verschiedenen DPP-4-Inhibitoren (Sitagliptin, Linagliptin und Vildagliptin) gepoolt werden und dann die Studie von Hermansen 2007 mit Sitagliptin als Leitsubstanz separat betrachtet werden.

4.3.2.1.3.1 Ausmaß der Blutzuckerkontrolle (HbA1c) – 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 4-67: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff+Saxagliptin	Sulfonylharnstoff+Insulin	Insulin	Sulfonylharnstoff
1	CV181040	•			•
1	Shank 1995		•		•
4	Birkeland 1994 Tovi 1998 Nathan 1999 Turner 1998 (UKPDS 24)			• • • •	• • • •

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

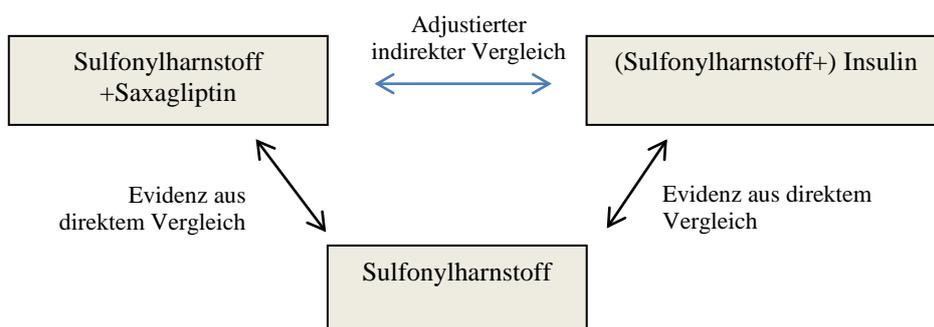
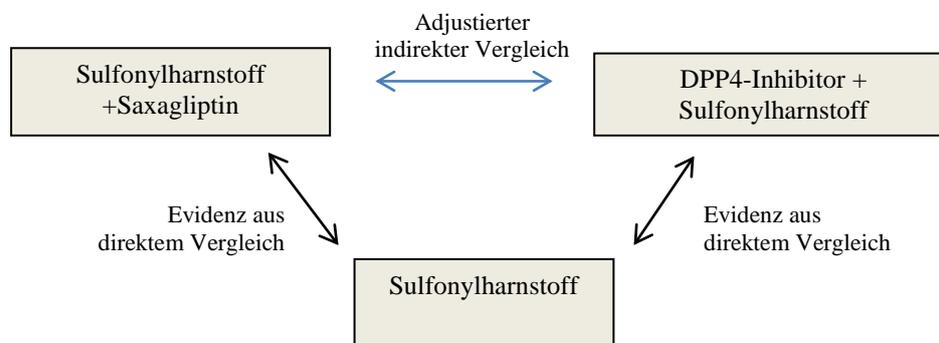


Tabelle 4-68: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff +Saxagliptin	Sulfonylharnstoff + DPP-4-Inhibitor	Sulfonylharnstoff + Placebo
1	CV181040	•		•
3	Hermansen et al 2007 Garber et al. 2008 Lewin 2012		• • •	• • •



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

Tabelle 4-69: Operationalisierung von Ausmaß der Blutzuckerkontrolle HbA1c

Studie	Operationalisierung
Studie CV181040	siehe Abschnitt 4.3.1.3.1.1
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin	
Shank 1995	keine Veränderung des HbA1c berichtet
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie	
Birkeland 1994	Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 52
Tovi 1998	Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 24 und 52
Nathan 1995	Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 36
Turner 1998 (UKPDS 24)	nur mediane HbA1c-Werte
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor	
Hermansen et al. 2007	Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 24
Garber et al. 2008	Änderung des absoluten HbA1c-Wertes (in %) von Studienbeginn bis Woche 24
Lewin 2012	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-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-70: 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 CV181040 siehe Abschnitt 4.3.1.3.1.1						
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin						
Shank 1995	hoch	unklar	nein**	unklar*	nein	hoch
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie						
Birkeland 1994	hoch	nein	nein**	unklar*	nein	hoch
Tovi 1998	niedrig	nein	nein**	unklar*	nein	hoch
Nathan 1995	hoch	nein	nein**	unklar*	nein	hoch
Turner 1998 (UKPDS 24)	hoch	nein	ja	unklar*	nein	hoch
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	niedrig	ja	ja	unklar*	nein	niedrig
Garber et al. 2008	niedrig	ja	ja	unklar*	nein	niedrig
Lewin 2012	niedrig	ja	ja	unklar*	nein	niedrig

* es liegt nur die Publikation der Studie vor (kein Studienbericht); ** aufgrund der geringen Fallzahl kann kein valides Ergebnis berechnet werden

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für den zu bewertenden Endpunkt Ausmaß der Blutzuckerkontrolle (HbA1c) für alle Insulin-Studien als „hoch“ bewertet, da aufgrund des hohen Verzerrungspotenzial auf Studienebene eine endpunktspezifische Verzerrung nicht ausgeschlossen werden kann. Auch die geringe Fallzahl in den Studien kann eine Verzerrung verursachen.

Für die DPP-4-Studien war der Endpunkterheber jeweils verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es gab auch keine sonstigen das Verzerrungspotenzial beeinflussenden Faktoren. Daher wird hier ein niedriges Verzerrungspotential angenommen.

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 alle Studien mit der zweckmäßigen Vergleichstherapie 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 4-71: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Saxagliptin				Komparator: Sulfonylharnstoff				Zweckmäßige Vergleichstherapie: Insulin bzw. alternative Vergleichstherapie DPP-4-Inhibitor				Gruppen- unterschied % Mittelwert (SE)
	N	Aus- gangs- wert % Mittel- wert (SE/SD)	End- wert % Mittel- wert (SE/SD)	Ver- änder- ung % Mittel- wert ^o (SE)	N	Aus- gangs- wert % Mittel- wert (SE/SD)	End- wert % Mittel- wert (SE/SD)	Ver- änder- ung % Mittel- wert (SE/SD)	N	Aus- gangs- wert % Mittel- wert (SE/SD)	End- wert % Mittel- wert (SE/SD)	Ver- änder- ung % Mittel- wert (SE/SD)	
CV181040													
24 Wochen	5 mg: 250	8,48 (0,056 ^a)	7,84 (0,074 ^a)	-0,64 (0,059 ^a)	264	8,44 (0,055 ^a)	8,52 (0,077 ^a)	0,08 (0,059 ^a)					-0,72 (0,082) [-0,88;-0,56]
50 Wochen	5 mg: 250	8,48 (0,056 ^a)	8,14 (0,074 ^a)	-0,33 (0,060 ^a)	264	8,44 (0,055 ^a)	8,74 (0,071 ^a)	0,30 (0,058 ^a)					-0,63 (0,083) [-0,79;-0,47]
Studien für indirekten Vergleich Sulfonylharnstoff +Saxagliptin vs. Sulfonylharnstoff+Insulin													
Shank 1995 24 Wochen					10	nicht berichtet	8,4 (0,6 ^a)	nicht berichtet	9	nicht berichtet	7,1 (0,3 ^a)	nicht berichtet	nicht berechenbar
Studien für indirekten Vergleich Sulfonylharnstoff +Saxagliptin vs. Insulin-Monotherapie													
Birkeland 1994 52 Wochen					16	8,6 (0,3 ^a)	9,5 (0,4 ^a)	0,9 (0,4 ^a)^^	18	8,8 (0,3 ^a)	7,9 (0,3 ^a)	-0,9 (0,3 ^a)^^	-1,80 (0,5000) [-2,78;-0,82]^
Tovi 1998 24 Wochen					16	9,1 (1,2 ^b)	nicht berichtet	-0,2 [-0,8;-0,4]	19	9,3 (1,4 ^b)	nicht berichtet	-2,0 [-2,7;-1,3]	-1,80 (0,4703) [-2,72;-0,88]^
52 Wochen					16	9,1 (1,2 ^b)	nicht berichtet	0,3 [-0,4;0,9]	19	9,3 (1,4 ^b)	nicht berichtet	-2,0 [-2,8;-1,2]	-2,30 (0,5102) [-3,30;-1,30]^

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Saxagliptin		Komparator: Sulfonylharnstoff			Zweckmäßige Vergleichstherapie:				Gruppen- unterschied % Mittelwert (SE)	
						Insulin bzw. alternative Vergleichstherapie DPP-4-Inhibitor					
Nathan 1988¹ 36 Wochen			16	10,5 (2,4) ^b	7,63 (1,7) ^b	-2,87 (2,4) ^b ^^	15	10,3 (1,5) ^b	7,14 (1,1) ^b	-3,16 (1,5) ^b ^^	-0,29 (0,7141) [-1,69;1,11]^
Turner 1998¹ (UKPDS 24) 52 Wochen			231	10,5 [9,3 – 12,1]**	6,8 [6,0 – 8,6]**	nicht berichtet	178	10,6 [9,5- 12,3]**	7,0 [6,1 – 7,8]**	nicht berichtet	nicht berechenbar
Shank 1995 24 Wochen			10	nicht bericht et	8,4 (0,6) ^a	nicht berichtet	9	nicht berichtet	7,8 (0,4) ^a	nicht berichtet	nicht berechenbar
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor											
Hermansen et al. 2007			106	8,42 (0,79) ^b	nicht berichtet	0,27 [0,09;0,4 5]	106	8,43 (0,80) ^b	nicht berichtet	-0,30 [-0,48;- 0,12]	-0,57 [-0,82;-0,32]
Garber et al. 2008			144	8,5 (1,0) ^b	nicht berichtet	0,07 (0,09) ^a	132	8,5 (0,9) ^b	nicht berichtet	-0,58 (0,1) ^a	-0,64 (0,13) ^a
Lewin 2012			84	8,6 (0,72) ^b	nicht berichtet	-0,07 (0,10) ^a	161	8,61 (0,85) ^b	nicht berichtet	-0,54 (0,07) ^a	-0,47 (0,12) [-0,70; -0,24]

° adjustiert nach Ausgangswert; ¹ Behandlungsnaiv; ^a Standardfehler (SE); ^b Standardabweichung (SD); ** Median [Interquartilsabstand] → kein Mittelwert/SD abschätzbar;

^^ Streuungsmaß geschätzt aus dem Mittelwert der Standardabweichungen der anderen Studien der Vergleichstherapie oder das größere Streuungsmaß der Vorher-Nachher- Messung, falls kein Mittelwert von mindestens 2 Studien berechenbar ist; ^ eigene Berechnung

Sulfonylharnstoff+Insulin versus Sulfonylharnstoff

Die Studie von Shank 1995 berichtet jeweils die Endwerte der Behandlungsgruppen nach 6 Monaten. Wegen fehlenden Ausgangswerten oder durch Angabe von einer Veränderung zum Ausgangswert kann kein Effektschätzer berechnet werden. Aufgrund dessen konnte diese Studie nicht in der Meta-Analyse oder in einen indirekten Vergleich berücksichtigt werden.

Insulin versus Sulfonylharnstoff

In der Studie von Birkeland 1994 zeigt die Insulin-Gruppe eine Senkung des HbA1c-Wertes um -0,9%, wobei es eine Zunahme von 0,9% in der Sulfonylharnstoff-Gruppe gab. Diese Gruppenunterschiede sind statistisch signifikant.

In der Studie von Tovi 1998 gab es einen signifikanten Unterschied in den Behandlungsgruppen von -1,8% nach sechs Monaten und -2,3% nach zwölf Monaten bezüglich der HbA1c-Veränderung zu Gunsten der Insulin-Gruppe. Dort gab es eine durchschnittliche Senkung von -2,0% nach sechs Monaten und nach zwölf Monaten.

In der Studie von Nathan 1988 konnte der HbA1c-Wert in beiden Behandlungsgruppen signifikant verringert werden (-3,16% vs. -2,87%), jedoch war der Unterschied zwischen den Behandlungsgruppen nicht signifikant. Die Population von Nathan wurde nur 36 Wochen untersucht und die Studienpopulation war behandlungsnaiv. Die Studie konnte aufgrund der unterschiedlichen Studiendauer und Patientenpopulation nicht in der Meta-Analyse berücksichtigt werden.

Auch die UKPDS-Studie schloss nur behandlungsnaive Patienten ein. Die Studie berichtete ebenfalls eine Reduzierung der medianen HbA1c-Werte in beiden Behandlungsgruppen innerhalb eines Jahres. Die Ausgangs-HbA1c-Werte betragen in der Sulfonylharnstoff-Gruppe 10,5% und in der Insulin-Gruppe 10,6%, wobei diese sich nach zwölf Monaten auf 6,8% bzw. 7,0% reduzierten. Dadurch dass keine Mittelwerte und Standardabweichungen angegeben wurden, konnte die UKPDS-Studie nicht in der Meta-Analyse berücksichtigt werden.

In der Meta-Analyse konnten nur Birkeland 1994 und Tovi 1998 aufgrund des ähnlichen Studiendesigns berücksichtigt werden.

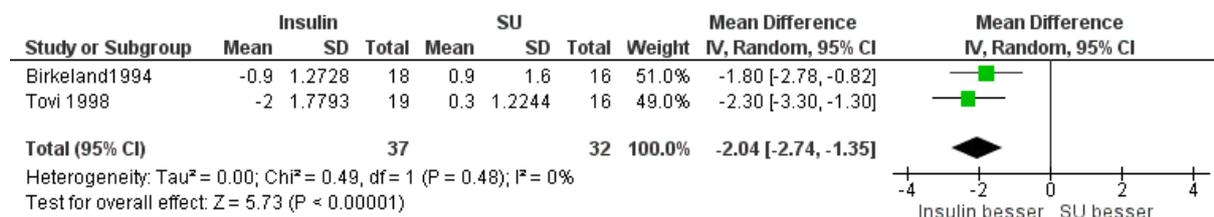


Abbildung 4: Meta-Analyse der eingeschlossenen Studien zum Effekt von Insulin vs. Sulfonylharnstoff auf den Endpunkt Ausmaß der Blutzuckerkontrolle

Der Gesamtschätzer der Meta-Analyse zeigte eine signifikante Reduzierung des HbA1c-Wertes in der Insulin-Behandlungsgruppe gegenüber der Sulfonylharnstoff-Gruppe bei homogener Datenlage ($I^2=0\%$, $p=0,48$).

Sulfonylharnstoff +DPP-4-Inhibitor versus Sulfonylharnstoff:

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. Auch die Meta-Analyse wies auf keine Heterogenität der Studienergebnisse hin ($I^2=0\%$; $p=0.61$), weshalb eine Zusammenfassung der Ergebnisse in einen gemeinsamen Effektschätzer gerechtfertigt erschien.

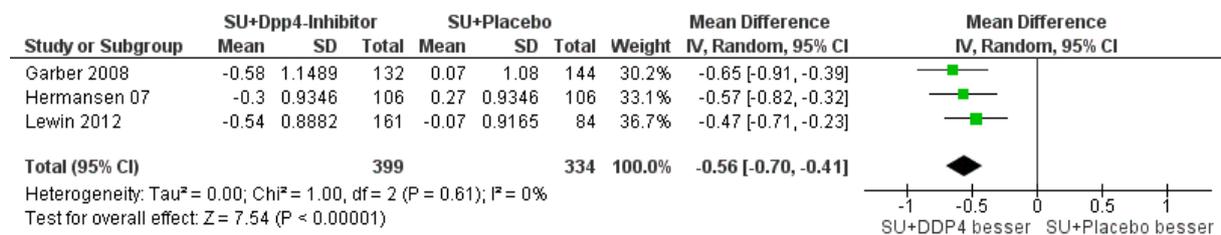


Abbildung 5: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt Ausmaß der Blutzuckerkontrolle

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.*

Ergebnisse des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Insulin-Monotherapie:

Tabelle 4-72: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) - indirekte Vergleiche

Mittelwertdifferenz des direkten Vergleiches der Studie CV181040 nach 50 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Mittelwertdifferenz des direkten Vergleiches (Insulin versus Sulfonylharnstoff)	Mittelwertdifferenz des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Insulin HbA1c% [95%-KI]
-0,63 (0,083)	-2,0449 (0,3571) ¹	1,41 [0,70;2,13]
-0,63 (0,083)	-0,29 (0,7141) ²	-0,34 [-1,75;1,07]

¹ Gepoolte Effektschätzer der Meta-Analyse von Birkeland 1994 und Tovi 1998 (nach 52 Wochen)

² Effektschätzer von Nathan 1988 (nach 36 Wochen)

KI: Konfidenzintervall

Der indirekte Vergleich mit Studie CV181040 und der Gesamtschätzer der Meta-Analyse von Birkeland 1994 und Tovi 1998 zeigte ein signifikantes Ergebnis zugunsten der Insulin-Behandlungsgruppe gegenüber der Sulfonylharnstoff+Saxagliptin-Gruppe. Da der HbA1c-senkenden Effekte von Antidiabetika u. a. stark vom HbA1c-Ausgangswert abhängt, muss das Ergebnis aber hinterfragt werden, da der durchschnittliche HbA1c-Ausgangswert insbesondere in der Studie von Tovi 1998 höher war (9,1% in der Sulfonylharnstoff-Gruppe und 9,3% in der Insulin-Gruppe) als der in der Studie CV181040 (8,48% in der Saxagliptin-Gruppe und 8,44% in der Placebo-Gruppe) und in der Studie von Birkeland 1994 (8,6% in der Sulfonylharnstoff-Gruppe und 8,8% in der Insulin-Gruppe).

Vergleicht man die Studie CV181040 mit Nathan 1988, der jedoch ein anderes Patientenkollektiv (behandlungsnaiv) einschloss sowie eine andere Studiendauer berichtet, zeigte sich kein Unterschied bezüglich der HbA1c-Senkung zwischen den Behandlungsgruppen.

Ergebnisse des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor

Tabelle 4-73: Ergebnisse für Ausmaß der Blutzuckerkontrolle (HbA1c) - indirekte Vergleiche

Mittelwertdifferenz des direkten Vergleiches der Studie CV181040 nach 24 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Mittelwertdifferenz des direkten Vergleiches (Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff)	Mittelwertdifferenz des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor HbA1c% [95%-KI]
-0,72 (0,082)	-0,56 (0,7023) ¹	-0,16 [-1,55;1,23]
-0,72 (0,082)	-0,57 (0,1284) ²	-0,15 [-0,45;0,15]

¹ Gepoolter Effektschätzer der Meta-Analyse von Hermansen, Garber und Lewin 2012

² Effektschätzer von Hermansen 2007 (Sitagliptin als Leitsubstanz für die DPP-4-Inhibitoren)

KI: Konfidenzintervall

Der indirekte Vergleich in Bezug auf den Endpunkt Ausmaß der Blutzuckerkontrolle (HbA1c) zeigte 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 – 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 4-74: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff +Saxagliptin vs. Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff +Saxagliptin	Sulfonylharnstoff+ Insulin	Insulin	Sulfonylharnstoff
1	CV181040	•			•
1	Shank 1995		•		•
4	Birkeland 1994 Tovi 1998 Nathan 1999 Turner 1998 (UKPDS 24)			• • • •	• • • •

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

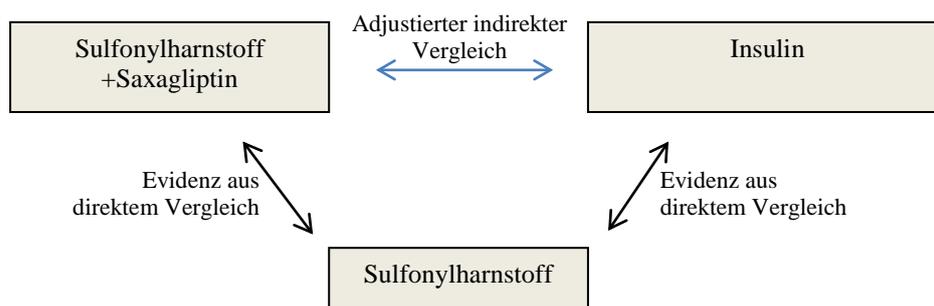
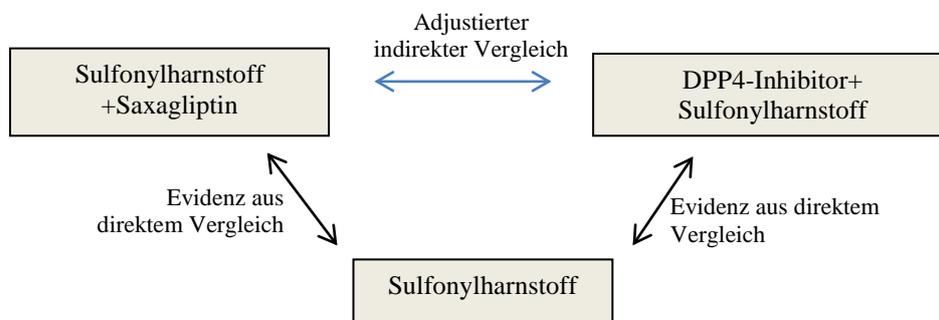


Tabelle 4-75: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff +Saxagliptin	Sulfonylharnstoff +DPP-4-Inhibitor	Sulfonylharnstoff +Placebo
1	CV181040	•		•
3	Hermansen et al 2007 Garber et al. 2008 Lewin 2012		• • •	• • •



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

Tabelle 4-76: Operationalisierung von Hypoglykämien

Studie	Operationalisierung
Studie CV181040 siehe Abschnitt 4.3.1.3.1.2	
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin	
Shank 1995	Nicht berichtet
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie	
Birkeland 1994	Anzahl der Patienten mit Hypoglykämie
Tovi 1998	Anzahl der hypoglykämischen Symptome
Nathan 1988	Mittlere Anzahl an Hypoglykämien in den 9 Monaten pro Behandlungsgruppe
Turner 1998 (UKPDS 24)	Mittlerer jährlicher Anteil an Patienten mit Hypoglykämien (Beobachtungszeit war 6 Jahre)
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor	
Hermansen 2007	<u>Gesamt:</u> Durch den Patienten berichtete hypoglykämische Ereignisse <u>Schwer:</u> Vom Studienarzt als schwer eingeschätzt oder Ereignisse, die medizinische Hilfe erforderten
Garber 2008	<u>Bestätigt:</u> Hypoglykämie-Symptome und Blutzuckerwert (aus Selbstmessung) $\leq 3,1$ mmol/L
Lewin 2012	<u>Gesamt:</u> Durch den Studienarzt als Hypoglykämie eingeschätzt*

*Nähere Angaben lagen hier aufgrund der fehlenden Vollpublikation nicht vor.

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

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

Tabelle 4-77: 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 CV181040 siehe Abschnitt 4.3.1.3.1.2						
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin						
Birkeland 1994	hoch	nein	nein**	unklar*	nein	hoch
Tovi 1998	niedrig	nein	nein**	unklar*	nein	hoch
Nathan 1988	hoch	nein	nein**	unklar*	nein	hoch
Turner 1998 (UKPDS 24)	hoch	nein	ja	unklar*	nein	hoch
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen 2007	niedrig	ja	ja	unklar*	nein	niedrig
Garber 2008	niedrig	ja	ja	unklar*	nein	niedrig
Lewin 2012	niedrig	ja	ja	unklar*	nein	niedrig

*es lag nur die Publikation der Studie vor (kein Studienbericht) ** aufgrund der geringen Fallzahl kann kein valides Ergebnis berechnet werden

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für den zu bewertenden Endpunkt Hypoglykämien für alle Insulin-Studien als „hoch“ bewertet, da aufgrund des hohen Verzerrungspotenzial auf Studienebene eine endpunktspezifische Verzerrung nicht ausgeschlossen werden kann. Auch die geringe Fallzahl in den Studien kann eine Verzerrung verursachen.

Für die DPP-4-Studien war der Endpunkterheber jeweils verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es gab auch keine sonstigen das Verzerrungspotenzial beeinflussenden Faktoren. 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 alle Studien mit der zweckmäßigen Vergleichstherapie 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 4-78: Ergebnisse für Hypoglykämien aus RCT für indirekte Vergleiche

Studie	Zu bewertendes Arzneimittel: Sulfonylharnstoff+Saxagliptin		Komparator: Sulfonylharnstoff		Zweckmäßige Vergleichstherapie: Insulin bzw. alternative Vergleichstherapie DPP-4-Inhibitor		Odds Ratio [95%-KI]
	N	n (%)	N	n (%)	N	n (%)	
CV181040							
Patientenberichtete Hypoglykämien							
24 Wochen	5 mg:253	37 (14,6)	267	27 (10,1)			1,52 [0,90;2,58]
50 Wochen	5 mg:253	48 (19,0)	267	44 (16,5)			1,19 [0,76;1,86]
Bestätigte Hypoglykämien							
24 Wochen	5 mg:253	2 (0,8)	267	2 (0,7)			1,06 [0,15;7,55]
50 Wochen	5 mg:253	5 (2,0)	267	7 (2,6)			0,75 [0,23;2,39]
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin							
Birkeland 1994							
52 Wochen			16	0	18	8	26,71[1,39;513,09]
Tovi 1998							
52 Wochen			16	nicht berichtet	19	0	nicht berechenbar
Nathan 1988							
36 Wochen			15	nicht berichtet	15	nicht berichtet	nicht berechenbar
Turner 1998** (UKPDS 24)							
52 Wochen			nicht berichtet	21 (durchschnittliche jährliche Hypoglykämie-Rate)	nicht berichtet	52 (durchschnittliche jährliche Hypoglykämie-Rate)	nicht berechenbar

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

Studie	Zu bewertendes Arzneimittel: Sulfonylharnstoff+Saxagliptin		Komparator: Sulfonylharnstoff		Zweckmäßige Vergleichstherapie: Insulin bzw. alternative Vergleichstherapie DPP-4-Inhibitor		Odds Ratio [95%-KI]
	N	n (%)	N	n (%)	N	n (%)	
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor							
Hermansen 2007							
gesamt			106	3 (2,8)	106	8 (7,5)	2,80 [0,72;10,87]*
schwer			106	0	106	0	nicht berechenbar
Garber 2008							
bestätigt			176	1 (0,6)	170	2 (1,2)	2,08 [0,19;23,19]*
Lewin 2012							
gesamt			84	4 (4,8)	161	9 (5,6)	1,18 [0,35;3,97]

* eigene Berechnung; **Behandlungsguppen: Glibenclamid und langwirkendes Insulin; KI: Konfidenzintervall

Insulin versus Sulfonylharnstoff+Insulin

Eine Studie (Shank 1995) berichtet den Vergleich Insulin versus Sulfonylharnstoff+Insulin jedoch nicht die Anzahl der Patienten mit Hypoglykämien.

Insulin versus Sulfonylharnstoff

In der Studie von Birkeland 1994 wurde in der Insulin-Gruppe bei 8 von 18 Patienten eine patientenberichtete Hypoglykämie beobachtet.

In der Studie von Tovi 1998 gab es zu Beginn in beiden Gruppen die gleiche Anzahl an Hypoglykämien (im Mittel 1,21 in der Insulin-Gruppe und 1,63 in der Sulfonylharnstoff-Gruppe), danach stiegen die Symptome bei der Insulingruppe an (im Mittel 2,0 in der Insulin-Gruppe und 2,44 in der Sulfonylharnstoff-Gruppe). Nach Studienende verringerte sich jedoch die Anzahl in der Insulin-Gruppe wieder deutlich (im Mittel 0,84 in der Insulin-Gruppe und 2,25 in der Sulfonylharnstoff-Gruppe).

In der Studie von Nathan 1988 wurde eine mittlere Anzahl an Hypoglykämien (Standardabweichung) von 1,8 (2,9) in der Insulin-Gruppe und 1,63 (1,6) in der Sulfonylharnstoff-Gruppe erreicht, was einen nicht signifikanten Unterschied der Behandlungsgruppen ergab ($p=0,84$). Es gab in keiner Behandlungsgruppe schwere Hypoglykämien.

In der UKPDS-Studie sowie in der Studie von Shank 1995 wurde die Anzahl an Patienten mit mindestens einer Hypoglykämie nicht berichtet.

Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff:

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 (patientenberichtet):

Diese Gesamtrate wurde in Studie CV181040 und in den Studien von Hermansen 2007 und in der Publikation von Lewin 2012 berichtet. Die Informationen zur Definition des Endpunktes in der Publikation von Lewin 2012 waren aufgrund der fehlenden Information eingeschränkt, es zeigten sich aber keine Hinweise auf bedeutsame Heterogenität der Ergebnisse ($I^2=0\%$; $p=0,35$). Die Zusammenfassung der Ergebnisse in einem gemeinsamen Effektschätzer im Rahmen einer Meta-Analyse erschien damit gerechtfertigt. Zusätzlich wurde der indirekte Vergleich mit der DPP-4-Leitsubstanz Sitagliptin berechnet.

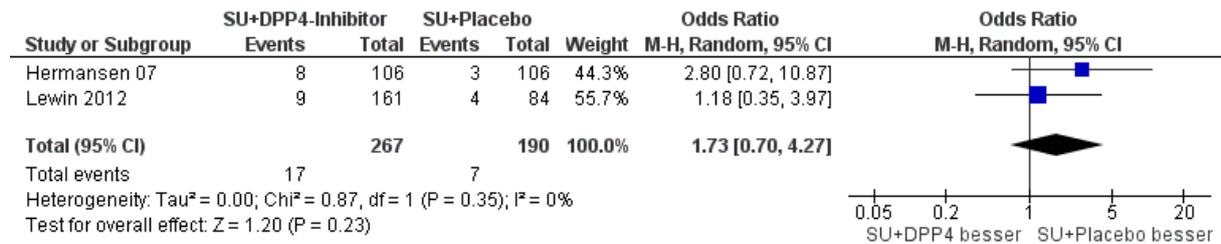


Abbildung 6: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt Hypoglykämien (patientenberichtet)

Die Meta-Analyse verdeutlicht, dass die Patienten der Sulfonylharnstoff+DPP-4-Inhibitor-Gruppe keinen signifikanten Unterschied für das Auftreten von Hypoglykämien gegenüber der Sulfonylharnstoff+Placebo-Gruppe haben. Da es keinen Hinweis auf eine bedeutsame Heterogenität der Ergebnisse gibt ($I^2=0\%$; $p=0,35$), ist die Zusammenfassung in einem gemeinsamen Effektschätzer im Rahmen der Meta-Analyse gerechtfertigt. Jedoch zeigt der Schätzer mit 1,73 und einem 95%-KI von [0,70;4,27] keinen signifikanten Unterschied in den Behandlungsgruppen, was auch schon in den Einzelbetrachtungen gezeigt wurde.

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.*

Ergebnisse des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Insulin:

Tabelle 4-79: Ergebnisse für Hypoglykämien - indirekte Vergleiche

Odds Ratio des direkten Vergleiches der Studie CV181040 nach 50 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Odds Ratio des direkten Vergleiches (Insulin versus Sulfonylharnstoff)	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Insulin [95%-KI]
1,19 (0,2300)	26,71 (1,5078)*	0,04 [0,00;0,88]

* Effektschätzer von Birkeland 1994; KI: Konfidenzintervall

Der indirekte Vergleich der Studie CV181040 und der Publikation von Birkeland 1994 ergab ein signifikantes Ergebnis zugunsten von Sulfonylharnstoff+Saxagliptin gegenüber Insulin, jedoch sollte das Ergebnis aufgrund der geringen Stichprobengröße (n=34) in der Studie von Birkeland 1994 und der Tatsache, dass die Intensität der Blutzuckerkontrolle unter Sulfonylharnstoff+Saxagliptin geringer war, mit Vorsicht interpretiert werden.

Ergebnisse des indirekten Vergleichs Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor

Tabelle 4-80: Ergebnisse für Hypoglykämien (patientenberichtet/bestätigt)- indirekte Vergleiche

Odds Ratio des direkten Vergleiches der Studie CV181040 nach 24 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Odds Ratio des direkten Vergleiches (Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff)	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor OR [95%-KI]
Patientenberichtete Hypoglykämien		
1,52 (0,2699)	1,73 (0,4602)*	0,88 [0,31;2,50]
1,52 (0,2699)	2,80 (0,6916)**	0,54 [0,13;2,33]
Bestätigte Hypoglykämien		
1,06 (1,0039)	2,08 (1,2295)***	0,51 [0,02;11,38]

* Gepoolter Effektschätzer der Meta-Analyse von Hermansen, Garber und Lewin 2012

** Effektschätzer von Hermansen 2007 (Sitagliptin als Leitsubstanz für die DPP4-Inhibitoren)

*** Effektschätzer von Garber 2008

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 alternativen Vergleichstherapie (DPP-4-Inhibitoren als Gruppe bzw. Sitagliptin).

4.3.2.1.3.1.2 Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien

Im Folgenden werden zunächst die in den vorherigen Abschnitten 4.3.2.1.3.1 und 4.3.2.1.3.1.1 berichteten Endpunkte Ausmaß der Blutzuckerkontrolle und Hypoglykämien gemeinsam betrachtet, da das Auftreten von Hypoglykämien nur unter Berücksichtigung der Blutzuckerkontrolle sinnvoll beurteilt werden kann (siehe Abschnitt 4.2.5.2).

Tabelle 4-81: Ergebnisse für gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien aus RCT mit dem zu bewertenden Arzneimittel – indirekter Vergleich

Zeitpunkte	Studie(n)	Blutzuckerkontrolle	Hypoglykämien	Gesamtbeurteilung
Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie				
52 Wochen	CV181040 Birkeland 1994	Nachteil	Vorteil	Kein Zusatznutzen
Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor				
24 Wochen	CV181040 Hermansen 2007 Garber 2008 Lewin 2012	Kein Unterschied	Kein Unterschied	Kein Zusatznutzen

Kein Unterschied: kein statistisch signifikanter Unterschied zwischen den Gruppen;

Nachteil: Statistisch signifikanter Unterschied zu Ungunsten von Sulfonylharnstoff+Saxagliptin

Vorteil: Statistisch signifikanter Unterschied zu Gunsten von Sulfonylharnstoff+Saxagliptin;

Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie:

Hinsichtlich der Änderung des HbA1c-Wertes ergab sich im indirekten Vergleich von Sulfonylharnstoff+Saxagliptin gegenüber Insulin ein Nachteil in der Blutzuckersenkung für Sulfonylharnstoff+Saxagliptin. Bezüglich der Hypoglykämien ergab sich ein Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin gegenüber Insulin, jedoch wurde der Endpunkt bei nur einer Studie mit einer geringen Patientenzahl adäquat berichtet. Insgesamt ergibt sich daher kein Zusatznutzen zugunsten einer der beiden Therapieoptionen.

Sulfonylharnstoff+Saxagliptin 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 keinen Vorteil für eine der beiden Behandlungsoptionen.

4.3.2.1.3.1.3 Gewichtsveränderung – 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**.

Tabelle 4-82: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff +Saxagliptin	Sulfonylharnstoff +Insulin	Insulin	Sulfonylharnstoff
1	CV181040	•			•
1	Shank 1995		•		•
4	Birkeland 1994 Tovi 1998 Nathan 1988 Turner 1998 (UKPDS 24)			• • • •	• • • •

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

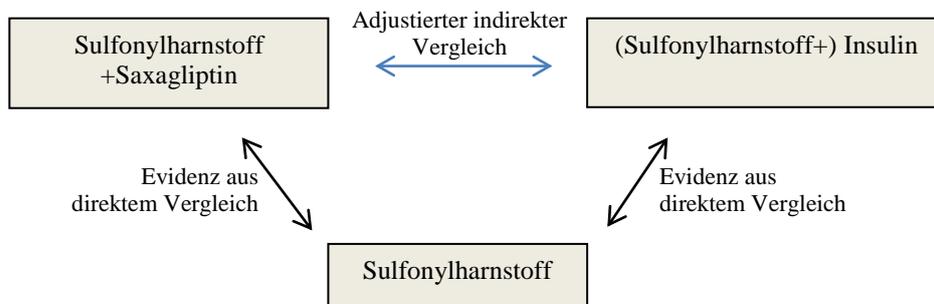
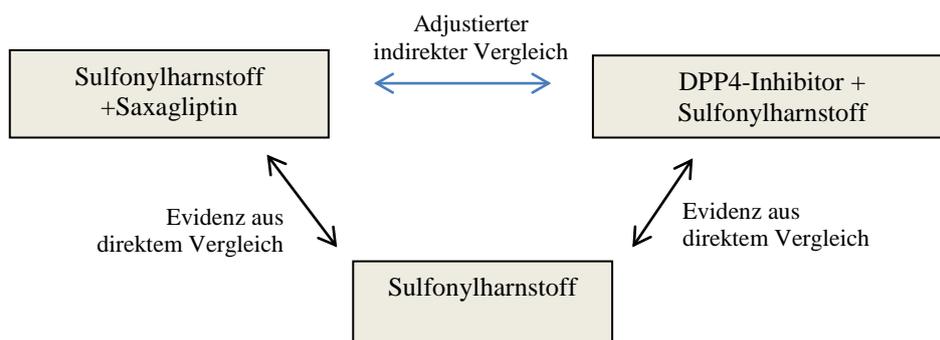


Tabelle 4-83: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff + Saxagliptin	Sulfonylharnstoff + DPP-4-Inhibitor	Sulfonylharnstoff + Placebo
1	CV181040	•		•
3	Hermansen et al 2007 Garber et al. 2008 Lewin 2012		• • •	• • •



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

Tabelle 4-84: Operationalisierung von Gewichtsveränderung

Studie	Operationalisierung
Studie CV181040	siehe Abschnitt 4.3.1.3.1.2
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin	
Shank 1995	Veränderung des Gewichtes nach 6 Monaten
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie	
Birkeland 1994	Veränderung des Gewichtes alle 3 Monate bis nach 12 Monaten
Tovi 1998	Veränderung des Gewichtes nach 6 und 12 Monaten
Nathan 1988	Veränderung des Gewichtes nach 9 Monaten
Turner 1998 (UKPDS 24)	Veränderung des Gewichtes nach 12 Monaten
Studien für indirekten Vergleich Sulfonylharnstoff +Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor	
Hermansen et al. 2007	Änderung des Gewichtes in kg von Studienbeginn bis Woche 24
Garber et al. 2008	Änderung des Gewichtes in kg von Studienbeginn bis Woche 24
Lewin 2012	Änderung des Gewichtes in kg von Studienbeginn bis Woche 1/8

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

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

Tabelle 4-85: 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 CV181040 siehe Abschnitt 4.3.1.3.1.2						
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin						
Shank 1995	hoch	unklar	nein**	unklar*	nein	hoch
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie						
Birkeland 1994	hoch	nein	nein**	unklar*	nein	hoch
Tovi 1998	niedrig	nein	nein**	unklar*	nein	hoch
Nathan 1995	hoch	nein	nein**	unklar*	nein	hoch
Turner 1998 (UKPDS 24)	hoch	nein	ja	unklar*	nein	hoch
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	niedrig	ja	ja	unklar*	nein	niedrig
Garber et al. 2008	niedrig	ja	ja	unklar*	nein	niedrig

*es liegt nur die Publikation der Studie vor (kein Studienbericht) ** aufgrund der geringen Fallzahl kann kein valides Ergebnis berechnet werden

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für den zu bewertenden Endpunkt Gewichtsveränderung für alle Insulin-Studien als „hoch“ bewertet, da aufgrund des hohen Verzerrungspotenzial auf Studienebene eine endpunktspezifische Verzerrung nicht ausgeschlossen werden kann. Auch die geringe Fallzahl in den Studien kann eine Verzerrung verursachen.

Für die DPP-4-Studien war der Endpunkterheber jeweils verblindet, das ITT-Prinzip wurde adäquat umgesetzt wurde und es gab auch keine sonstigen das Verzerrungspotenzial beeinflussenden Faktoren. 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 alle Studien mit der zweckmäßigen Vergleichstherapie 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 4-86: Ergebnisse für Gewichtsveränderung aus RCT für indirekte Vergleiche

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Saxagliptin				Komparator: Sulfonylharnstoff				Zweckmäßige Vergleichstherapie: Insulin bzw. alternative Vergleichstherapie DPP-4-Inhibitor				Gruppen- unterschied kg Mittelwert (SE)
	N	Ausgangs- wert kg Mittelwert (SE/SD)	End- wert kg Mittel- wert (SE/ SD)	Ver- änder- ung kg Mittel- wert (SE/SD)	N	Aus- gangs- wert kg Mittel- wert (SE/SD)	Endwert kg Mittel- wert (SE/SD)	Veränder- ung kg Mittelwert (SE/SD)	N	Aus- gangs- wert kg Mittel- wert (SE/SD)	End- wert kg Mittel- wert (SE/SD)	Ver- änder- ung kg Mittel- wert (SE/SD)	
CV181040													
24 Wochen	253	76,2 (1,11 ^a)	77,0 (1,10 ^a)	0,8 (0,13 ^a)	265	75,6 (1,07 ^a)	76,0 (1,09 ^a)	0,3 (0,13 ^a)					0,5 (0,19) [0,1;0,8]
50 Wochen	253	76,2 (1,11 ^a)	77,1 (1,11 ^a)	0,9 (0,15 ^a)	265	75,6 (1,07 ^a)	75,9 (1,09 ^a)	0,3 (0,14 ^a)					0,60 (0,2052) [0,20;1,00]*
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin													
Shank 1995 24 Wochen					10	79,2 (3,6 ^a)	79,8 (3,5 ^a)	0,6 (3,6 ^a)*	9	87,3 (5,5 ^a)	91,8 (5,8 ^a)	4,5(5,8 ^a)*	3,90 (6,8264) [-9,48;17,28]
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie													
Birkeland 1994 52 Wochen					16	77,1 (3,6 ^a)	76,6 (3,4 ^a)	-0,5 (3,6 ^a)*	18	75,3 (3,1 ^a)	79,2 (3,3 ^a)	3,9 (3,3 ^a)*	4,40 (4,8836) [-13,97;5,17] ^o
Tovi 1998 24 Wochen					16	78,3 (15,3 ^b)	n.b.	-1,2 [-2,3;-0,1]	19	72,9 (14,1 ^b)	n.b.	2,0 [0,4;3,5]	3,90 (0,7844) [2,36;5,44] ^o
52 Wochen					16	78,3	n.b.	-1,9	19	72,9	n.b.	2,9	4,80 (1.0817)

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

Studie	Zu bewertende Arzneimittel: Sulfonylharnstoff+Saxagliptin			Komparator: Sulfonylharnstoff			Zweckmäßige Vergleichstherapie: Insulin bzw. alternative Vergleichstherapie DPP-4-Inhibitor			Gruppen- unterschied kg Mittelwert (SE)				
	N	Ausgangs- wert kg Mittelwert (SE/SD)	End- wert kg Mittel- wert (SE/ SD)	Ver- änder- ung kg Mittel- wert (SE/SD)	N	Aus- gangs- wert kg Mittel- wert (SE/SD)	Endwert kg Mittel- wert (SE/SD)	Veränder- ung kg Mittelwert (SE/SD)	N		Aus- gangs- wert kg Mittel- wert (SE/SD)	End- wert kg Mittel- wert (SE/SD)	Ver- änder- ung kg Mittel- wert (SE/SD)	
						(15,3 ^b)		[-3,4;-0,5]			(14,1 ^b)		[1,0;4,7]	[2,68;6,92] ^o
Nathan 1988** 36 Wochen					16	89,3 (15 ^b)	93,1 (15 ^b)	3,8 (15 ^b)*	15	80,6 (16 ^b)	86,0 (16 ^b)	5,4 (16 ^b)*		1,6 (5,2475) [-8,68;11,88] ^o
Turner 1998** (UKPDS 24) 52 Wochen					231	71 (15 ^b)	75 (15 ^b)	4 (15 ^b)*	178	71 (16 ^b)	76 (17 ^b)	5 (17 ^b)*		1,00 (1,6117) [-2,16;4,16] ^o
Shank 1995 24 Wochen					10	79,2 (3,6 ^a)	79,8 (3,5 ^a)	0,6 (3,6 ^a)*	9	74,9 (2,6 ^a)	75,5 (2,9 ^a)	0,6 (2,9 ^a)*		0,00 (4,6228) [-9,06;9,06] ^o
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor														
Hermansen 2007					106	85,1 (22,6) ^b		0,0 (0,33 ^a)* [-0,6;0,7]	106	85,8 (22,5) ^b		1,1 (0,33 ^a)* [0,5;1,8]		1,10 (0,4667) [0,19;2,01] ^o
Garber 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] ^o
Lewin 2012					82	76,1 (17,0) ^b		-0,01 (17,0) ^b *	158	74,5 _b (17,0)		0,43 (17,0) ^b *		0,44 (2,3138) [-4,09;4,97] ^o

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

* Streuungsmaß geschätzt aus dem Mittelwert der Standardabweichungen der anderen Studien der Vergleichstherapie oder das größere Streuungsmaß der Vorher-Nachher-Messung, falls kein Mittelwert von mindestens 2 Studien berechenbar ist

** behandlungsnaiv

Insulin versus Sulfonylharnstoff+Insulin:

Beide Behandlungsgruppen in der Publikation von Shank 1995 zeigten eine geringe Gewichtszunahme (jeweils um 0,6 kg) während einer 6-monatigen Behandlungsphase.

Insulin versus Sulfonylharnstoff

In der Studie von Birkeland 1994 nahm die Insulin-Gruppe deutlich um 3,9 kg an Gewicht zu und die Sulfonylharnstoff-Gruppe um -0,5 kg ab.

In der Studie von Tovi 1998 nahmen die Patienten der Insulin-Gruppe zwar durchschnittlich 2,0 kg nach sechs Monaten und 2,9 kg nach zwölf Monaten zu und in der Sulfonylharnstoff-Gruppe durchschnittlich -1,2 kg nach sechs Monaten und -1,9 kg nach zwölf Monaten ab.

In der Studie von Nathan 1988 (behandlungsnaive Patienten) nahm die Insulin-Gruppe zwar etwas mehr an Gewicht zu (5,4 kg), aber im Vergleich zur Sulfonylharnstoff-Gruppe mit einer Gewichtszunahme von 3,8 kg ist dieser Effekt nicht signifikant unterschiedlich.

Turner 1998 (UKPDS 24) zeigte nach einem Jahr bei einem behandlungsnaiven Patientenkollektiv in beiden Behandlungsgruppen eine Gewichtszunahme von vier und fünf Kilogramm; es ergab sich aber kein signifikanter Unterschied zwischen den Gruppen.

Die Studie von Shank 1995 berichtete in den Gruppen Insulin und Sulfonylharnstoff die Gewichtveränderung nur nach sechs Monaten und kann daher in der Meta-Analyse nicht berücksichtigt werden. Nach sechs Monaten zeigte sich kein Vorteil bezüglich der Gewichtveränderung in einer der beiden Behandlungsoptionen.

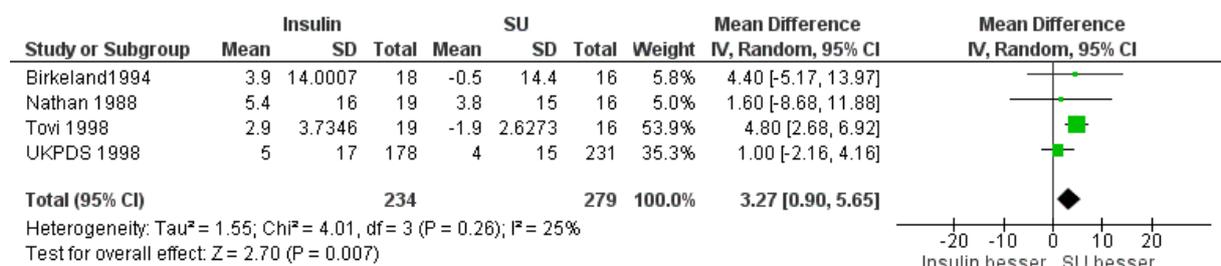


Abbildung 7: Meta-Analyse der eingeschlossenen Studien zum Effekt von Insulin vs. Sulfonylharnstoff auf den Endpunkt Gewichtsveränderung

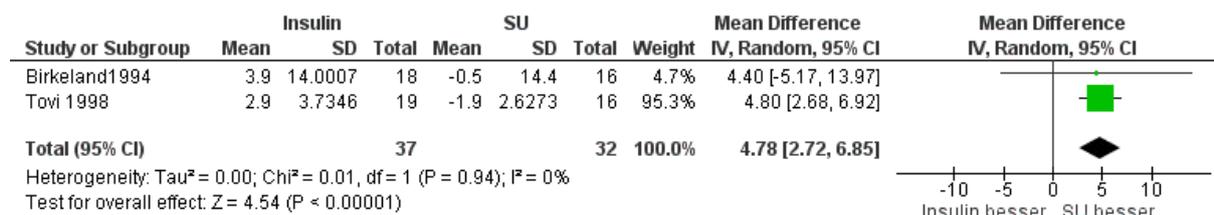


Abbildung 8: Meta-Analyse der eingeschlossenen Studien zum Effekt von Insulin vs. Sulfonylharnstoff auf den Endpunkt Gewichtsveränderung (vorbehandelte Patienten)

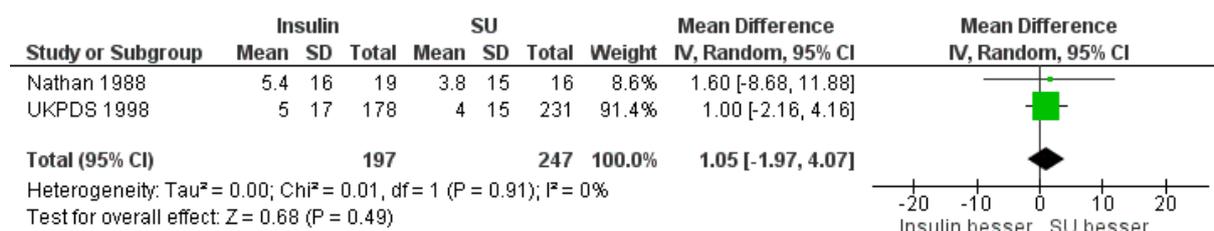


Abbildung 9: Meta-Analyse der eingeschlossenen Studien zum Effekt von Insulin vs. Sulfonylharnstoff auf den Endpunkt Gewichtsveränderung (behandlungsnaive Patienten)

Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff:

Für den Endpunkt Gewichtsveränderung standen zwei der identifizierten DPP-4-Studien des Vergleiches Sulfonylharnstoff+Placebo versus Sulfonylharnstoff+DPP-4-Inhibitor zur Verfügung. Die Studien von Garber 2008 und Hermansen 2007 berichteten die Veränderungen inklusive Streuungsmaß wohingegen die Publikation von Lewin 2012 keine Standardabweichung in der Publikation berichtete. Die Publikationen wiesen in Bezug auf Veränderungen des Körpergewichts auf einen Nachteil von DPP-4-Inhibitoren+Sulfonylharnstoff hin. Der beobachtete Gruppenunterschied war aber nur in der Studie von Hermansen 2007 statistisch signifikant. Mit einem Wert von I²=38% weisen die beiden Einzelergebnisse auf eine moderate Heterogenität hin, daher muss man mit der Interpretation des gemeinsamen Effektschätzers 0,68 [-0,11;1,46] vorsichtig sein.

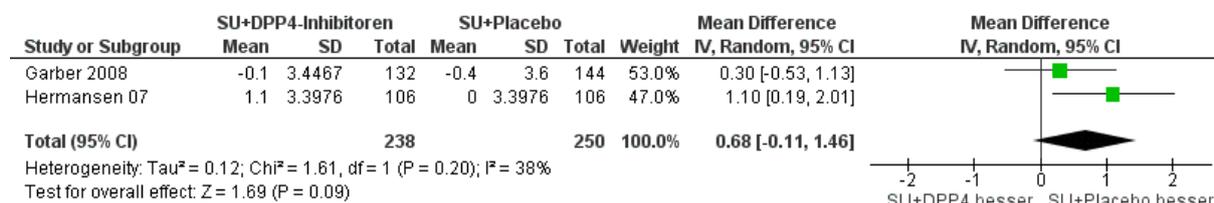


Abbildung 10: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt Gewichtsveränderung

Zusätzlich wurde die Publikation von Lewin 2012 (die Standardabweichungen sind nicht angegeben, so wurde das Effektmaß geschätzt) noch mit in die Meta-Analyse aufgenommen.

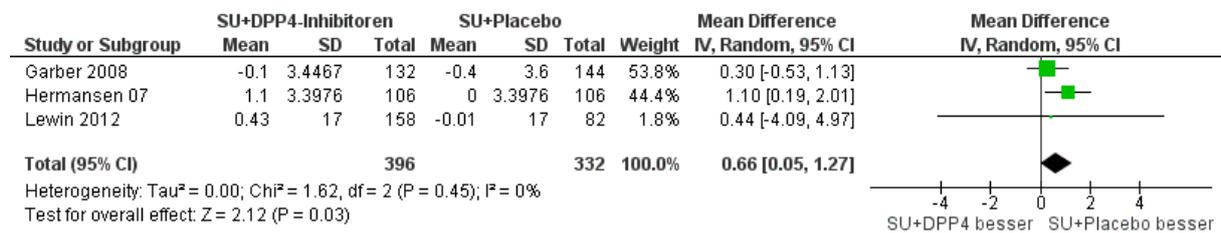


Abbildung 11: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt Gewichtsveränderung

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.

Ergebnisse des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+Insulin:

Tabelle 4-87: Ergebnisse für Gewichtsveränderung - indirekte Vergleiche

Mittelwertdifferenz des direkten Vergleiches der Studie CV181040 nach 24 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Mittelwertdifferenz des direkten Vergleiches (Sulfonylharnstoff+Insulin versus Sulfonylharnstoff)	Mittelwertdifferenz des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+Insulin kg [95%-KI]
0,5 (0,19)	3,90 (6,8264) ^a	-3,40 [-16,78;9,98]

^a Effektschätzer von Shank 1995; KI: Konfidenzintervall

Auch der indirekte Vergleich Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+Insulin ergab keinen statistisch signifikanten Vorteil zugunsten einer der beiden Behandlungsoptionen

Ergebnisse des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Insulin

Tabelle 4-88: Ergebnisse für Gewichtsveränderung - indirekte Vergleiche

Mittelwertdifferenz des direkten Vergleiches der Studie CV181040 nach 50 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Mittelwertdifferenz des direkten Vergleiches (Insulin versus Sulfonylharnstoff) nach 52 Wochen	Mittelwertdifferenz des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Insulin kg [95%-KI]
0,60 (0,2052)	4,78 (1,0542) ^a	-4,18 [-6,29;-2,07]
0,60 (0,2052)	1,05 (1,5407) ^b	-0,45 [-3,50;2,59]
0,60 (0,2052)	3,27 (0,9011) ^c	-2,67 [-4,48;-0,86]

^a Gepoolter Effektschätzer der Meta-Analyse von Birkeland 94 und Tovi 98 (vorbehandelte Patienten)^b Gepoolte Effektschätzer der Meta-Analyse von Nathan 88 und UKPDS 98 (behandlungsnaive Patienten)^c Gesamtschätzer von Birkeland 94, Tovi 98, Nathan 88 und UKPDS 98

KI: Konfidenzintervall

Der indirekte Vergleich Sulfonylharnstoffe+Saxagliptin versus Insulin zeigte bei den vorbehandelten Patienten ein signifikantes Ergebnis zugunsten von Sulfonylharnstoff+Saxagliptin

Ergebnisse des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor

Tabelle 4-89: Ergebnisse für Gewichtsveränderung - indirekte Vergleiche

Mittelwertdifferenz des direkten Vergleiches der Studie CV181040 nach 24 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Mittelwertdifferenz des direkten Vergleiches (Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff)	Mittelwertdifferenz des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor kg [95%-KI]
0,5 (0,19)	0,68 (0,3993) ^a	-0,18 [-1,04;0,69]
0,5 (0,19)	0,66 (0,3111) ^b	-0,16 [-0,87; 0,56]
0,5 (0,19)	1,10 (0,4667) ^c	-0,60 [-1,59;0,39]

^a Gepoolter Effektschätzer der Meta-Analyse von Hermansen 07, Garber 08^b Gepoolter Effektschätzer der Meta-Analyse von Hermansen 07, Garber 08 und Lewin 2012^c Effektschätzer von Hermansen 2007 (Sitagliptin als Leitsubstanz für die DPP-4-Inhibitoren)

KI: Konfidenzintervall

Der indirekte Vergleich von Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor zeigte keinen signifikanten Unterschied zugunsten einer der beiden Behandlungsoptionen, gleiches galt für den indirekten Vergleich gegen die Leitsubstanz Sitagliptin.

4.3.2.1.3.1.4 Unerwünschte Ereignisse – 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.

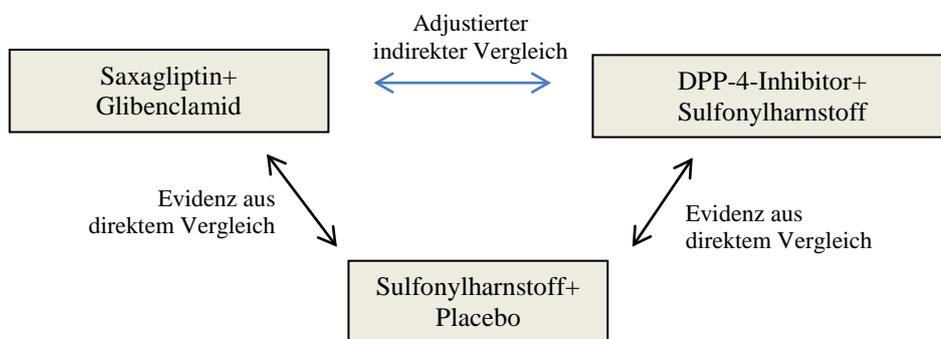
Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Ein indirekter Vergleich zwischen Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+Insulin bzw. Insulin kann aufgrund fehlender Daten nicht erstellt werden.

Tabelle 4-90: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich (Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor) herangezogen wurden

Anzahl Studien	Referenzen Studien	Sulfonylharnstoff+Saxagliptin	Sulfonylharnstoff+DPP-4-Inhibitor	Sulfonylharnstoff+Placebo
1	CV181040	•		•
3	Hermansen et al 2007 Garber et al. 2008 Lewin 2012		• • •	• • •

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 4-91: Operationalisierung von Unerwünschte Ereignisse

Studie	Operationalisierung
Studie CV181040 siehe Abschnitt 4.3.1.3.1.2	
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+Insulin	
Shank 1995	Nicht berichtet
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Insulin-Monotherapie	
Birkeland 1994	Nicht berichtet
Tovi 1998	Nicht berichtet
Nathan 1988	Nicht berichtet
Turner 1998 (UKPDS 24)	Nicht berichtet
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor	
Hermansen et al. 2007	Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis oder einer Laborveränderung (innerhalb der Behandlungsphase) Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase) Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrachen (innerhalb der Behandlungsphase)
Garber et al. 2008	Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase) Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase) Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrachen (innerhalb der Behandlungsphase)
Lewin 2012	Gesamtrate unerwünschter Ereignisse: Anzahl der Patienten mit mindestens einem unerwünschten Ereignis (innerhalb der Behandlungsphase) Schwerwiegende unerwünschte Ereignisse: Anzahl der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (innerhalb der Behandlungsphase) Studienabbruch wegen unerwünschter Ereignisse: Anzahl Patienten, die aufgrund eines unerwünschten Ereignisses die Studie abbrachen (innerhalb der Behandlungsphase)

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

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

Tabelle 4-92: 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 CV181040 siehe Abschnitt 4.3.1.3.1.2						
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor						
Hermansen et al. 2007	niedrig	ja	ja	unklar*	nein	niedrig
Garber et al. 2008	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.

Bezüglich der Vergleichbarkeit der Studien für die Zwecke eines indirekten Vergleiches muss berücksichtigt werden, dass keine Angabe gemacht wurde, ob Hypoglykämien in die Gesamtzahl der unerwünschten Ereignisse eingeschlossen waren oder nicht (Garber 2008; Hermansen 2007; Lewin 2012). Auch sind in der Publikation von Hermansen 2007 Laborveränderung als unerwünschtes Ereignis berichtet worden. Daher 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 zusammenfassend.

Tabelle 4-93: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel

Studie	N	Gesamtrate unerwünschter Ereignisse, n (%)	Schwerwiegende unerwünschte Ereignisse, n (%)	Studienabbruch aufgrund unerwünschter Ereignisse, n (%)
CV181040				
Nach 24 Wochen				
Sulfonylharnstoff+Saxagliptin 5 mg	253	183 (72,3)	6 (2,4)	8 (3,2)
Sulfonylharnstoff+Placebo 5 mg: OR [95%-KI]	267	205 (76,8) 0,79 [0,53;1,17]	6 (2,2) 1,06 [0,34;3,32]	4 (1,5) 2,15 [0,64;7,22]
Nach 76 Wochen				
Sulfonylharnstoff+Saxagliptin 5 mg	253	215 (85,0)	12 (4,6)	16 (6,3)
Sulfonylharnstoff+Placebo 5 mg: OR [95%-KI]	267	237 (88,8) 0,72 [0,43;1,20]	23 (8,6) 0,53 [0,26;1,09]	12 (4,5) 1,43 [0,66;3,10]
Studien für indirekten Vergleich Sulfonylharnstoff+Saxagliptin vs. Sulfonylharnstoff+DPP-4-Inhibitor				
Hermansen et al. 2007				
Sulfonylharnstoff+DPP-4-Inhibitor	106	59 (55,7)	5 (4,7)	3 (2,8)
Sulfonylharnstoff	106	43 (40,6)	6 (5,7)	1 (0,9)
OR [95%-KI]		1,84 [1,07;3,17]	0,83 [0,24;2,79]	3,06 [0,31;29,88]
Garber et al. 2008				
Sulfonylharnstoff+DPP-4-Inhibitor	170	114 (67,1)	5 (2,9)	4 (2,4)
Sulfonylharnstoff	176	113 (64,2)	9 (5,1)	3 (1,7)
OR [95%-KI]		1,13 [0,73;1,77]	0,56 [0,18;1,71]	1,39 [0,31;6,3]
Lewin 2012				
Sulfonylharnstoff+DPP-4-Inhibitor	161	68 (42,2)	5 (3,11)	5 (3,11)
Sulfonylharnstoff	84	36 (42,9)	1 (1,19)	3 (3,6)
OR [95%-KI]		0,97 [0,57;1,66]	2,66 [0,31;23,15]	

KI: Konfidenzintervall

Sulfonylharnstoff+Insulin oder Insulin-Monotherapie versus Sulfonylharnstoff:

Keine der Studien, identifiziert für die zweckmäßige Vergleichstherapie berichtet die Gesamtrate unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse oder die Studienabbrüche aufgrund unerwünschter Ereignisse.

Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff:

Die Anzahl unerwünschter Ereignisse war nur in der Studie von Hermansen 2007 in der Sulfonylharnstoff+DPP-4-Inhibitor-Gruppe höher als in der Sulfonylgruppe+Placebo (OR [95%-KI]: 1,84 [1,07;3,17]). Des Weiteren fanden sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen.

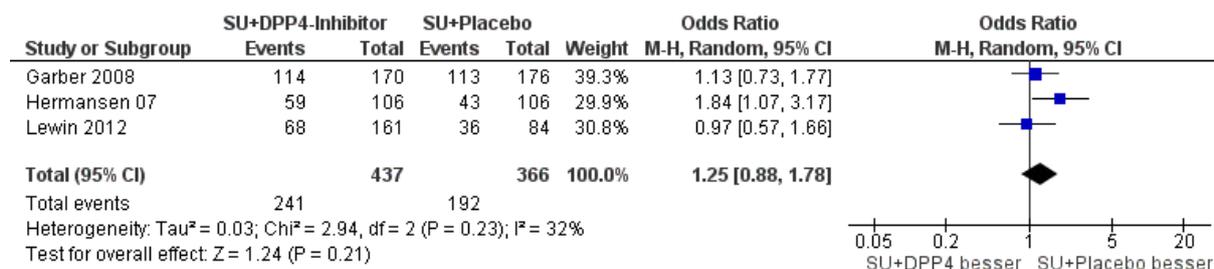


Abbildung 12: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate unerwünschter Ereignisse“

Da es gegensätzliche Einzeleffekte der Studien in der Meta-Analyse gibt und eine moderate Heterogenität ($I^2=32\%$; $p=0,23$) besteht, ist nur bedingt sinnvoll, den gemeinsamen Effektschätzer für unerwünschte Ereignisse zu betrachten. Der Effektschätzer konnte keinen signifikanten Unterschied in den Behandlungsgruppen entdecken.

Bezüglich der schwerwiegenden unerwünschten Ereignisse gab es in den einzelnen Studien keine signifikanten Unterschiede.

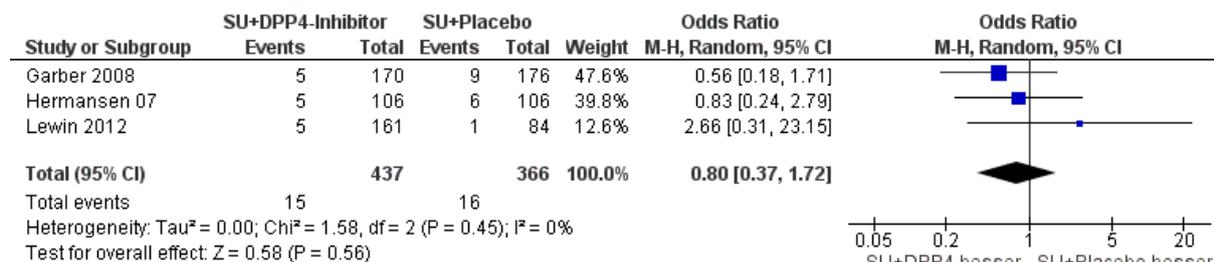


Abbildung 13: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Gesamtrate schwerwiegender unerwünschter Ereignisse“

Auch bei der Betrachtung der schwerwiegenden unerwünschten Ereignisse in einer Meta-Analyse besteht kein Hinweis auf Heterogenität ($I^2=0\%$; $p=0,45$). Jedoch sind die Einzeleffekte gegensätzlich und der gemeinsame Effektschätzer von 0,8 nicht signifikant. Daraus folgt, dass es keine signifikanten Unterschiede zwischen den zwei Behandlungsgruppen gibt.

Studienabbrüche aufgrund unerwünschter Ereignisse berichteten ebenfalls alle Publikationen.

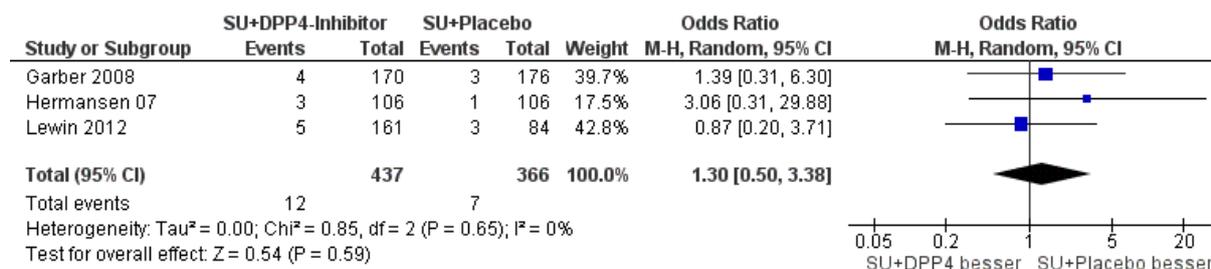


Abbildung 14: Meta-Analyse der eingeschlossenen Studien zum Effekt von Sulfonylharnstoff+DPP-4-Inhibitor vs. Sulfonylharnstoff+Placebo auf den Endpunkt „Studienabbrüche aufgrund unerwünschter Ereignisse“

Die Meta-Analyse bezüglich der Studienabbrüche aufgrund unerwünschter Ereignisse weist keine Heterogenität auf ($I^2=0\%$; $p=0,57$). Zudem sind die Einzel- und der gemeinsame Effektschätzer nicht signifikant. Trotzdem lässt sich ein leichter Trend zu mehr Studienabbrüchen in der Sulfonylharnstoff-Gruppe erkennen.

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 4-94: Ergebnisse für Gesamtrate unerwünschter Ereignisse –indirekter Vergleich

Odds Ratio des direkten Vergleiches der Studie CV181040 nach 24 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Odds Ratio des direkten Vergleiches (Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff)	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor [95%-KI]
0,79 (0,2019)	1,25 (0,1803) ^a	0,36 [0,37;1,07]
0,79 (0,2019)	1,84 (0,2781) ^b	0,43 [0,22;0,84]

^a Gepoolter Effektschätzer der Meta-Analyse von Hermansen 2007, Garber 2008 und Lewin 2012

^b Effektschätzer von Hermansen 2007 (Sitagliptin als Leitsubstanz für die DPP4-Inhibitoren)

KI: Konfidenzintervall

Der indirekte Vergleich in Bezug auf den Endpunkt „Gesamtrate unerwünschter Ereignisse“ ergab keinen statistisch signifikanten Unterschied zwischen der zu bewertenden Therapie und der alternativen Vergleichstherapie (DPP-4-Inhibitoren als Gruppe). Jedoch zeigte der indirekte Vergleich Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+Sitagliptin (als Einzelsubstanz) einen signifikanten Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin. Dieser Unterschied wird aber in Frage gestellt, da der Endpunkt in der Studie von Hermansen 2007 sehr wahrscheinlich anders operationalisiert wurde als in der Studie CV181040.

Tabelle 4-95: Ergebnisse für schwerwiegende unerwünschte Ereignisse –indirekter Vergleich

Odds Ratio des indirekten Vergleiches der Studie CV181040 nach 24 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Odds Ratio des direkten Vergleiches (Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff)	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor [95%-KI]
1,06 (0,5841)	0,80 (0,3921) ^a	1,33 [0,33;5,26]
1,06 (0,5841)	0,83 (0,6217) ^b	1,28 [0,24;6,82]

^a Gepoolter Effektschätzer der Meta-Analyse von Hermansen 2007, Garber 2008 und Lewin 2012

^b Effektschätzer von Hermansen 2007 (Sitagliptin als Leitsubstanz für die DPP-4-Inhibitoren)

KI: Konfidenzintervall

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

Tabelle 4-96: Ergebnisse für Studienabbrüche aufgrund unerwünschter Ereignisse –indirekter Vergleich

Odds Ratio des indirekten Vergleiches der Studie CV181040 nach 24 Wochen (Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff)	Odds Ratio des direkten Vergleiches (Sulfonylharnstoff+DPP-4-Inhibitor versus Sulfonylharnstoff)	Odds Ratio des indirekten Vergleiches Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+DPP-4-Inhibitor [95%-KI]
2,15 (0,6188)	1,3023 (0,4862) ^a	1,65 [0,35;7,71]
2,15 (0,6188)	3,0583 (1,1630) ^b	0,70 [0,05;9,28]

^a Gepoolter Effektschätzer der Meta-Analyse von Hermansen 2007, Garber 2008 und Lewin 2012

^b Effektschätzer von Hermansen 2007 (Sitagliptin als Leitsubstanz für die DPP4-Inhibitoren)

KI: Konfidenzintervall

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

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 4.3.1.2.1.

Es konnten keine Subgruppenanalysen auf Basis indirekter Vergleiche vorgelegt werden, da für die im indirekten Vergleich verwendeten Studien für die Vergleichstherapie 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

Nicht zutreffend.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-97: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Paralleltät 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 4-98: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-99: 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.

Nicht zutreffend.

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 4.3.1.2.1.

Nicht zutreffend.

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*

Nicht zutreffend.

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.

Nicht zutreffend.

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 4-100: 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.

Nicht zutreffend.

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.

Nicht zutreffend.

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

4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.2.1.

Nicht zutreffend.

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 Saxagliptin wurde ein indirekter Vergleich herangezogen, da für das Anwendungsgebiet Saxagliptin Add-on zum Sulfonylharnstoff keine Studien vorlagen, die einen direkten Vergleich zwischen dem zu bewertenden Arzneimittel und der zweckmäßigen Vergleichstherapie ermöglichen. Als zweckmäßige Vergleichstherapie wurde Sulfonylharnstoff+Insulin oder Insulin-Monotherapie genannt. Zudem wurde ein indirekter Vergleich für die alternative Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitoren aufgestellt.

Anhand folgender Endpunkte wurde die Wirksamkeit und Sicherheit von Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+Insulin oder Insulin-Monotherapie bzw. Sulfonylharnstoff+DPP-4-Inhibitoren gegenübergestellt:

Ausmaß der Blutzuckerkontrolle:

Das Ausmaß der Blutzuckerkontrolle konnte für den indirekten Vergleich zu Sulfonylharnstoff+Insulin nicht berechnet werden, da die Studie von Shank 1995 keinen HbA1c-Ausgangswert oder eine Veränderung des HbA1c vom Studienbeginn bis Woche 24 berichtet.

Für den indirekten Vergleich mit der zweckmäßigen Vergleichstherapie Insulin konnten die Studien von Shank 1995 und Turner 1998 aufgrund fehlender Angaben (weder ein HbA1c-Ausgangswert noch eine Veränderung des HbA1c-Wertes bzw. Mittelwerte oder Standardabweichungen) nicht herangezogen werden. Die Studien von Birkeland 1994 und Tovi 1998 wurden aufgrund ihres ähnlichen Studiendesigns in einer Meta-Analyse berücksichtigt. Im indirekten Vergleich zur Studie CV181040 und dem Gesamtschätzer der Meta-Analyse zeigte sich ein signifikantes Ergebnis zugunsten der Insulin-Behandlungsgruppe nach 50 Wochen (MWD [95%-KI]: 1,41% [0,70;2,13]). Die Studie von Nathan 1988 wurde nicht in die Meta-Analyse eingeschlossen, da es sich bei der Patientenpopulation um behandlungsnaive Patienten handelt. Betrachtet man ungeachtet des unterschiedlichen Patientenkollektivs den indirekten Vergleich mit der Studie CV181040, zeigte sich kein Unterschied im Bezug auf die Senkung des HbA1c-Wertes (MWD [95%-KI]: -0,34% [-1,75;1,07]).

Der indirekte Vergleich mit der alternativen Vergleichstherapie Sulfonylharnstoff+anderer DPP-4-Inhibitor konnte mit dem Gesamtschätzer der Meta-Analyse (Studien von Hermansen 2007, Garber 2008 und Lewin 2012) sowie der Studie CV181040 durchgeführt werden. Es ergab sich für den Endpunkt „Ausmaß der Blutzuckerkontrolle“ weder für den Vergleich mit Sitagliptin als Einzelsubstanz noch mit allen DPP-4-Inhibitoren als Gruppe Add-on Sulfonylharnstoff ein statistisch signifikanter Unterschied zwischen den Behandlungsoptionen (MWD [95%-KI]: -0,15% [-0,45;0,15] bzw. -0,16% [-1,55;1,23]).

Hypoglykämien:

Der Endpunkt Hypoglykämien wurde in der Studie von Shank 1995 nicht berichtet, so dass ein indirekter Vergleich mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Insulin nicht möglich war.

Für den indirekten Vergleich mit Insulin konnte die Studie von Birkeland 1994 herangezogen werden. Es zeigte sich ein signifikantes Ergebnis nach 50 Wochen zugunsten von Sulfonylharnstoff+Saxagliptin gegenüber Insulin (OR [95%-KI]: 0,04 [0,00;0,88]). Das Ergebnis sollte jedoch wegen der geringen Studienpopulation in der Studie von Birkeland 1994 (n=34) mit Vorsicht interpretiert werden. Die Studien von Tovi 1998, Nathan 1988 und Turner 1998 konnten aufgrund der unterschiedlichen Operationalisierung des Endpunktes Hypoglykämien nicht mit der Studie CV181040 verglichen werden.

Für die alternative Vergleichstherapie Sulfonylharnstoff+anderer DPP-4-Inhibitor ergab sich weder in Bezug auf die patientenberichteten Hypoglykämien (Sitagliptin (OR [95%-KI]: 0,54 [0,13;2,33]) bzw. DPP-4-Inhibitoren als Gruppe (OR [95%-KI]: 0,88 [0,31;2,50]),

noch in Bezug auf die bestätigten Hypoglykämien (OR [95%-KI]: 0,51 [0,02;11,38]) ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien:

Die gemeinsame Betrachtung der Blutzuckerkontrolle und der Hypoglykämien war für die Vergleichstherapie Sulfonylharnstoff+Insulin nicht möglich.

Der indirekte Vergleich von Sulfonylharnstoff+Saxagliptin mit Insulin ergab einen Nachteil in der Blutzuckersenkung für Saxagliptin in Kombination mit Sulfonylharnstoff. Dagegen zeigte sich ein Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin gegenüber Insulin in Bezug auf die Anzahl der Hypoglykämien. Insgesamt ergibt sich somit kein Zusatznutzen zugunsten einer Behandlungsoption in der gemeinsamen Betrachtung der Blutzuckerkontrolle und der Hypoglykämien.

Weder im Bezug auf die Blutzuckersenkung noch auf die Anzahl der Hypoglykämien konnte ein statistisch signifikanter Unterschied zwischen Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+Sitagliptin bzw. Sulfonylharnstoff+andere DPP-4-Inhibitoren als Gruppe gezeigt werden. Bei der gemeinsamen Betrachtung der Blutzuckerkontrolle und der Hypoglykämien ergibt sich somit kein Vorteil für eine der Therapiemöglichkeiten.

Gewichtsveränderung:

Für den Endpunkt Gewichtsveränderung konnte im indirekten Vergleich von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Insulin nach 24 Wochen kein statistisch signifikanter Vorteil zugunsten einer der beiden Behandlungsoptionen ermittelt werden (MWD [95%-KI]: -3,40 kg [-16,78;9,98]).

Für den indirekten Vergleich von Sulfonylharnstoff+Saxagliptin und Insulin wurden die Studien von Birkeland 1994 und Tovi 1998 aufgrund des ähnlichen Studiendesigns in einer Meta-Analyse zusammengefasst und der gepoolte Effektschätzer wurde mit der Studie CV181040 verglichen. Das Ergebnis zeigte nach 50 Wochen keinen statistisch signifikanten Unterschied (MWD [95%-KI]: 0,43 kg [-1,95;2,81]). In den Studien von Nathan 1988 und Turner 1998 werden behandlungsnaive Patienten untersucht. Im Vergleich der Meta-Analyse dieser Studien zur Studie CV181040 ergab sich kein statistisch signifikanter Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin (MWD [95%-KI]: -0,45 kg [-3,50;2,59]). Fasst man alle Studien in einer Meta-Analyse zusammen, zeigte sich im indirekten Vergleich des Gesamtschätzers mit der Studie CV181040 kein signifikanter Unterschied zugunsten einer Behandlungsgruppe (MWD [95%-KI]: 0,10 kg [-1,80;1,99]).

Der indirekte Vergleich von Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+andere DPP-4-Inhibitoren zeigte keinen signifikanten Unterschied zwischen den Behandlungsoptionen (MWD [95%-KI]: -0,60 kg [-1,59;0,39] für Sulfonylharnstoff+Sitagliptin, bzw. MWD [95%-KI]: -0,18 kg [-1,04;0,69] für Sulfonylharnstoff+andere DPP-4-Inhibitoren als Gruppe).

Unerwünschte Ereignisse:

Der Endpunkt „unerwünschte Ereignisse“ wurde weder in den Studien zur zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Insulin noch in den Studien zur Insulin-Monotherapie berichtet.

Der indirekte Vergleich von Sulfonylharnstoff+Saxagliptin mit Sulfonylharnstoff+Sitagliptin (als Einzelsubstanz) in Bezug auf den Endpunkt Gesamtrate unerwünschter Ereignisse ergab einen signifikanten Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin (OR [95%-KI]: 0,43 [0,22;0,84]). Der Vergleich zu DPP-4-Inhibitoren als Gruppe ergab keinen statistisch signifikanten Unterschied (OR [95%-KI]: 0,36 [0,37;1,07]). Im Bezug auf den Endpunkt Gesamtrate schwerwiegender unerwünschter Ereignisse ergab sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (OR [95%-KI]: 1,28 [0,24;6,82] für Sulfonylharnstoff+Sitagliptin, bzw. OR [95%-KI]: 1,33 [0,33;5,26] für Sulfonylharnstoff+DPP-4-Inhibitoren als Gruppe). Ferner konnte kein statistisch signifikanter Unterschied für den Endpunkt Studienabbrüche aufgrund unerwünschter Ereignisse gezeigt werden (OR [95%-KI]: 0,70 [0,05;9,28] für Sulfonylharnstoff+Sitagliptin, bzw. OR [95%-KI]: 1,65 [0,35;7,71] für Sulfonylharnstoff+DPP-4-Inhibitoren als Gruppe).

Bezüglich der Gesamtrate unerwünschter Ereignisse ergab sich somit ein Zusatznutzen für Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Sitagliptin. Im Vergleich zu DPP-4-Inhibitoren als Gruppe für die Endpunkte „Gesamtrate unerwünschter Ereignisse“, „schwerwiegende unerwünschte Ereignisse“ und „Studienabbruch aufgrund unerwünschter Ereignisse“ zeigte sich weder ein Über- noch eine Unterlegenheit für eine der beiden Behandlungsoptionen.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens**4.4.1 Beurteilung der Aussagekraft der Nachweise**

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Zur Beantwortung der primären Fragestellung lag keine direkt vergleichende Studie vor, daher wurden adjustierte indirekte Vergleiche Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+Insulin oder Insulin-Monotherapie durchgeführt.

Die Studie CV10140 mit dem Vergleich Sulfonylharnstoff+Saxagliptin versus Sulfonylharnstoff+Insulin ist eine randomisierte, kontrollierte Studie und entspricht demnach Evidenzstufe Ib. Nach Prüfung aller möglichen Verzerrungsaspekte ist sowohl auf Studienebene als auch auf Endpunktebene von einem niedrigen Verzerrungspotenzial auszugehen. Die betrachteten Endpunkte sind in ihren Operationalisierungen valide, wiesen in den Studien ein geringes endpunktspezifisches Verzerrungspotenzial auf und stellen, wie in Abschnitt 4.3.1.2.2 begründet, patientenrelevante Endpunkte dar. Die Validität der herangezogenen Endpunkte ist somit insgesamt als hoch einzuschätzen:

Um eine gleichwertige Beweissicherheit auch für die Nachweise zur zweckmäßigen Vergleichstherapie zu erzielen, wurden für die zweckmäßige Vergleichstherapie ebenfalls nur RCT gesucht. Hierzu konnte für die zweckmäßige Vergleichstherapie Sulfonylharnstoff+Insulin nur eine RCT selektiert werden. Diese erwies sich als eine doppelblinde Studie mit einer geringen Fallzahl, so dass eine Verzerrung nicht ausgeschlossen werden konnte.

Für die **zweckmäßige Vergleichstherapie** Insulin wurden insgesamt fünf RCT identifiziert, von denen zwei nur behandlungsnaive Patienten einschlossen und somit nur eingeschränkt für den indirekten Vergleich geeignet waren. Bei den weiteren Studien konnte aufgrund der geringen Fallzahl und des oft offenen Studiendesigns ebenfalls eine Verzerrung nicht ausgeschlossen werden.

Einschränkungen in Bezug auf die Aussagekraft der Nachweise ergaben sich für die Studien zur zweckmäßigen Vergleichstherapie durch eine zum Teil hohe Heterogenität der Studien, einen offenen Studiendesign und die geringe Fallzahl in einigen Studien.

Für alle diskutierten Endpunkte können deswegen gegenüber der zweckmäßigen Vergleichstherapie, im Falle statistisch signifikanter Ergebnisse in den adjustierten indirekten Vergleichen, Aussagen gemacht werden, die maximal als Anhaltspunkte gewertet werden können.

Zur Beantwortung der sekundären Fragestellung für den Vergleich mit Sulfonylharnstoff+Saxagliptin und der alternativen Vergleichstherapie Sulfonylharnstoff+andere DPP-4-Inhibitoren lag ebenfalls keine direkt vergleichende Studie vor, daher wurden adjustierte indirekte Vergleiche durchgeführt. Für die alternative Vergleichstherapie wurden insgesamt drei Studien identifiziert, die alle ein niedriges Verzerrungspotenzial auf Studienebene aufwiesen und aufgrund ihres Designs (doppelblinde, randomisierte, kontrollierte Studien) eine hohe Evidenzstufe (Ib) hatten. Diese Studien waren im Studiendesign, in den Operationalisierungen der Endpunkte und in ihrer Studienpopulation homogen, so dass sie in einer Meta-Analyse zusammengefasst werden konnten. Daher konnte für jeden Endpunkt ein indirekter Vergleich mit der niedrig verzerrten Studie CV108040 des zu bewertenden Arzneimittels und den Studien zur alternativen Vergleichstherapie, bei denen für jeden Endpunkt mindestens zwei Studien mittels einer Meta-Analyse zusammengefasst werden konnte, durchgeführt werden.

Infolgedessen können für alle diskutierten Endpunkte, im Falle statistisch signifikanter Ergebnisse in den adjustierten indirekten Vergleichen, Aussagen gemacht werden, die 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, 2011). 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 und für die die Anwendung von Metformin ungeeignet erscheint.

Im Dossier betrachtete patientenrelevante Endpunkte sind das Auftreten von Hypoglykämien, Gewichtsveränderungen, unerwünschte Ereignisse sowie die kardialen Ereignisse. Darüber hinaus wurden Angaben zum HbA1c-Wert als Maß für die langfristige Blutzuckersenkung extrahiert und dargestellt. Diese Angaben dienen zur Interpretation der Ergebnisse hinsichtlich patientenrelevanter Endpunkte, insbesondere bezüglich des Auftretens von Hypoglykämien. Hierfür wurde auch der Composite-Endpunkt „Erreichen des HbA1c-Zielwerts (<6,5%/<7%) ohne Auftreten von Hypoglykämien“ hinzugezogen.

Obwohl der Schwerpunkt dieses Abschnitts auf der Beschreibung des Zusatznutzens liegt, soll der Nutzen von Sulfonylharnstoff+Saxagliptin ebenfalls dargestellt werden. Die beschriebenen Ergebnisse beziehen sich dabei jeweils auf die Saxagliptin 5 mg-Gruppe.

Nutzen von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Placebo

Die Darstellung des medizinischen Nutzens von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Placebo erfolgte mittels der Methodik der evidenzbasierten Medizin und mittels patientenrelevanter Endpunkte. Saxagliptin 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, 2009).

Die Ergebnisse zum Nutzen des zu bewertenden Arzneimittels Saxagliptin in Kombination mit Sulfonylharnstoffen basieren auf einer RCT der Phase III (CV181040), die mit der Kombination von Glibenclamid und Saxagliptin durchgeführt wurde.

Eine optimierte Blutzuckerkontrolle (und damit die HbA1c-Einstellung), wird von der DDG empfohlen (Zielvorgabe HbA1c <6,5% ohne das vermehrte Auftreten von Hypoglykämien und Gewichtszunahmen). Kann dieser Zielwert unter Monotherapie mit einem OAD nicht erreicht werden, wird die Kombinationstherapie mit einem anderen OAD bzw. einem GLP-1-Agonisten empfohlen (Matthaei et al., 2009). Der aktuelle Entwurf der 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 großen klinischen Studie UKPDS wurde gezeigt, dass sich bereits bei einer Senkung des HbA1c um 1% das Risiko mikrovaskulärer Erkrankungen um 37% und das Risiko makrovaskulärer Erkrankungen um 12-16% verringert (Stratton et al., 2000). In der ADVANCE Studie bewirkte die intensiviertere antihyperglykämische Therapie eine signifikante Reduktion des Neuauftretens bzw. der Progression der Nephropathie und des kombinierten Endpunktes „mikro- und makrovaskuläre Erkrankungen“ (Patel et al., 2008). Eine optimale Blutzuckereinstellung stellt somit ein wichtiges Therapieziel dar, insbesondere wenn sie unter Vermeidung von Nebenwirkungen wie z. B. Hypoglykämien erreicht werden kann.

Für das **Ausmaß der Blutzuckerkontrolle** zeigte die Studie CV181040 sowohl nach 24 als auch nach 76 Wochen (LOCF) statistisch signifikante (MWD [95%-KI]: -0,72% [-0,88;-0,56] / -0,56% [-0,72;-0,39]) und klinisch relevante Unterschiede zwischen Glibenclamid+Saxagliptin und Glibenclamid+Placebo. Auch der Anteil von Patienten, die aufgrund unzureichender Blutzuckerkontrolle die Studie abgebrochen haben oder eine Notfallmedikation erhielten, ergab in der Glibenclamid+Saxagliptin-Gruppe gegenüber der Glibenclamid+Placebo-Gruppe nach 24 und 76 Wochen signifikant geringere Raten (OR [95%-KI] nach 24 Wochen 0,47 [0,31;0,72] und nach 76 Wochen 0,55 [0,38;0,80]).

Bei der Anzahl von Patienten mit mindestens einer patientenberichteten oder bestätigten **Hypoglykämie** gab es weder nach 24 noch nach 76 Wochen statistisch signifikante Unterschiede zwischen den Gruppen. Auch für die Anzahl von Studienabbrüchen aufgrund einer Hypoglykämie gab es keine signifikanten Unterschiede.

Beim Composite-Endpunkt „Erreichen des HbA1c-Zielwerts ohne Auftreten von Hypoglykämien“ gab es zu beiden Zeitpunkten einen signifikanten Unterschied zugunsten von Glibenclamid+Saxagliptin (OR [95%-KI] nach 24 bzw. 76 Wochen: 3,54 [1,28;9,80] bzw. 2,82 [1,16;6,87]).

In beiden Gruppen zeigte sich eine, wenn auch geringe, **Gewichtszunahme**, die in der Glibenclamid+Saxagliptin-Gruppe sowohl nach 24 als auch nach 76 Wochen etwas höher (0,8 kg und 1,0 kg) lag als in der Glibenclamid+Placebo-Gruppe (0,3 kg und 0,3 kg). Der Unterschied war für beide Zeitpunkte (LOCF) signifikant (MWD [95%-KI] nach 24 Wochen: 0,5 kg [0,1;0,8] und nach 76 Wochen: 0,70 kg [0,28;1,12]). Die Analyse mit den Observed-Cases ergab lediglich nach 76 Wochen einen signifikanten Unterschied (MWD [95%-KI]: 1,80 kg [0,73;3,23]).

Bei den **unerwünschten Ereignissen** gab es keine statistisch signifikanten Unterschiede zwischen den Gruppen. So betrug für die Gesamtrate unerwünschter Ereignisse das OR [95%-KI] nach 24 Wochen 0,71 [0,47;1,06] und nach 76 Wochen 0,72 [0,43;1,20]). Für die Gesamtrate der schwerwiegenden unerwünschten Ereignisse ergab sich ein OR [95%-KI] von 0,90 [0,30;2,72] bzw. 0,53 [0,26;1,09] und für die Abbruchrate aufgrund unerwünschter Ereignisse ein OR [95%-KI] von 1,06 [0,26;4,27] bzw. 1,43 [0,66;3,10], jeweils nach 24 bzw. 76 Wochen).

Bei der Anzahl der **kardialen Ereignisse** gab es keine statistisch signifikanten Unterschiede zwischen den Gruppen (OR [95%-KI] nach 24 Wochen: 1,06 [0,43;2,59] und nach 76 Wochen: 1,06 [0,51;2,21]).

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

Zusammenfassend kann man sagen, dass die Behandlung mit Saxagliptin bei Patienten, deren Blutzucker mit Sulfonylharnstoffen alleine nicht mehr ausreichend kontrolliert werden kann, sicher und wirksam ist. Die Kombination von Sulfonylharnstoff+Saxagliptin senkt den Blutzucker effektiv, ohne die Gesamtrate an Hypoglykämien signifikant zu erhöhen. Die Gewichtszunahme ist bei Behandlung mit Glibenclamid+Saxagliptin nur geringfügig und unwesentlich höher als in der Kontrollgruppe, und die Rate an unerwünschten Ereignissen ist in beiden Behandlungsgruppen ähnlich.

Zusatznutzen von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Insulin oder Insulin-Monotherapie bzw. Sulfonylharnstoff+DPP-4-Inhibitor

Im Folgenden werden die Ergebnisse zum medizinischen Zusatznutzen genauer betrachtet.

Sie basieren auf adjustierten indirekten Vergleichen zwischen Saxagliptin in Kombination mit einem Sulfonylharnstoff und der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Insulin oder Insulin-Monotherapie (wenn eine Insulintherapie angezeigt ist) bzw. der alternativen Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor (wenn eine Insulintherapie noch nicht angezeigt ist).

Die systematische Literaturrecherche identifizierte fünf relevante Studien für den indirekten Vergleich mit der zweckmäßigen Vergleichstherapie Sulfonylharnstoff+Insulin oder Insulin-Monotherapie (Nathan et al., 1988; Birkeland et al., 1994; Shank et al., 1995; Tovi et al., 1998; Turner et al., 1998). Alle Studien ermöglichten einen Vergleich gegen Insulin-Monotherapie. Die Studie von Shank 1995 erlaubte zudem einen Vergleich gegen Sulfonylharnstoff+Insulin. Diese Studie verwendete Glipizid als Sulfonylharnstoff. Wenn möglich, wurden die einzelnen Studien meta-analysiert und ein Gesamtschätzer berechnet. Für den indirekten Vergleich mit der alternativen Vergleichstherapie Sulfonylharnstoff+DPP-4-Inhibitor wurden insgesamt drei relevante Studien identifiziert, mit unterschiedlichen DPP-4-Inhibitoren (Sitagliptin, Vildagliptin und Linagliptin), für welche ein Gesamtschätzer berechnet werden konnte. In zwei der Studien wurde Glimepirid als Sulfonylharnstoff benutzt (Hermansen et al., 2007; Garber et al., 2008). Die Publikation von Lewin 2012 (Lewin et al., 2012) verwendete diverse Sulfonylharnstoffe als Basistherapie. Für die alternative Vergleichstherapie wurde sowohl ein Vergleich zur Gruppe Sulfonylharnstoff+andere DPP-4-Inhibitoren, als auch zur Leitsubstanz Sulfonylharnstoff+Sitagliptin durchgeführt.

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 intensiviertere Therapie zur Blutzuckerkontrolle ein höheres Risiko für das Erleiden schwerwiegender Hypoglykämien zur Folge haben kann (UK Prospective Diabetes Study Group, 1998b; Patel et al., 2008; Duckworth et al., 2009; Ismail-Beigi et al., 2010). Dabei gilt es zu berücksichtigen, dass insbesondere eine Insulin- und Sulfonylharnstofftherapie ein hohes Risiko für das Auftreten von Hypoglykämien birgt (AkdÄ, 2009). Dies gilt auch für eine Kombination eines DPP-4-Inhibitors mit einem Sulfonylharnstoff und ist auf den Wirkmechanismus der Sulfonylharnstoffe zurückzuführen (AkdÄ, 2009; Bristol-Myers Squibb / AstraZeneca, 2013).

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

Der indirekte Vergleich gegen Insulin-Monotherapie konnte für den Endpunkt „Ausmaß der Blutzuckerkontrolle (HbA1c)“ mit zwei Studien (Birkeland 1994 und Tovi 1998) durchgeführt werden, die meta-analysiert wurden. Im indirekten Vergleich zeigte sich ein statistisch signifikantes Ergebnis zugunsten der Insulin-Behandlungsgruppe nach 50 Wochen (MWD [95%-KI]: 1,41% [0,70;2,13]). Da die HbA1c-senkenden Effekte von Antidiabetika u. a. stark vom Ausgangs-HbA1c abhängen, muss das Ergebnis aber hinterfragt werden, da der Ausgangs-HbA1c-Wert insbesondere in der Studie von Tovi 1998 höher war. Darüber hinaus zeigte der Vergleich mit der Studie von Nathan 1988, welche behandlungsnaive Patienten einschloss und deshalb separat betrachtet wurde, keinen statistisch signifikanten Unterschied mehr (MWD [95%-KI]: -0,34% [-1,75;1,07]). Ein indirekter Vergleich gegen Sulfonylharnstoff+Insulin war aufgrund fehlender Angaben in der Publikation von Shank 1995 nicht möglich. Ein indirekter Vergleich für das Auftreten von Hypoglykämien war nur gegen Insulin-Monotherapie mit der Studie von Birkeland 1994 möglich. Es zeigte sich ein signifikantes Ergebnis nach 50 Wochen zugunsten von Sulfonylharnstoff+Saxagliptin gegenüber Insulin (OR [95%-KI]: 0,04 [0,00;0,88]).

Für den indirekten Vergleich von Sulfonylharnstoff+Saxagliptin mit der alternativen Vergleichstherapie Sulfonylharnstoff+anderer DPP-4-Inhibitor konnten alle drei Studien herangezogen und meta-analysiert werden. Der Vergleich ergab für den Endpunkt Blutzuckerkontrolle keine statistisch signifikanten Unterschiede zwischen der zu bewertenden Therapie und Sulfonylharnstoff+einem anderen DPP-4-Inhibitor als Gruppe (MWD [95%-KI]: -0,16% [-1,55;1,23]) bzw. Sulfonylharnstoff+Sitagliptin (MWD [95%-KI]: -0,15% [-0,45;0,15]). Auch in Bezug auf das Auftreten von Hypoglykämien zeigte sich weder für die patientenberichteten Hypoglykämien (DPP-4-Inhibitoren als Gruppe OR [95%-KI]: 0,88 [0,31;2,50] bzw. Sitagliptin OR [95%-KI]: 0,54 [0,13;2,33]), noch für die bestätigten Hypoglykämien (OR [95%-KI]: 0,51 [0,02;11,38], Vergleich nur mit Garber 2008 möglich) ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Gemessen an dem Ziel, eine adäquate Blutzuckerkontrolle bei gleichzeitiger Hypoglykämie-Vermeidung zu erreichen, ergibt sich aus der gemeinsamen Betrachtung dieser Endpunkte kein Hinweis auf eine Über- oder Unterlegenheit, weder für Sulfonylharnstoff+Saxagliptin gegenüber der zweckmäßigen Vergleichstherapie noch für Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+anderer DPP-4-Inhibitoren.

Betrachtung der Gewichtsveränderung

Die Adipositas (Fettleibigkeit), und hier insbesondere die abdominelle Adipositas, ist ein wichtiger Risikofaktor zur Entwicklung eines Diabetes mellitus Typ 2 (Freemantle et al., 2008). Die Adipositas führt zusammen mit Bewegungsmangel 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 (Porzolt et al., 2010). Einige Therapieoptionen haben diesbezüglich günstige Wirkungen auf das Körpergewicht, wie Metformin oder DPP-4-Inhibitoren. DPP-4-Inhibitoren werden dabei als gewichtneutral beschrieben. Einige andere Wirkstoffklassen sind ungünstig, wie etwa Sulfonylharnstoffe, Insulin oder Pioglitazon (AkdÄ, 2009; Bundesärztekammer (BÄK) et al., 2012), 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). In der Nationalen VersorgungsLeitlinie (Konsultationsfassung) wird dementsprechend bei einem Body-Mass-Index (BMI) von 27–35 kg/m² eine Gewichtsabnahme von 5% empfohlen (Bundesärztekammer (BÄK) et al., 2012).

Für den Endpunkt „Gewichtsveränderung“ war ein indirekter Vergleich gegen Sulfonylharnstoff+Insulin sowie gegen Insulin-Monotherapie möglich. Der Vergleich von Sulfonylharnstoff+Saxagliptin gegenüber Sulfonylharnstoff+Insulin nach 24 Wochen zeigte keinen statistisch signifikanten Unterschied (MWD [95%-KI]: -3,40 kg [-16,78;9,98]).

Gegen Insulin-Monotherapie gab es nach 50 Wochen keine statistisch signifikanten Unterschiede, wenn nur die Studien von Nathan 1988 und Turner 1998 betrachtet wurden, welche behandlungsnaive Patienten einschlossen (MWD [95%-KI]: -0,45 kg [-3,50;2,59]).

Wurde der Vergleich nur mit den Studien von Birkeland 1994 und Tovi 1998 durchgeführt, so zeigte sich eine statistisch signifikante Gewichtsreduktion in der mit Sulfonylharnstoff+Saxagliptin behandelten Patientengruppe gegenüber den mit Insulin behandelten Patienten (MWD [95%-KI]: -4,18 kg [-6,29;-2,07]).

Bei der Betrachtung aller vier Studien ergab sich ebenfalls ein statistisch signifikanter Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin (MWD [95%-KI]: -2,67 kg [-4,48;-0,86]).

Der indirekte Vergleich mit der alternativen Vergleichstherapie konnte mit zwei der drei Studien durchgeführt werden und ergab keine statistisch signifikanten Unterschiede zwischen Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+DPP-4-Inhibitor als Gruppe (MWD [95%-KI]: -0,18 kg [-1,04;0,69]) wie auch der Kombination Sulfonylharnstoff+Sitagliptin (MWD [95%-KI]: -0,60 kg [-1,59;0,39]).

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).

Gemessen an dem Ziel, eine Reduktion des Körpergewichts zu erreichen, ergibt sich gegenüber zweckmäßigen Vergleichstherapie eine bedeutsame Vermeidung nicht schwerwiegender Nebenwirkungen und damit ein Anhaltspunkt für einen beträchtlichen Zusatznutzen für Sulfonylharnstoff+Saxagliptin.

Bei den behandlungsnaiven Patienten zeigt sich dieser Effekt nicht. Da diese nicht zur Zielpopulation gehören, werden die indirekten Vergleiche, wo diese Studien enthalten sind, jedoch nicht für die Quantifizierung des Zusatznutzens herangezogen.

Gegenüber Sulfonylharnstoff+andere DPP-4-Inhibitoren ergibt sich keine Über- oder Unterlegenheit für Sulfonylharnstoff+Saxagliptin.

Betrachtung der unerwünschten Ereignisse

Keine der Studien zur zweckmäßigen Vergleichstherapie berichtet Ergebnisse zu unerwünschten Ereignissen, die für die Berechnung eines indirekten Vergleiches verwendbar gewesen wären.

Für die alternative Vergleichstherapie konnten alle drei Studien verwendet werden. Es ergaben sich keine signifikanten Unterschiede zwischen Sulfonylharnstoff+Saxagliptin und Sulfonylharnstoff+andere DPP-4-Inhibitoren als Gruppe, weder für die Gesamtrate unerwünschter Ereignisse (OR [95%-KI]: 0,36 [0,37;1,07]), schwerwiegender unerwünschter Ereignisse (OR [95%-KI]: 1,33 [0,33;5,26]) noch für Studienabbrüche aufgrund von unerwünschten Ereignissen (OR [95%-KI]: 1,65 [0,35;7,71]). Der Vergleich gegen Sulfonylharnstoff+Sitagliptin als Einzelsubstanz zeigte hingegen einen signifikanten Vorteil zugunsten von Sulfonylharnstoff+Saxagliptin (OR [95%-KI]: 0,43 [0,22;0,84]). Dieser Unterschied wird aber in Frage gestellt, da der Endpunkt in der Studie von Hermansen 2007 sehr wahrscheinlich anders operationalisiert wurde als in der Studie CV181040. Für schwerwiegende unerwünschte Ereignisse (OR [95%-KI]: 1,28 [0,24;6,82]) und Studienabbrüche aufgrund unerwünschter Ereignisse (OR [95%-KI]: 0,70 [0,05;9,28]) gab es auch gegen Sulfonylharnstoff+Sitagliptin keine signifikanten Unterschiede.

Insgesamt ergibt sich für das Auftreten unerwünschter Ereignisse (Gesamtrate) kein Hinweis auf einen Zusatznutzen gegenüber Sulfonylharnstoff+andere DPP-4-Inhibitoren.

Gesamtbetrachtung des Zusatznutzens

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

Die Ergebnisse zum Zusatznutzen zeigen, dass es für die Endpunkte „Ausmaß der Blutzuckerkontrolle (HbA1c)“ und „Hypoglykämien“ keinen Anhaltspunkt auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen (Sulfonylharnstoff+Saxagliptin gegen zweckmäßige Vergleichstherapie) gibt.

Für den Endpunkt „Gewichtsveränderungen“ ergibt sich für die Zielpopulation ein Anhaltspunkt auf einen beträchtlichen Zusatznutzen gegenüber der zweckmäßigen

Vergleichstherapie. Wie bereits erwähnt muss bei den Ergebnissen berücksichtigt werden, dass die Behandlung mit DPP-4-Inhibitoren als Monosubstanz weder mit einem erhöhten Hypoglykämie-Risiko noch mit Gewichtszunahmen assoziiert ist. Für die Behandlung mit Sulfonylharnstoffen bzw. mit Insulin werden diese Nebenwirkungen jedoch häufig berichtet (AkdÄ, 2009; Matthaei et al., 2009; Bundesärztekammer (BÄK) et al., 2012). Dies gilt damit auch für die Kombination von DPP-4-Inhibitoren mit Sulfonylharnstoffen und ist auf die Sulfonylharnstoffwirkung zurückzuführen (AkdÄ, 2009; Bristol-Myers Squibb / AstraZeneca, 2013).

Darüber hinaus ist eine Insulintherapie für den Patienten komplexer als eine Behandlung mit OAD. Im Gegensatz zu einer Therapie mit OAD sollte jede Insulintherapie von einer Blutglukoseselbstkontrolle begleitet sein, was eine intensive Schulung voraussetzt. Je nach Art der Insulintherapie sind hierfür 2-4 Blutzuckermessungen pro Tag erforderlich. Dabei hat insbesondere eine konventionelle Insulintherapie einen negativen Einfluss auf die Flexibilität der Patienten, da ein geregelter Tagesablauf mit festgesetzten Spritz- und Esszeiten sowie meistens auch Zwischenmahlzeiten zur Vermeidung von Hypoglykämien notwendig sind. Durch eine intensiviertere Insulintherapie kann man zwar eine deutlich größere Flexibilität im Alltag erreichen, dafür muss häufiger Insulin gespritzt und auch häufiger der Blutzucker gemessen werden (Matthaei et al., 2009).

Gegenüber der alternativen Vergleichstherapie zeigt sich für die Endpunkte „Ausmaß der Blutzuckerkontrolle (HbA1c)“, „Hypoglykämien“, „Gewichtsveränderungen“ und „unerwünschte Ereignisse“ kein Hinweis auf eine Über- oder Unterlegenheit für eine der beiden Behandlungsoptionen (Sulfonylharnstoff+Saxagliptin gegen Sulfonylharnstoff+andere DPP-4-Inhibitoren).

Aus der Zusammenschau aller Ergebnisse lässt sich ableiten, dass Saxagliptin in Kombination mit einem Sulfonylharnstoff eine wirksame und gut verträgliche Alternative zur zweckmäßigen Vergleichstherapie darstellt, die im Vergleich zu dieser in bedeutsamem Maße zur Gewichtsreduktion beitragen kann.

Insgesamt ergibt sich ein Anhaltspunkt auf einen beträchtlichen Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie und kein Hinweis auf einen Zusatznutzen gegenüber der alternativen Vergleichstherapie.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-101: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit Typ-2-Diabetes mellitus, für die die Anwendung von Metformin ungeeignet erscheint und die unter einer Sulfonylharnstoff-Monotherapie zusammen mit einer Diät und Bewegung einen nicht ausreichend kontrollierten Blutzucker aufweisen. (gegenüber der zweckmäßigen Vergleichstherapie)	Beträchtlicher Zusatznutzen
Erwachsene Patienten mit Typ-2-Diabetes mellitus, für die die Anwendung von Metformin ungeeignet erscheint und die unter einer Sulfonylharnstoff-Monotherapie zusammen mit einer Diät und Bewegung einen nicht ausreichend kontrollierten Blutzucker aufweisen und für die eine Insulintherapie noch nicht angezeigt ist. (gegenüber Sulfonylharnstoff+andere DPP-4-Inhibitoren)	Kein Zusatznutzen

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 4-102: 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+Saxagliptin) und der zweckmäßigen Vergleichstherapie (Sulfonylharnstoff+Insulin oder Insulin-Monotherapie) bzw. der alternativen Vergleichstherapie (Sulfonylharnstoff+DPP-4-Inhibitor) vor.

Mit Sulfonylharnstoff (+Placebo) als Brückenkomparator wurde deshalb ein adjustierter 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 nicht randomisierter 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.

Nicht zutreffend

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.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

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.

- 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.

Nicht zutreffend

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	Datenquelle
CV181040	<i>A Multicenter, Randomized, Double-blind, Placebo-controlled Phase 3 Trial to Evaluate the Efficacy and Safety of Saxagliptin in Combination with Glyburide in Subjects with Type 2 Diabetes Who Have Inadequate Glycemic Control on Glyburide Alone.</i>	(Bristol-Myers Squibb, 2008)
Birkeland 1994	<i>A long-term, randomized, comparative study of insulin versus sulfonylurea therapy in type 2 diabetes.</i>	(Birkeland et al., 1994)
Tovi 1998	<i>Well-being and symptoms in elderly Type 2 diabetes patients with poor metabolic control: Effect of insulin treatment.</i>	(Tovi et al., 1998)
Nathan 1988	<i>Glyburide or insulin for metabolic control in non-insulin-dependent diabetes mellitus. A randomized, double-blind study.</i>	(Nathan et al., 1988)
Turner 1998 (UKPDS24)	<i>United Kingdom prospective diabetes study 24: A 6-year, randomized, controlled trial comparing sulfonylurea, insulin, and metformin therapy in patients with newly diagnosed type 2 diabetes that could not be controlled with diet therapy.</i>	(Turner et al., 1998)
Shank 1995	<i>Bedtime insulin/daytime glipizide. Effective therapy for sulfonylurea failures in NIDDM.</i>	(Shank et al., 1995)

Studie	Titel	Datenquelle
Garber 2008 (CLAF237A2305)	<i>Effects of vildagliptin on glucose control in patients with type 2 diabetes inadequately controlled with a sulphonylurea</i>	(Garber et al., 2008)
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, multicenter, randomized, double-blind, placebo-controlled trial.</i>	(Lewin et al., 2012)

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).

Referenzliste

- [1] AkdÄ 2009. Empfehlungen zur antihyperglykämischen Therapie des Diabetes mellitus Typ 2. *Arzneiverordnung in der Praxis, Band 36, Sonderheft 1 (Therapieempfehlungen)*.
- [2] AstraZeneca / Bristol-Myers Squibb 2012. Zusatzanalysen (Composite-Endpunkt: Erreichen des HbA1c-Zielwerts (<6,5% / <7%) ohne Auftreten von Hypoglykämien sowie Subgruppenauswertung nach Alter und Geschlecht) für die Studie CV181040.
- [3] Birkeland, K. I., Hanssen, K. F., Urdal, P., et al. 1994. A long-term, randomized, comparative study of insulin versus sulfonylurea therapy in type 2 diabetes. *Journal of internal medicine*, 236, 305-13.
- [4] Bristol-Myers Squibb. 2008. A Multicenter, Randomized, Double-blind, Placebo-controlled Phase 3 Trial to Evaluate the Efficacy and Safety of Saxagliptin in Combination with Glyburide in Subjects with Type 2 Diabetes Who Have Inadequate Glycemic Control on Glyburide Alone.
- [5] Bristol-Myers Squibb / AstraZeneca 2013. Fachinformation Onglyza.
- [6] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV) & Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) 2012. Nationale VersorgungsLeitlinie Therapieplanung bei Typ-2-Diabetes. 1.0. Konsultationsfassung.
- [7] Chacra, A. R., Tan, G. H., Apanovitch, A., et al. 2009. 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. *Int J Clin Pract*, 63, 1395-406.
- [8] 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.
- [9] ClinicalTrials.gov. 2013a. *Efficacy and Safety of Vildagliptin in Combination With Glimpiride in Patients With Type 2 Diabetes* [Online]. Adresse: <http://clinicaltrials.gov/show/NCT00099944> [Aufgerufen am 05.03.2013].
- [10] ClinicalTrials.gov. 2013b. *Efficacy and Safety of Vildagliptin in Combination With Insulin in Patients With Type 2 Diabetes* [Online]. Adresse: <http://www.clinicaltrials.gov/ct2/show/NCT00099931> [Aufgerufen am 05.03.2013].
- [11] ClinicalTrials.gov. 2013c. *Does Saxagliptin Reduce the Risk of Cardiovascular Events When Used Alone or Added to Other Diabetes Medications (SAVOR- TIMI 53)* [Online]. Adresse: <http://clinicaltrials.gov/show/NCT01107886> [Aufgerufen am 05.03.2013].
- [12] ClinicalTrials.gov. 2013d. *Effects of DPP-4 Inhibition on Triglycerides* [Online]. Adresse: <http://clinicaltrials.gov/show/NCT01527747> [Aufgerufen am 05.03.2013].
- [13] ClinicalTrials.gov. 2013e. *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://www.clinicaltrials.gov/ct2/show/study/NCT00819091> [Aufgerufen am 05.03.2013].
- [14] Duckworth, W., Abaira, 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.
- [15] European Medicines Agency. 2009. EPAR_Public Assessment Report for Onglyza. Verfügbar: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/001039/WC500044319.pdf.
- [16] FDA 2008. Guidance for Industry: Diabetes Mellitus: Developing Drugs and Therapeutic Biologics for Treatment and Prevention (DRAFT GUIDANCE).
- [17] Freemantle, N., Holmes, J., Hockey, A., et al. 2008. How strong is the association between abdominal obesity and the incidence of type 2 diabetes? *Int J Clin Pract*, 62, 1391-6.
- [18] G-BA. 2010. Beschluss des Gemeinsamen Bundesausschusses über die Einleitung eines Stellungnahmeverfahrens zur Änderung der Arzneimittel-Richtlinie (AM-RL) in Anlage IV: Therapiehinweis zu Gliptinen bzw. DPP-4 Inhibitoren.
- [19] G-BA. 2012. Therapiehinweise gemäß § 92 Abs. 2 Satz 7 SGB V i. V. m. § 17 AM-RL zur wirtschaftlichen Verordnungsweise von Arzneimitteln.
- [20] 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.
- [21] Glenny, A. M., Altman, D. G., Song, F., et al. 2005. Indirect comparisons of competing interventions. *Health Technol Assess*, 9, 1-134, iii-iv.
- [22] 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.
- [23] 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.

- [24] Higgins, J. P., Thompson, S. G., Deeks, J. J., et al. 2003. Measuring inconsistency in meta-analyses. *BMJ*, 327, 557-60.
- [25] Higgins, J. P. & Green, S. 2008. *Cochrane Handbook for Systematic Reviews of Interventions*. , The Cochrane Collaboration.
- [26] ICTRP Search Portal. 2013a. *Efficacy and Safety of Vildagliptin in Combination With Glimpiride in Patients With Type 2 Diabetes* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00099944> [Aufgerufen am 05.03.2013].
- [27] ICTRP Search Portal. 2013b. *Does Saxagliptin Reduce the Risk of Cardiovascular Events When Used Alone or Added to Other Diabetes Medications* [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2009-017358-10-HU> [Aufgerufen am 05.03.2013].
- [28] IQWiG. 2009a. Urin- und Blutzuckerselbstmessung bei Diabetes mellitus Typ 2
- [29] IQWiG 2009b. Glinide zur Behandlung des Diabetes mellitus Typ 2 - Abschlussbericht A05-05C. Version 1.0.
- [30] IQWiG. 2011. Allgemeine Methoden Version 4.0 vom 23.09.2011. Verfügbar: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf.
- [31] 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.
- [32] Kerner, W., Brückel, J. & Böhm, O. 2004. Definition, Klassifikation und Diagnostik des Diabetes mellitus. *Evidenzbasierte Leitlinie DDG – Aktualisierung 10/2004*, 1-10.
- [33] Klipper, W. 2012. Niederschrift zum Beratungsgespräch gemäß § 8 AM-Nutzen V. Beratungsanforderung 2012-B-027- Saxagliptin zur Behandlung des Diabetes mellitus Typ 2.
- [34] Lewin, A. J., Arvay, 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.
- [35] Matthaei, S., Bierwirth, R., Fritsche, A., et al. 2009. Medikamentöse antihyperglykämische Therapie des Diabetes mellitus Typ 2. *Diabetologie* 2009, 4, 32-64.
- [36] Nathan, D. M., Roussel, A. & Godine, J. E. 1988. Glyburide or insulin for metabolic control in non-insulin-dependent diabetes mellitus. A randomized, double-blind study. *Annals of internal medicine*, 108, 334-40.
- [37] 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.
- [38] 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*.
- [39] Porzolt, F., Clouth, J., Deutschmann, M., et al. 2010. Preferences of diabetes patients and physicians: a feasibility study to identify the key indicators for appraisal of health care values. *Health and quality of life outcomes*, 8, 125.

- [40] Shank, M. L., Del Prato, S. & DeFronzo, R. A. 1995. Bedtime insulin/daytime glipizide. Effective therapy for sulfonylurea failures in NIDDM. *Diabetes*, 44, 165-72.
- [41] 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.
- [42] 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.
- [43] Tovi, J., Engfeldt, P., Tovi, J., et al. 1998. Well-being and symptoms in elderly Type 2 diabetes patients with poor metabolic control: Effect of insulin treatment. *Practical Diabetes International*, 15, 73-77.
- [44] Turner, R., Murchison, L., Wright, A. D., et al. 1998. United Kingdom prospective diabetes study 24: A 6-year, randomized, controlled trial comparing sulfonylurea, insulin, and metformin therapy in patients with newly diagnosed type 2 diabetes that could not be controlled with diet therapy. *Annals of internal medicine*, 128, 165-175.
- [45] 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.
- [46] UK Prospective Diabetes Study Group 1998b. 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.
- [47] 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.
- [48] 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.

Suchstrategien zur Suche nach RCT mit dem zu bewertenden Arzneimittel

Die Suchstrategien sind in Tabelle 4-103, Tabelle 4-104 und Tabelle 4-105 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 von BMS/AZ durchgeführten Studien die Datengrundlage für die Bewertung des medizinischen Nutzens und Zusatznutzens von Saxagliptin im Abschnitt 4.4.2 ab. Die Suche wurde für Abschnitt 4.3.1.1.2 durchgeführt.

Tabelle 4-103: DIMDI Suchstrategie Embase 1947 bis heute: am 04. Februar 2013

Datenbankname	Embase
Suchoberfläche	DIMDI
Datum der Suche	04.02.2013
Zeitsegment	1947 bis 2013 Woche 6
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)

#	Suchbegriffe	Ergebnis
1	EM47	26336326
2	FT=Diabetes mellitus OR (CT D "Diabetes mellitus" OR UT="Diabetes mellitus" OR IT="Diabetes mellitus" OR SH="Diabetes mellitus")	504987
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")	371230
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	109840
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	4794
6	2 OR 3 OR 4 OR 5	505210
7	FT=Saxagliptin OR (CT D "Saxagliptin" OR UT="Saxagliptin" OR IT="Saxagliptin" OR SH="Saxagliptin")	840
8	FT=Onglyza OR (CT D "Onglyza" OR UT="Onglyza" OR IT="Onglyza" OR SH="Onglyza")	807
9	7 OR 8	840
10	6 AND 9	760
11	random?/(Ti;Ab)	688304
12	double-blind?/(Ti;Ab)	126127
13	placebo?/(Ti;Ab;Ct) OR CT=placebo?	283534
14	11 OR 12 OR 13	886060
15	10 AND 14	339
16	(15) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	334
17	DT=Review	1852670
18	16 NOT 17	127

#	Suchbegriffe	Ergebnis
19	18 NOT SU=Medline	125

Tabelle 4-104: DIMDI Suchstrategie MEDLINE 1950 bis heute: am 04. Februar 2013

Datenbankname	Medline Direct
Suchoberfläche	DIMDI
Datum der Suche	04.02.2013
Zeitsegment	1950 bis 2013 Woche 6
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- Strategy minimizing difference between sensitivity and specificity (Wong et al., 2006) (angepasst für DIMDI)

#	Suchbegriffe	Ergebnis
1	ME60	22474774
2	FT=Diabetes mellitus OR (CT D "Diabetes mellitus" OR UT="Diabetes mellitus" OR IT="Diabetes mellitus" OR SH="Diabetes mellitus")	324074
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")	156068
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	75791
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	4687
6	2 OR 3 OR 4 OR 5	324588
7	FT=Saxagliptin OR (CT D "Saxagliptin" OR UT="Saxagliptin" OR IT="Saxagliptin" OR SH="Saxagliptin")	192
8	FT=Onglyza OR (CT D "Onglyza" OR UT="Onglyza" OR IT="Onglyza" OR SH="Onglyza")	13
9	7 OR 8	192
10	6 AND 9	156
11	DT=Randomized controlled trial	338147
12	Randomized/(Ti;Ab;Ct;TE) OR CT=Randomized?	402275

#	Suchbegriffe	Ergebnis
13	Placebo/(Ti;Ab;Ct;TE) OR CT=Placebo?	160072
14	11 OR 12 OR 13	632724
15	10 AND 14	65
16	(15) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	65

Tabelle 4-105: DIMDI Suchstrategie Cochrane 1948 bis heute: am 04. Februar 2013

Datenbankname	Cochrane
Suchoberfläche	DIMDI
Datum der Suche	04.02.2013
Zeitsegment	CCTR93: 1948 bis 2013 Woche 6
Suchfilter	Aufgrund der geringen Trefferzahl wurde kein Filter verwendet

#	Suchbegriffe	Ergebnis
1	CCTR93	677775
2	FT=Diabetes mellitus OR (CT D "Diabetes mellitus" OR UT="Diabetes mellitus" OR IT="Diabetes mellitus" OR SH="Diabetes mellitus")	16502
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 02"; "DIABETES MELLITUS, NON INSULIN DEPENDENT"; "DIABETESMELLITUS, NONINSULIN DEPENDENT"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, TYPE II"; "DIABETES MELLITUS, TYPE 2")	8872
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	6876
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	368
6	2 OR 3 OR 4 OR 5	16584
7	FT=Saxagliptin OR (CT D "Saxagliptin" OR UT="Saxagliptin" OR IT="Saxagliptin" OR SH="Saxagliptin")	23
8	FT=Onglyza OR (CT D "Onglyza" OR UT="Onglyza" OR IT="Onglyza" OR SH="Onglyza")	2
9	7 OR 8	23
10	6 AND 9	19
11	(10) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	19

Suchstrategien zur Suche nach RCT mit der zweckmäßigen Vergleichstherapie

Die Suchstrategien sind in Tabelle 4-106, Tabelle 4-107 und Tabelle 4-108 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 von BMS/AZ durchgeführten Studien die Datengrundlage für die Bewertung des medizinischen Nutzens und Zusatznutzens von Saxagliptin im Abschnitt 4.4.2 ab. Die Suche wurde für Abschnitt 4.3.2.1.1.2 durchgeführt.

Tabelle 4-106: DIMDI Suchstrategie Embase 1947 bis heute: am 07. Februar 2013

Datenbankname	Embase
Suchoberfläche	DIMDI
Datum der Suche	07.02.2013
Zeitsegment	1947 bis 2013 Woche 6
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)

#	Suchbegriffe	Ergebnis
1	EM47	26351429
2	FT=Diabetes mellitus OR (CT D "Diabetes mellitus" OR UT="Diabetes mellitus" OR IT="Diabetes mellitus" OR SH="Diabetes mellitus")	505460
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 2"; "DIABETES MELLITUS TYPE II"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, TYPE 2"; "DIABETES MELLITUS, TYPE II"; "DIABETES MELLITUS,NON INSULIN DEPENDENT")	371577
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	109997
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	4816
6	2 OR 3 OR 4 OR 5	505684
7	FT=Insulin OR (CT D "Insulin" OR UT= "Insulin" OR IT= "Insulin" OR SH= "Insulin")	459806
8	CT=("SULFONYLURIEA"; "SULFONYLUREA COMPOUNDS")	15356
9	FT=Sulfonylurea OR (CT D "Sulfonylurea" OR UT= "Sulfonylurea" OR IT= "Sulfonylurea" OR SH= "Sulfonylurea")	17777
10	FT=Sulphonylurea OR (CT D "Sulphonylurea" OR UT= "Sulphonylurea" OR IT= "Sulphonylurea" OR SH= "Sulphonylurea")	9170
11	FT=Sulfonylharnstoff OR (CT D "Sulfonylharnstoff" OR UT= "Sulfonylharnstoff" OR IT= "Sulfonylharnstoff" OR SH= "Sulfonylharnstoff")	88

#	Suchbegriffe	Ergebnis
12	FT=Glibenclamid OR (CT D "Glibenclamid" OR UT= "Glibenclamid" OR IT= "Glibenclamid" OR SH= "Glibenclamid")	18149
13	FT=Glibenclamide OR (CT D "Glibenclamide" OR UT= "Glibenclamide" OR IT= "Glibenclamide" OR SH= "Glibenclamide")	19113
14	FT=Glibenclamid? OR (CT D "Glibenclamid?" OR UT= "Glibenclamid?" OR IT= "Glibenclamid?" OR SH= "Glibenclamid?")	19124
15	FT=Glyburid OR (CT D "Glyburid" OR UT= "Glyburid" OR IT= "Glyburid" OR SH= " Glyburid")	1
16	FT=Glyburide OR (CT D "Glyburide" OR UT= "Glyburide" OR IT= "Glyburide" OR SH= " Glyburide")	18294
17	FT=Glyburid? OR (CT D "Glyburid?" OR UT= "Glyburid?" OR IT= "Glyburid?" OR SH= " Glyburid?")	18429
18	FT=Glimepirid OR (CT D "Glimepirid" OR UT= "Glimepirid" OR IT= "Glimepirid" OR SH= "Glimepirid")	27
19	FT=Glimepiride OR (CT D "Glimepiride" OR UT= "Glimepiride" OR IT= "Glimepiride" OR SH= "Glimepiride")	3804
20	FT=Glimepirid? OR (CT D "Glimepirid?" OR UT= "Glimepirid?" OR IT= "Glimepirid?" OR SH= "Glimepirid?")	3807
21	FT=Glipizid OR (CT D "Glipizid" OR UT= "Glipizid" OR IT= "Glipizid" OR SH= "Glipizid")	11
22	FT=Glipizide OR (CT D "Glipizide" OR UT= "Glipizide" OR IT= "Glipizide" OR SH= "Glipizide")	4322
23	FT=Glipizid? OR (CT D "Glipizid?" OR UT= "Glipizid?" OR IT= "Glipizid?" OR SH= "Glipizid?")	4327
24	FT=Gliclazid OR (CT D "Gliclazid" OR UT= "Gliclazid" OR IT= "Gliclazid" OR SH= "Gliclazid")	4
25	FT=Gliclazide OR (CT D "Gliclazide" OR UT= "Gliclazide" OR IT= "Gliclazide" OR SH= "Gliclazide")	3920
26	FT=Gliclazid? OR (CT D "Gliclazid?" OR UT= "Gliclazid?" OR IT= "Gliclazid?" OR SH= "Gliclazid?")	3920
27	FT= Gliquidon OR (CT D "Gliquidon" OR UT= "Gliquidon" OR IT= "Gliquidon" OR SH= "Gliquidon")	14
28	FT= Gliquidone OR (CT D "Gliquidone" OR UT= "Gliquidone" OR IT= "Gliquidone" OR SH= "Gliquidone")	578
29	FT= Gliquidon? OR (CT D "Gliquidon?" OR UT= "Gliquidon?" OR IT= "Gliquidon?" OR SH= "Gliquidon?")	579
30	8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29	36723
31	7 AND 30	20745
32	6 AND 31	18184
33	random?/(Ti;Ab)	689046
34	double-blind?/(Ti;Ab)	126205
35	placebo?/(Ti;Ab;Ct) OR CT=placebo?	283783
36	33 OR 34 OR 35	886970
37	32 AND 36	4127

#	Suchbegriffe	Ergebnis
38	(37) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	3797
39	DT="REVIEW"	1853948
40	38 NOT 39	2512
41	40 NOT SU=Medline	2428

Tabelle 4-107: DIMDI Suchstrategie MEDLINE 1950 bis heute: am 07. Februar 2013

Datenbankname	Medline Direct
Suchoberfläche	DIMDI
Datum der Suche	07.02.2013
Zeitsegment	1950 bis 2013 Woche 6
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- Strategy minimizing difference between sensitivity and specificity (Wong et al., 2006) (angepasst für DIMDI)

#	Suchbegriffe	Ergebnis
1	ME60	22484519
2	FT=diabetes mellitus OR (CT D "diabetes mellitus" OR UT="diabetes mellitus" OR IT="diabetes mellitus" OR SH="diabetes mellitus")	324226
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")	156140
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	75832
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	4696
6	2 OR 3 OR 4 OR 5	324743
7	FT=Insulin OR (CT D "Insulin" OR UT= "Insulin" OR IT= "Insulin" OR SH= "Insulin")	293225
8	CT=("SULFONYLUREA COMPOUNDS"; "SULFONYLURIEA CPDS")	4354
9	FT=Sulfonylurea OR (CT D "Sulfonylurea" OR UT= "Sulfonylurea" OR IT= "Sulfonylurea" OR SH= "Sulfonylurea")	7258
10	FT=Sulphonylurea OR (CT D "Sulphonylurea" OR UT= "Sulphonylurea" OR IT= "Sulphonylurea" OR SH= "Sulphonylurea")	1316
11	(FT=Sulfonylharnstoff OR (CT D "Sulfonylharnstoff" OR UT= "Sulfonylharnstoff" OR IT= "Sulfonylharnstoff" OR SH=	45

#	Suchbegriffe	Ergebnis
	"Sulfonylharnstoff") OR (CTG D "Sulfonylharnstoff" OR UTG="Sulfonylharnstoff" OR ITG="Sulfonylharnstoff" OR SHG="Sulfonylharnstoff")	
12	(FT=Glibenclamid OR (CT D "Glibenclamid" OR UT="Glibenclamid" OR IT="Glibenclamid" OR SH="Glibenclamid")) OR (CTG D "Glibenclamid" OR UTG="Glibenclamid" OR ITG="Glibenclamid" OR SHG="Glibenclamid")	5409
13	FT=Glibenclamide OR (CT D "Glibenclamide" OR UT="Glibenclamide" OR IT="Glibenclamide" OR SH="Glibenclamide")	8382
14	FT=Glibenclamid? OR (CT D "Glibenclamid?" OR UT="Glibenclamid?" OR IT="Glibenclamid?" OR SH="Glibenclamid?")	8396
15	(FT=Glyburid OR (CT D "Glyburid" OR UT="Glyburid" OR IT="Glyburid" OR SH="Glyburid")) OR (CTG D "Glyburid" OR UTG="Glyburid" OR ITG="Glyburid" OR SHG="Glyburid")	5390
16	FT=Glyburide OR (CT D "Glyburide" OR UT="Glyburide" OR IT="Glyburide" OR SH="Glyburide")	5817
17	FT=Glyburid? OR (CT D "Glyburid?" OR UT="Glyburid?" OR IT="Glyburid?" OR SH="Glyburid?")	5817
18	(FT=Glimepirid OR (CT D "Glimepirid" OR UT="Glimepirid" OR IT="Glimepirid" OR SH="Glimepirid")) OR (CTG D "Glimepirid" OR UTG="Glimepirid" OR ITG="Glimepirid" OR SHG="Glimepirid")	11
19	FT=Glimepiride OR (CT D "Glimepiride" OR UT="Glimepiride" OR IT="Glimepiride" OR SH="Glimepiride")	794
20	FT=Glimepirid? OR (CT D "Glimepirid?" OR UT="Glimepirid?" OR IT="Glimepirid?" OR SH="Glimepirid?")	799
21	(FT=Glipizid OR (CT D "Glipizid" OR UT="Glipizid" OR IT="Glipizid" OR SH="Glipizid")) OR (CTG D "Glipizid" OR UTG="Glipizid" OR ITG="Glipizid" OR SHG="Glipizid")	612
22	FT=Glipizide OR (CT D "Glipizide" OR UT="Glipizide" OR IT="Glipizide" OR SH="Glipizide")	917
23	FT=Glipizid? OR (CT D "Glipizid?" OR UT="Glipizid?" OR IT="Glipizid?" OR SH="Glipizid?")	922
24	(FT=Gliclazid OR (CT D "Gliclazid" OR UT="Gliclazid" OR IT="Gliclazid" OR SH="Gliclazid")) OR (CTG D "Gliclazid" OR UTG="Gliclazid" OR ITG="Gliclazid" OR SHG="Gliclazid")	701
25	FT=Gliclazide OR (CT D "Gliclazide" OR UT="Gliclazide" OR IT="Gliclazide" OR SH="Gliclazide")	967
26	FT=Gliclazid? OR (CT D "Gliclazid?" OR UT="Gliclazid?" OR IT="Gliclazid?" OR SH="Gliclazid?")	967
27	(FT=Gliquidon OR (CT D "Gliquidon" OR UT="Gliquidon" OR IT="Gliquidon" OR SH="Gliquidon")) OR (CTG D "Gliquidon" OR UTG="Gliquidon" OR ITG="Gliquidon" OR SHG="Gliquidon")	10
28	FT=Gliquidone OR (CT D "Gliquidone" OR UT="Gliquidone" OR IT="Gliquidone" OR SH="Gliquidone")	132
29	FT=Gliquidon? OR (CT D "Gliquidon?" OR UT="Gliquidon?" OR IT="Gliquidon?" OR SH="Gliquidon?")	137
30	8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29	15660

#	Suchbegriffe	Ergebnis
31	7 AND 30	5778
32	6 AND 31	4139
33	DT=Randomized controlled trial	338316
34	Randomized/(Ti;Ab;Ct;TE) OR CT=Randomized?	402564
35	Placebo/(Ti;Ab;Ct;TE) OR CT=Placebo?	160157
36	33 OR 34 OR 35	633098
37	32 AND 36	984
38	(37) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	912

Tabelle 4-108: DIMDI Suchstrategie Cochrane 1948 bis heute: am 07. Februar 2013

Datenbankname	Cochrane
Suchoberfläche	DIMDI
Datum der Suche	07.02.2013
Zeitsegment	CCTR93: 1948 bis 2013 Woche 6
Suchfilter	Aufgrund der geringen Trefferzahl wurde kein Filter verwendet

#	Suchbegriffe	Ergebnis
1	CCTR93	67775
2	FT=Diabetes mellitus OR (CT D "Diabetes mellitus" OR UT="Diabetes mellitus" OR IT="Diabetes mellitus" OR SH="Diabetes mellitus")	16502
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")	8872
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	6876
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	368
6	2 OR 3 OR 4 OR 5	16584

#	Suchbegriffe	Ergebnis
7	FT=Insulin OR (CT D "Insulin" OR UT= "Insulin" OR IT= "Insulin" OR SH= "Insulin")	16768
8	CT=("SULFONYLUREA COMPOUNDS"; "SULFONYLURIEA CPDS"	466
9	FT=Sulfonylurea OR (CT D "Sulfonylurea" OR UT= "Sulfonylurea" OR IT= "Sulfonylurea" OR SH= "Sulfonylurea")	820
10	FT=Sulphonylurea OR (CT D "Sulphonylurea" OR UT= "Sulphonylurea" OR IT= "Sulphonylurea" OR SH= "Sulphonylurea")	247
11	FT=Sulfonylharnstoff OR (CTG D "Sulfonylharnstoff" OR UTG= "Sulfonylharnstoff" OR ITG= "Sulfonylharnstoff" OR SHG= "Sulfonylharnstoff")	14
12	(FT=Glibenclamid OR (CT D "Glibenclamid" OR UT= "Glibenclamid" OR IT= "Glibenclamid" OR SH= "Glibenclamid")) OR (CTG D "Glibenclamid" OR UTG= "Glibenclamid" OR ITG= "Glibenclamid" OR SHG= "Glibenclamid")	37
13	FT=Glibenclamide OR (CT D "Glibenclamide" OR UT= "Glibenclamide" OR IT= "Glibenclamide" OR SH= "Glibenclamide")	623
14	FT=Glibenclamid? OR (CT D "Glibenclamid?" OR UT= "Glibenclamid?" OR IT= "Glibenclamid?" OR SH= "Glibenclamid?")	623
15	FT=Glyburide OR (CT D "Glyburide" OR UT= "Glyburide" OR IT= "Glyburide" OR SH= " Glyburide")	687
16	FT=Glyburid? OR (CT D "Glyburid?" OR UT= "Glyburid?" OR IT= "Glyburid?" OR SH= " Glyburid?")	692
17	(FT=Glimepirid OR (CT D "Glimepirid" OR UT= "Glimepirid" OR IT= "Glimepirid" OR SH= "Glimepirid")) OR (CTG D "Glimepirid" OR UTG= "Glimepirid" OR ITG= "Glimepirid" OR SHG= "Glimepirid")	3
18	FT=Glimepiride OR (CT D "Glimepiride" OR UT= "Glimepiride" OR IT= "Glimepiride" OR SH= "Glimepiride")	207
19	FT=Glimepirid? OR (CT D "Glimepirid?" OR UT= "Glimepirid?" OR IT= "Glimepirid?" OR SH= "Glimepirid?")	208
20	(FT=Glipizid OR (CT D "Glipizid" OR UT= "Glipizid" OR IT= "Glipizid" OR SH= "Glipizid")) OR (CTG D "Glipizid" OR UTG= "Glipizid" OR ITG= "Glipizid" OR SHG= "Glipizid")	0
21	FT=Glipizide OR (CT D "Glipizide" OR UT= "Glipizide" OR IT= "Glipizide" OR SH= "Glipizide")	196
22	FT=Glipizid? OR (CT D "Glipizid?" OR UT= "Glipizid?" OR IT= "Glipizid?" OR SH= "Glipizid?")	196
23	(FT=Gliclazid OR (CT D "Gliclazid" OR UT= "Gliclazid" OR IT= "Gliclazid" OR SH= "Gliclazid")) OR (CTG D "Gliclazid" OR UTG= "Gliclazid" OR ITG= "Gliclazid" OR SHG= "Gliclazid")	0
24	FT=Gliclazide OR (CT D "Gliclazide" OR UT= "Gliclazide" OR IT= "Gliclazide" OR SH= "Gliclazide")	205

#	Suchbegriffe	Ergebnis
25	FT=Gliclazid? OR (CT D "Gliclazid?" OR UT= "Gliclazid?" OR IT= "Gliclazid?" OR SH= "Gliclazid?")	205
26	(FT= Gliquidon OR (CT D "Gliquidon" OR UT= "Gliquidon" OR IT= "Gliquidon" OR SH= "Gliquidon")) OR (CTG D "Gliquidon" OR UTG= "Gliquidon" OR ITG= "Gliquidon" OR SHG= "Gliquidon")	2
27	FT= Gliquidone OR (CT D "Gliquidone" OR UT= "Gliquidone" OR IT= "Gliquidone" OR SH= "Gliquidone")	20
28	FT= Gliquidon? OR (CT D "Gliquidon?" OR UT= "Gliquidon?" OR IT= "Gliquidon?" OR SH= "Gliquidon?")	20
29	8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28	1602
30	7 AND 29	1118
31	6 AND 30	1007
32	31 AND LA=(ENGLISH; GERMAN) AND pps=Mensch	955

Suchstrategien zur Suche nach RCT mit der alternativen Vergleichstherapie (DPP-4-Inhibitoren)

Die Suchstrategien sind in Tabelle 4-109, Tabelle 4-110 und Tabelle 4-111 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 von BMS/AZ durchgeführten Studien die Datengrundlage für die Bewertung des medizinischen Nutzens und Zusatznutzens von Saxagliptin im Abschnitt 4.4.2 ab. Die Suche wurde für Abschnitt 4.3.2.1.1.2 durchgeführt.

Tabelle 4-109: DIMDI Suchstrategie Embase 1947 bis heute: am 04. Februar 2013

Datenbankname	Embase
Suchoberfläche	DIMDI
Datum der Suche	04.02.2013
Zeitsegment	1947 bis 2013 Woche 6
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)

#	Suchbegriffe	Ergebnis
1	EM47	26336326

2	FT=Diabetes mellitus OR (CT D "Diabetes mellitus" OR UT="Diabetes mellitus" OR IT="Diabetes mellitus" OR SH="Diabetes mellitus")	504987
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 2"; "DIABETES MELLITUS TYPE II"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, TYPE 2"; "DIABETES MELLITUS, TYPE II"; "DIABETES MELLITUS,NON INSULIN DEPENDENT")	371230
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	109840
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	4794
6	2 OR 3 OR 4 OR 5	505210
7	CT=("DIPEPTIDYL PEPTIDASE 4 INHIBITOR"; "DIPEPTIDYL PEPTIDASE IV INHIBITOR"; "DIPEPTIDYL PEPTIDASE IV INIHIBITOR"; "DIPEPTIDYL PEPTIDASE IV INHIBITORS"; "DIPEPTIDYL-PEPTIDASE IV INHIBITORS"; "DIPEPTIDYLPEPTIDASE 4 INHIBITOR"; "DIPEPTIDYLPEPTIDASE IV INHIBITOR")	2263
8	FT=Dipeptidyl Peptidase 4 Inhibitor? OR (CT D "Dipeptidyl Peptidase 4 Inhibitor"? OR UT="Dipeptidyl Peptidase 4 Inhibitor"? OR IT="Dipeptidyl Peptidase 4 Inhibitor"? OR SH="Dipeptidyl Peptidase 4 Inhibitor"?)	4403
9	FT=Dipeptidyl-Peptidase 4 Inhibitor? OR (CT D "Dipeptidyl- Peptidase 4 Inhibitor"? OR UT="Dipeptidyl-Peptidase 4 Inhibitor"? OR IT="Dipeptidyl-Peptidase 4 Inhibitor"? OR SH="Dipeptidyl- Peptidase 4 Inhibitor"?)	479
10	FT=Dipeptidyl#Peptidase 4 Inhibitor? OR (CT D "Dipeptidyl"#Peptidase 4 Inhibitor"? OR UT="Dipeptidyl"#Peptidase 4 Inhibitor"? OR IT="Dipeptidyl"#Peptidase 4 Inhibitor"? OR SH="Dipeptidyl"#Peptidase 4 Inhibitor"?)	4392
11	FT=Dipeptidyl#Peptidase#4#Inhibitor? OR (CT D "Dipeptidyl"#Peptidase"#4"#Inhibitor"? OR UT="Dipeptidyl"#Peptidase"#4"#Inhibitor"? OR IT="Dipeptidyl"#Peptidase"#4"#Inhibitor"? OR SH="Dipeptidyl"#Peptidase"#4"#Inhibitor"?)	4390
12	FT=Dipeptidyl Peptidase IV Inhibitor? OR (CT D "Dipeptidyl Peptidase IV Inhibitor"? OR UT="Dipeptidyl Peptidase IV Inhibitor"? OR IT="Dipeptidyl Peptidase IV Inhibitor"? OR SH="Dipeptidyl Peptidase IV Inhibitor"?)	4430
13	FT=Dipeptidyl-Peptidase IV Inhibitor? OR (CT D "Dipeptidyl- Peptidase IV Inhibitor"? OR UT="Dipeptidyl-Peptidase IV Inhibitor"? OR IT="Dipeptidyl-Peptidase IV Inhibitor"? OR SH="Dipeptidyl-Peptidase IV Inhibitor"?)	4430
14	FT=Dipeptidyl#Peptidase IV Inhibitor? OR (CT D "Dipeptidyl"#Peptidase IV Inhibitor"? OR UT="Dipeptidyl"#Peptidase IV Inhibitor"? OR IT="Dipeptidyl"#Peptidase IV Inhibitor"? OR SH="Dipeptidyl"#Peptidase IV Inhibitor"?)	4392
15	FT=Dipeptidyl#Peptidase#IV#Inhibitor? OR (CT D "Dipeptidyl"#Peptidase"#IV"#Inhibitor"? OR UT="Dipeptidyl"#Peptidase"#IV"#Inhibitor"? OR IT="Dipeptidyl"#Peptidase"#IV"#Inhibitor"? OR SH="Dipeptidyl"#Peptidase"#IV"#Inhibitor"?)	4389
16	FT=DPP-4-Inhibitor? OR (CT D "DPP-4-Inhibitor"? OR UT="DPP- 4-	784

	Inhibitor"? OR IT="DPP-4-Inhibitor"? OR SH="DPP- 4-Inhibitor"?)	
17	FT=DPP-IV-Inhibitor? OR (CT D "DPP-IV-Inhibitor"? OR UT="DPPIV-Inhibitor"? OR IT="DPP-IV-Inhibitor"? OR SH="DPPIV- Inhibitor"?)	423
18	FT=Gliptin? OR (CT D "Gliptin"? OR UT="Gliptin"? OR IT="Gliptin"? OR SH="Gliptin"?)	4401
19	FT=Sitagliptin? OR (CT D "Sitagliptin"? OR UT="Sitagliptin"? OR IT="Sitagliptin"? OR SH="Sitagliptin"?)	2507
20	FT=Vildagliptin? OR (CT D "Vildagliptin"? OR UT="Vildagliptin"? OR IT="Vildagliptin"? OR SH="Vildagliptin"?)	1520
21	FT=Linagliptin? OR (CT D "Linagliptin"? OR UT="Linagliptin"? OR IT="Linagliptin"? OR SH="Linagliptin"?)	365
22	7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21	4606
23	6 AND 22	3714
24	random?/(Ti;Ab)	688304
25	double-blind?/(Ti;Ab)	126127
26	placebo?/(Ti;Ab;Ct) OR CT=placebo?	283534
27	24 OR 25 OR 26	886060
28	23 AND 27	1265
29	(28) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	1204
30	DT="REVIEW"	1852670
31	29 NOT 30	624
32	31 NOT SU=Medline	611

Tabelle 4-110: DIMDI Suchstrategie MEDLINE 1950 bis heute: am 04. Februar 2013

Datenbankname	Medline Direct
Suchoberfläche	DIMDI
Datum der Suche	04.02.2013
Zeitsegment	1950 bis 2013 Woche 6
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- Strategy minimizing difference between sensitivity and specificity (Wong et al., 2006) (angepasst für DIMDI)

#	Suchbegriffe	Ergebnis
1	ME60	22474774
2	FT=diabetes mellitus OR (CT D "diabetes mellitus" OR UT="diabetes mellitus" OR IT="diabetes mellitus" OR SH="diabetes mellitus")	324074
3	CT=("DIABETES MELLITUS"; "DIABETES MELLITUS TYPE 02";	156068

	"DIABETES MELLITUS, NON INSULIN DEPENDENT"; "DIABETES MELLITUS, NONINSULIN DEPENDENT"; "DIABETES MELLITUS, NON-INSULIN-DEPENDENT"; "DIABETES MELLITUS, TYPE II"; "DIABETES MELLITUS, TYPE 2")	
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	75791
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	4687
6	2 OR 3 OR 4 OR 5	324588
7	FT=Dipeptidyl#Peptidase#4 Inhibitor? OR (CT D "Dipeptidyl#"Peptidase#"4 Inhibitor"? OR UT="Dipeptidyl#"Peptidase#"4 Inhibitor"? OR IT="Dipeptidyl#"Peptidase#"4 Inhibitor"? OR SH="Dipeptidyl#"Peptidase#"4 Inhibitor"?)	1177
8	FT=Dipeptidyl-Peptidase IV Inhibitor? OR (CT D "Dipeptidyl- Peptidase IV Inhibitor"? OR UT="Dipeptidyl-Peptidase IV Inhibitor"? OR IT="Dipeptidyl-Peptidase IV Inhibitor"? OR SH="Dipeptidyl-Peptidase IV Inhibitor"?)	1323
9	CT=("DIPEPTIDYL PEPTIDASE 4 INHIBITORS"; "DIPEPTIDYL PEPTIDASE IV INHIBITORS"; "DIPEPTIDYL-PEPTIDASE 4 INHIBITORS"; "DIPEPTIDYL-PEPTIDASE IV INHIBITORS")	1177
10	FT=Gliptin? OR (CT D "Gliptin"? OR UT="Gliptin"? OR IT="Gliptin"? OR SH="Gliptin"?)	1201
11	FT=DPP-4 Inhibitor? OR (CT D "DPP-4 Inhibitor"? OR UT="DPP-4 Inhibitor"? OR IT="DPP-4 Inhibitor"? OR SH="DPP-4 Inhibitor"?)	600
12	FT=DPP-IV Inhibitor? OR (CT D "DPP-IV Inhibitor"? OR UT="DPPIV Inhibitor"? OR IT="DPP-IV Inhibitor"? OR SH="DPP-IV Inhibitor"?)	331
13	FT=Dipeptidyl#Peptidase 4 Inhibitor? OR (CT D "Dipeptidyl#"Peptidase 4 Inhibitor"? OR UT="Dipeptidyl#"Peptidase 4 Inhibitor"? OR IT="Dipeptidyl#"Peptidase 4 Inhibitor"? OR SH="Dipeptidyl#"Peptidase 4 Inhibitor"?)	1184
14	FT=Sitagliptin? OR (CT D "Sitagliptin"? OR UT="Sitagliptin"? OR IT="Sitagliptin"? OR SH="Sitagliptin"?)	731
15	FT=Vildagliptin? OR (CT D "Vildagliptin"? OR UT="Vildagliptin"? OR IT="Vildagliptin"? OR SH="Vildagliptin"?)	418
16	FT=Linagliptin? OR (CT D "Linagliptin"? OR UT="Linagliptin"? OR IT="Linagliptin"? OR SH="Linagliptin"?)	146
17	7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16	2074
18	6 AND 17	1364
19	DT=Randomized controlled trial	338147
20	Randomized/(Ti;Ab;Ct;TE) OR CT=Randomized?	402275
21	Placebo/(Ti;Ab;Ct;TE) OR CT=Placebo?	160072
22	19 OR 20 OR 21	632724
23	18 AND 22	388
24	(23) AND (LA=ENGLISH OR LA=GERMAN) AND pps=Mensch	378

Tabelle 4-111: DIMDI Suchstrategie Cochrane 1950 bis heute: am 04. Februar 2013

Datenbankname	Cochrane
Suchoberfläche	DIMDI
Datum der Suche	04.02.2013
Zeitsegment	CCTR93: 1948 bis 2013 Woche 6
Suchfilter	Aufgrund der geringen Trefferzahl wurde kein Filter verwendet

#	Suchbegriffe	Ergebnis
1	CCTR93	67775
2	FT=Diabetes mellitus OR (CT D "Diabetes mellitus" OR UT="Diabetes mellitus" OR IT="Diabetes mellitus" OR SH="Diabetes mellitus")	16502
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")	8872
4	FT=NIDDM OR (CT D "NIDDM" OR UT="NIDDM" OR IT="NIDDM" OR SH="NIDDM")	6876
5	FT=T2DM OR (CT D "T2DM" OR UT="T2DM" OR IT="T2DM" OR SH="T2DM")	368
6	2 OR 3 OR 4 OR 5	16584
7	FT=Dipeptidyl#Peptidase#4 Inhibitor? OR (CT D "Dipeptidyl#"Peptidase#"4 Inhibitor"? OR UT="Dipeptidyl#"Peptidase#"4 Inhibitor"? OR IT="Dipeptidyl#"Peptidase#"4 Inhibitor"? OR SH="Dipeptidyl#"Peptidase#"4 Inhibitor"?)	126
8	FT=Dipeptidyl#Peptidase 4 Inhibitor? OR (CT D "Dipeptidyl#"Peptidase 4 Inhibitor"? OR UT="Dipeptidyl#"Peptidase 4 Inhibitor"? OR IT="Dipeptidyl#"Peptidase 4 Inhibitor"? OR SH="Dipeptidyl#"Peptidase 4 Inhibitor"?)	126
9	FT=Dipeptidyl-Peptidase IV Inhibitor? OR (CT D "Dipeptidyl- Peptidase IV Inhibitor"? OR UT="Dipeptidyl-Peptidase IV Inhibitor"? OR IT="Dipeptidyl-Peptidase IV Inhibitor"? OR SH="Dipeptidyl-Peptidase IV Inhibitor"?)	134
10	CT=("DIPEPTIDYL PEPTIDASE 4 INHIBITORS"; "DIPEPTIDYL PEPTIDASE IV INHIBITORS"; "DIPEPTIDYL-PEPTIDASE 4 INHIBITORS"; "DIPEPTIDYL-PEPTIDASE IV INHIBITORS")	126
11	FT=DPP-4 Inhibitor? OR (CT D "DPP-4 Inhibitor"? OR UT="DPP-4 Inhibitor"? OR IT="DPP-4 Inhibitor"? OR SH="DPP-4 Inhibitor"?)	49
12	FT=DPP-IV Inhibitor? OR (CT D "DPP-IV Inhibitor"? OR UT="DPP-IV Inhibitor"? OR IT="DPP-IV Inhibitor"? OR SH="DPP-IV Inhibitor"?)	10
13	FT=Gliptin? OR (CT D "Gliptin"? OR UT="Gliptin"? OR IT="Gliptin"? OR SH="Gliptin"?)	126

14	FT=Sitagliptin? OR (CT D "Sitagliptin"? OR UT="Sitagliptin"? OR IT="Sitagliptin"? OR SH="Sitagliptin"?)	80
15	FT=Vildagliptin? OR (CT D "Vildagliptin"? OR UT="Vildagliptin"? OR IT="Vildagliptin"? OR SH="Vildagliptin"?)	81
16	FT=Linagliptin? OR (CT D "Linagliptin"? OR UT="Linagliptin"? OR IT="Linagliptin"? OR SH="Linagliptin"?)	15
17	7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16	216
18	6 AND 17	171
19	18 AND LA=(ENGLISH; GERMAN) AND pps=Mensch	171

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

Die Studienregistersuche für das zu bewertende Arzneimittel wurde für Abschnitt 4.3.1.1.3 durchgeführt.

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	21.1.2013
Suchstrategie	Diabetes mellitus [Condition] AND Saxagliptin OR Onglyza OR BMS-477118 [Intervention] AND (Phase 2 OR Phase 3 OR Phase 4) [Phase]
Treffer	40

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	21.1.2013
Suchstrategie	Diabetes mellitus [Condition] AND Saxagliptin OR Onglyza OR BMS-477118 [Intervention] AND ALL [Recruitment stats]
Treffer	38

Studienregistersuche zur Durchführung eines indirekten Vergleiches mit Insulin

Die Studienregistersuche zur Durchführung eines indirekten Vergleiches mit Insulin wurde für Abschnitt 4.3.2.1.1.3 durchgeführt.

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	25.2.2013
Suchstrategie	Diabetes mellitus [Condition] AND Insulin AND (Glimepiride OR Glibenclamide OR Glipizide OR Gliclazide OR Gliquidone OR sulfonyleurea) [Intervention] AND (Phase 2 OR Phase 3 OR Phase 4) [Phase]
Treffer	78

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 Insulin und Sulfonylharnstoff eine eigene Suche durchgeführt werden:

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 25.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Insulin AND Glimpiride [Intervention]
AND ALL [Recruitment status]
Treffer 21

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 25.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Insulin AND Glibenclamide [Intervention]
AND ALL [Recruitment status]
Treffer 9

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 25.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Insulin AND Glipizide [Intervention]
AND ALL [Recruitment status]
Treffer 2

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 25.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Insulin AND Gliclazide [Intervention]
AND ALL [Recruitment status]
Treffer 5

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 25.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Insulin AND Gliquidone [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 25.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Insulin AND sulfonyleurea [Intervention]
AND ALL [Recruitment status]
Treffer 7

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 25.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Insulin AND sulphonyleurea [Intervention]
AND ALL [Recruitment status]
Treffer 2

Studienregistersuche zur Durchführung eines indirekten Vergleiches mit einer Kombinationstherapie aus Sulfonylharnstoffen und DPP-4-Inhibitoren

Die Studienregistersuche zur Durchführung eines indirekten Vergleiches mit einer Kombinationstherapie aus Sulfonylharnstoffen und DPP-4-Inhibitoren wurde für Abschnitt 4.3.2.1.1.3 durchgeführt.

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

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	12.2.2013
Suchstrategie	Diabetes mellitus [Condition] AND (Glimepiride OR Glibenclamide OR Glipizide OR Gliclazide OR Gliquidone OR sulfonylurea) AND (Sitagliptin OR Vildagliptin OR Linagliptin) [Intervention] AND (Phase 2 OR Phase 3 OR Phase 4) [Phase]
Treffer	62

Da die Webseite des WHO ICTRP nur sehr einfache Suchanfragen und kaum Verknüpfungen zwischen Suchbegriffen zulässt, musste für jede Kombination aus Sulfonylharnstoff und DPP-4-Inhibitor eine eigene Suche durchgeführt werden:

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	12.2.2013
Suchstrategie	Diabetes mellitus [Condition] AND Glimepiride AND Sitagliptin [Intervention] AND ALL [Recruitment status]
Treffer	23

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 12.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Glimepiride AND Vildagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 7

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 12.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Glimepiride AND Linagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 5

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 12.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Glibenclamide AND Sitagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 1

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 12.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Glibenclamide AND Vildagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 12.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Glibenclamide AND Linagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 12.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Glipizide AND Sitagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 9

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 12.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Glipizide AND Vildagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 12.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Glipizide AND Linagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 13.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Gliclazide AND Sitagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 3

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 13.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Gliclazide AND Vildagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 4

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 13.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Gliclazide AND Linagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 13.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Gliquidone AND Sitagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 13.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Gliquidone AND Vildagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 13.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND Gliquidone AND Linagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 13.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND sulfonyleurea AND Sitagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 2

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 13.2.2013
Suchstrategie Diabetes mellitus [Condition]
AND sulfonyleurea AND Vildagliptin [Intervention]
AND ALL [Recruitment status]
Treffer 0

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	13.2.2013
Suchstrategie	Diabetes mellitus [Condition] AND sulfonylurea AND Linagliptin [Intervention] AND ALL [Recruitment status]
Treffer	0
Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	13.2.2013
Suchstrategie	Diabetes mellitus [Condition] AND sulphonylurea AND Sitagliptin [Intervention] AND ALL [Recruitment status]
Treffer	1
Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	13.2.2013
Suchstrategie	Diabetes mellitus [Condition] AND sulphonylurea AND Vildagliptin [Intervention] AND ALL [Recruitment status]
Treffer	0
Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	13.2.2013
Suchstrategie	Diabetes mellitus [Condition] AND sulphonylurea AND Linagliptin [Intervention] AND ALL [Recruitment status]
Treffer	0

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 4-112: Im Volltext ausgeschlossene nicht relevante Publikationen für das zu bewertende Arzneimittel

	Im Volltext ausgeschlossene Publikationen	Grund
1	Saxagliptin (Onglyza) for type 2 diabetes. 2009. Medical Letter on Drugs and Therapeutics, 51, 85-86.	Keine RCT
2	FAIRMAN, K. A. & CURTISS, F. R. 2009. Call for comparative effectiveness research: lowering A1c with sitagliptin, saxagliptin, or cinnamon. United States.	Keine RCT
3	FREEMAN, M. K. & FREEMAN, M. K. 2011. Efficacy and safety of linagliptin (Tradjenta) in adults with type-2 diabetes mellitus. P and T, 36, 807-812.	Keine RCT
4	GOLDENBERG, M. M. & GOLDENBERG, M. M. 2009. Pharmaceutical approval update. P and T, 34, 569-574.	Keine RCT
5	SCHEEN, A. J., RADERMECKER, R. P., SCHEEN, A. J. & RADERMECKER, R. P. 2010. Addition of incretin therapy to metformin in type 2 diabetes. <i>The Lancet</i> , 375, 1410-1412.	Keine RCT
6	SCHMITZ, O. & SCHMITZ, O. 2008. The GLP-1 concept in the treatment of type 2 diabetes - Still standing at the gate of dawn? <i>Journal of Clinical Endocrinology and Metabolism</i> , 93, 375-377.	Keine RCT
7	WHALEN, K., ONGE, E. S., WHALEN, K. & ONGE, E. S. 2010. Saxagliptin (onglyza) for type 2 diabetes mellitus. <i>American Family Physician</i> , 81, 1483-1484.	Keine RCT

Tabelle 4-113: Im Volltext ausgeschlossene nicht relevante Publikationen für die zweckmäßige Vergleichstherapie

	Im Volltext ausgeschlossene Publikationen	Grund
Literatursuche für indirekten Vergleich Sulfonylharnstoff+Insulin oder Insulin-Monotherapie versus Sulfonylharnstoff		
1	DUCKWORTH, W., ABRAIRA, C., MORITZ, T., REDA, D., EMANUELE, N., REAVEN, P. D., ZIEVE, F. J., MARKS, J., DAVIS, S. N., HAYWARD, R., WARREN, S. R., GOLDMAN, S., MCCARREN, M., VITEK, M. E., HENDERSON, W. G., HUANG, G. D., DUCKWORTH, W., HAYWARD, R., GOLDMAN, S., WARREN, S. R., MCCARREN, M., VITEK, M. E., HENDERSON, W. G., HUANG, G. D., ABRAIRA, C., MORITZ, T., REDA, D., EMANUELE, N., REAVEN, P. D., ZIEVE, F. J., MARKS, J. & DAVIS, S. N. 2009. Glucose control and vascular complications in veterans with type 2 diabetes. <i>New England Journal of Medicine</i> , 360, 129-139.	falsche Intervention (weder Sulfonylharnstoff+Insulin noch Insulin-Monotherapie)
2	FRIEDEWALD, W. T., BUSE, J. B., BIGGER, J. T., BYINGTON, R. P., CUSHMAN, W. C., GERSTEIN, H. C., GINSBERG, H. N., GOFF, D. C., JR., GRIMM, R. H., JR., ISMAIL-BEIGI, F., PROBSTFIELD, J. L., SIMONS-MORTON, D. G., BUSE, J. B., BIGGER, J. T., BYINGTON, R. P., FRIEDEWALD, W. T., CUSHMAN, W. C., GERSTEIN, H. C., GINSBERG, H. N., GOFF, D. C., JR., GRIMM, R. H., JR., ISMAIL-BEIGI, F., PROBSTFIELD, J. L. & SIMONS-MORTON, D. G. 2008. Effects of intensive glucose lowering in type 2 diabetes. <i>New England Journal of Medicine</i> , 358, 2545-2559.	falsche Intervention (weder Sulfonylharnstoff+Insulin noch Insulin-Monotherapie), falsche Vergleichstherapie
3	GÆDE, P., LUND-ANDERSEN, H., PARVING, H. H., PEDERSEN, O., GÆDE, P., LUND-ANDERSEN, H., PARVING, H. H. & PEDERSEN, O. 2008. Effect of a multifactorial intervention on mortality in type 2 diabetes. <i>New England Journal of Medicine</i> , 358, 580-591.	falsche Intervention (weder Sulfonylharnstoff+Insulin noch Insulin-Monotherapie), falsche Vergleichstherapie
4	PEACEY, S. R., ROBINSON, R., BEDFORD, C., HARRIS, N. D., MACDONALD, I. A., HOLMAN, R. R. & HELLER, S. R. 2000. Does the choice of treatment for type 2 diabetes affect the physiological response to hypoglycemia? <i>UNITED STATES</i> .23, 1022-3.	zu geringe Studiendauer
5	RUTTER, M. K., NESTO, R. W., RUTTER, M. K. & NESTO, R. W. 2010. The BARI 2D study: A randomised trial of therapies for type 2 diabetes and coronary artery disease. <i>Diabetes and Vascular Disease Research</i> , 7, 69-72.	falsche Vergleichstherapie
6	ROTH, J., PROUT, T. E., GOLDFINE, I. J. D., ROTH, J., PROUT, T. E. & GOLDFINE, I. J. D. 1971. Sulfonylureas: effects in vivo and in vitro. <i>AnnInternJMed</i> , 75, 607-621.	keine RCT
7	RULL, J. A., GONZALEZ-MILLAN, H., QUIBRERA, R., GARCIA-VIVEROS, M. & LOZANO-CASTANEDA, O. 1970. Over-all therapeutic usefulness of glibenclamide, a new hypoglycemic sulfonylurea. <i>Diabetes</i> , 19, 264-70.	falsche Intervention (weder Sulfonylharnstoff+Insulin noch Insulin-Monotherapie)
Literatursuche für indirekten Vergleich Insulin-Monotherapie versus Sulfonylharnstoff		
8	CHEN, H. S., HSIAO, L. C., WU, T. E., LEE, S. H., JAP, T. S., LIN, H. D., CHEN, H. S., HSIAO, L. C., WU, T. E., LEE, S. H., JAP, T.	falsche Vergleichstherapie

	Im Volltext ausgeschlossene Publikationen	Grund
	S. & LIN, H. D. 2008. Beneficial effects of insulin on glycemic control and β -cell function in newly diagnosed type 2 diabetes with severe hyperglycemia after short-term intensive insulin therapy. <i>Diabetes Care</i> , 31, 1927-1932.	
9	VAN DER WAL, P. S., DRAEGER, K. E., MARTINI, C. & HEINE, R. J. 1995. Influence of oral glimepiride treatment on the interaction between glycaemia and insulin secretion in NIDDM patients.	falsche Intervention (keine Insulin-Monotherapie)
10	KOSHIBA, K., NOMURA, M., NAKAYA, Y. & ITO, S. 2006. Efficacy of glimepiride on insulin resistance, adipocytokines, and atherosclerosis. <i>The journal of medical investigation : JMI</i> , 53, 87-94.	keine patientenrelevanten Endpunkte
11	MERTZ, D. P. & BÜRVENICH, K. 1985. [Different levels of lipoprotein Lp(a) in adult diabetic patients depending on the therapeutic regimen] Unterschiedliche Werte von Lipoprotein Lp(a) bei erwachsenen Diabetikern in Abhängigkeit von der Therapieform. <i>Klinische Wochenschrift</i> , 63, 648-50.	keine RCT
12	POP-BUSUI, R., LU, J., LOPES, N., JONES, T. L. & INVESTIGATORS, B. D. 2009. Prevalence of diabetic peripheral neuropathy and relation to glycemic control therapies at baseline in the BARI 2D cohort. <i>Journal of the peripheral nervous system : JPNS</i> , 14, 1-13.	keine RCT
13	HAPP, J., BEYER, J., NEST, E., FRÖHLICH, A., ALTHOFF, P. H. & SCHÖFFLING, K. 1974. [Blood sugar, serum insulin, nonesterified fatty acids and somatotropin in daily profile in adult diabetics under monotherapy with various sulfonylureas] Blutzucker, Seruminsulin, unveresterte Fettsäuren und Wachstumshormon im Tagesprofil bei Erwachsenen-Diabetikern unter Monotherapie mit verschiedenen Sulfonylharnstoffen. <i>Arzneimittel-Forschung</i> , 24, 1228-34.	falsche Intervention (keine Insulin-Monotherapie)
14	RODIER, M., COLETTE, C., GOUZES, C., MICHEL, F., CRASTES PAULET A, D. E., MONNIER, L., RODIER, M., COLETTE, C., GOUZES, C., MICHEL, F., CRASTES PAULET A, D. E. & MONNIER, L. 1995. Effects of insulin therapy upon plasma lipid fatty acids and platelet aggregation in NIDDM with secondary failure to oral antidiabetic agents. <i>Diabetes Research and Clinical Practice</i> , 28, 19-28.	falsche Vergleichstherapie
15	PEACOCK, I., WATTS, R., SELBY, C. & TATTERSALL, R. B. 1987. Serum C-peptide after 6 months on glibenclamide remains higher than during insulin treatment. <i>Diabetes research (Edinburgh, Scotland)</i> , 6, 57-9.	keine patientenrelevanten Endpunkte
16	BAGG, W., PLANK, L. D., GAMBLE, G., DRURY, P. L., SHARPE, N., BRAATVEDT, G. D., BAGG, W., PLANK, L. D., GAMBLE, G., DRURY, P. L., SHARPE, N. & BRAATVEDT, G. D. 2001. The effects of intensive glycaemic control on body composition in patients with type 2 diabetes. <i>Diabetes, Obesity and Metabolism</i> , 3, 410-416.	falsche Intervention (keine Insulin-Monotherapie)
17	BAGG, W., WHALLEY, G. A., GAMBLE, G., DRURY, P. L., SHARPE, N. & BRAATVEDT, G. D. 2001. Effects of improved glycaemic control on endothelial function in patients with type 2	falsche Intervention (keine Insulin-Monotherapie)

	Im Volltext ausgeschlossene Publikationen	Grund
	diabetes. Internal medicine journal, 31, 322-8.	
18	HAUPT, E., PETZOLDT, R., TREIBER, J. & SCHÖFFLING, K. 1982. [Changes in blood glucose, insulin, cortisol, growth hormone and glucagon induced by sulfonylures] Sulfonylharnstoffinduzierte Veränderungen von Blutglukose, Insulin, Kortisol, Wachstumshormon und Glukagon. Die Medizinische Welt, 33, 71-7.	falsche Intervention (keine Insulin-Monotherapie)
19	SEINO, Y., IKEDA, M., KURAHACHI, H., TAMINATO, T., SAKURAI, H., GOTO, Y., INOUE, Y., KADOWAKI, S., MORI, K. & IMURA, H. 1978. Failure of suppress plasma glucagon concentrations by orally administered glucose in diabetic patients after treatment. Diabetes, 27, 1145-50.	falsche Patientenpopulation (gesunde vs. Typ-2-Diabetes)
20	HOULDEN, R., ROSS, S., HARRIS, S., YALE, J. F., SAURIOL, L. & GERSTEIN, H. C. 2007. Treatment satisfaction and quality of life using an early insulinization strategy with insulin glargine compared to an adjusted oral therapy in the management of Type 2 diabetes: The Canadian INSIGHT Study. Diabetes Research and Clinical Practice, 78, 254-258.	falsche Intervention (Insulin-Analoga; OAD gemischt)
21	TAYEK, J. & TAYEK, J. 2008. SUR receptor activity vs. incidence of hypoglycaemia and cardiovascular mortality with sulphonylurea therapy for diabetics. Diabetes, Obesity and Metabolism, 10, 1128-1129.	keine RCT
22	LIEDTKE, R. K., SORGER, M., MERK, F. & VETTER, H. 1990. [Transdermal administration of insulin in type II diabetics. Results of a clinical pilot study] Transdermale Applikation von Insulin bei Typ-II-Diabetikern. Ergebnisse einer klinischen Pilotstudie. Arzneimittel-Forschung, 40, 884-6.	falsche Intervention (Insulin-Pflaster)
23	JAIN, S. K., NAGI, D. K., SLAVIN, B. M., LUMB, P. J., YUDKIN, J. S., JAIN, S. K., NAGI, D. K., SLAVIN, B. M., LUMB, P. J. & YUDKIN, J. S. 1993. Insulin therapy in Type 2 diabetic subjects suppresses plasminogen activator inhibitor (PAI-1) activity and proinsulin-like molecules independently of glycaemic control. Diabetic Medicine, 10, 27-32.	keine passenden Endpunkte
24	JOYA-GALEANA, J., FERNANDEZ, M., CERVERA, A., REYNA, S., GHOSH, S., TRIPLITT, C., MUSI, N., DEFRONZO, R. A. & CERSOSIMO, E. 2011. Effects of insulin and oral anti-diabetic agents on glucose metabolism, vascular dysfunction and skeletal muscle inflammation in type 2 diabetic subjects. Diabetes/metabolism research and reviews, 27, 373-82.	falsche Vergleichstherapie
25	HAMELBECK, H., KLEIN, W., ZOLTOBROCKI, M. & SCHÖFFLING, K. 1982. [Glibenclamide-insulin combination in the management of secondary failure of sulfonyl-urea medication] Glibenclamid-Insulin-Kombinationsbehandlung beim Sekundärversagen der Sulfonylharnstoff-Therapie. Deutsche medizinische Wochenschrift (1946), 107, 1581-3.	falsche Intervention (keine Insulin-Monotherapie), falsche Vergleichstherapie
26	TRISCHITTA, V., ITALIA, S., MAZZARINO, S. & RIDDLE, M. C. 1992. Addition of insulin or metformin after secondary failure to glyburide. Annals of internal medicine, 117, 46.	falsche Intervention (keine Insulin-Monotherapie)

	Im Volltext ausgeschlossene Publikationen	Grund
27	CHOUDHARY, P., LONNEN, K., EMERY, C. J., FREEMAN, J. V., MCLEOD, K. M., HELLER, S. R., CHOUDHARY, P., LONNEN, K., EMERY, C. J., FREEMAN, J. V., MCLEOD, K. M. & HELLER, S. R. 2011. Relationship between interstitial and blood glucose during hypoglycemia in subjects with type 2 diabetes. <i>Diabetes Technology and Therapeutics</i> , 13, 1121-1127.	keine RCT
28	EELEY, E. A., STRATTON, I. M., HADDEN, D. R., TURNER, R. C. & HOLMAN, R. R. 1996. UKPDS 18: estimated dietary intake in type 2 diabetic patients randomly allocated to diet, sulphonylurea or insulin therapy. UK Prospective Diabetes Study Group. <i>Diabetic medicine : a journal of the British Diabetic Association</i> , 13, 656-62.	keine patientenrelevanten Endpunkte.
29	1985. U.K. prospective diabetes study. II. Reduction in HbA1c with basal insulin supplement, sulphonylurea, or biguanide therapy in maturity-onset diabetes. A multicenter study. <i>Diabetes</i> , 34, 793-8.	keine patientenrelevanten Endpunkte
30	SACHSE, G., WILLMS, B., SACHSE, G. & WILLMS, B. 1978. Effect of the alpha -glucosidase-inhibitor Bay-g-5421 on the carbohydrate metabolism of maturity-onset diabetics and insulin dependent diabetics. <i>Diabetologia</i> , 15, 302.	falsche Intervention (keine Insulin-Monotherapie)
31	RACHMAN, J., PAYNE, M. J., LEVY, J. C., BARROW, B. A., HOLMAN, R. R. & TURNER, R. C. 1998. Changes in amylin and amylin-like peptide concentrations and beta-cell function in response to sulphonylurea or insulin therapy in NIDDM. <i>Diabetes Care</i> , 21, 810-6.	keine patientenrelevanten Endpunkte
32	RIVELLESE, A. A., PATTI, L., ROMANO, G., INNELLI, F., DI MARINO, L., ANNUZZI, G., IAVICOLI, M., CORONEL, G. A. & RICCARDI, G. 2000. Effect of insulin and sulphonylurea therapy, at the same level of blood glucose control, on low density lipoprotein subfractions in type 2 diabetic patients. <i>The Journal of clinical endocrinology and metabolism</i> , 85, 4188-92.	keine patientenrelevanten Endpunkte
33	PANAHLOO, A., MOHAMED-ALI, V., ANDRÉS, C., DENVER, A. E. & YUDKIN, J. S. 1998. Effect of insulin versus sulphonylurea therapy on cardiovascular risk factors and fibrinolysis in type II diabetes. <i>Metabolism: clinical and experimental</i> , 47, 637-43.	keine patientenrelevanten Endpunkte
34	MD ISA, S. H., NAJIHAH, I., NAZAIMOON, W. M., KAMARUDIN, N. A., UMAR, N. A., MAT, N. H. & KHALID, B. A. 2006. Improvement in C-reactive protein and advanced glycosylation end-products in poorly controlled diabetics is independent of glucose control. <i>Diabetes Research and Clinical Practice</i> , 72, 48-52.	zu kurze Studiendauer, keine patientenrelevanten Endpunkte.
35	CARROLL, M. F., HARDY, K. J., BURGE, M. R. & SCHADE, D. S. 2002. Frequency of the dawn phenomenon in type 2 diabetes: Implications for diabetes therapy. <i>Diabetes technology & therapeutics</i> , 4, 595-605.	zu geringe Studiendauer, keine patientenrelevanten Endpunkte.
36	MATSUMOTO, K., NAKAMURA, H., UEKI, Y., TOMINAGA, T. & MIYAKE, S. 2001. Correction of hyperglycaemia reduces insulin resistance and serum soluble E-selectin levels in patients with Type 2 diabetes mellitus. <i>Diabetic medicine : a journal of the British Diabetic Association</i> , 18, 224-8.	keine RCT
37	USHAKOVA, O., SOKOLOVSKAYA, V., MOROZOVA, A., VALEEVA, F., ZANOZINA, O., SAZONOVA, O., ZHADANOVA,	falsche Vergleichstherapie

	Im Volltext ausgeschlossene Publikationen	Grund
	E., STARCEVA, M., KAZAKOVA, E., SAIFULLINA, M., SHAPIRO, I., TARASOV, A., AL-TAYAR, B. & STARKOVA, N. 2007. Comparison of biphasic insulin aspart 30 given three times daily or twice daily in combination with metformin versus oral antidiabetic drugs alone in patients with poorly controlled type 2 diabetes: a 16-week, randomized, open-label, parallel-group trial conducted in russia. <i>Clinical therapeutics</i> , 29, 2374-84.	
38	AAS, A. M., BERGSTAD, I., THORSBYT, P. M., JOHANNESSEN, Ø., SOLBERG, M., BIRKELANDT, K. I., AAS, A. M., BERGSTAD, I., THORSBYT, P. M., JOHANNESSEN, Ø., SOLBERG, M. & BIRKELANDT, K. I. 2005. An intensified lifestyle intervention programme may be superior to insulin treatment in poorly controlled Type 2 diabetic patients on oral hypoglycaemic agents: Results of a feasibility study. <i>Diabetic Medicine</i> , 22, 316-322.	falsche Vergleichstherapie
39	BOYD, K., ROGERS, C., BOREHAM, C., ANDREWS, W. J. & HADDEN, D. R. 1992. Insulin, glibenclamide or metformin treatment for non insulin dependent diabetes: heterogenous responses of standard measures of insulin action and insulin secretion before and after differing hypoglycaemic therapy. <i>Diabetes research (Edinburgh, Scotland)</i> , 19, 69-76.	keine patientenrelevanten Endpunkte.
40	RACHMAN, J., LEVY, J. C., BARROW, B. A., MANLEY, S. E. & TURNER, R. C. 1997. Relative hyperproinsulinemia of NIDDM persists despite the reduction of hyperglycemia with insulin or sulfonylurea therapy. <i>Diabetes</i> , 46, 1557-62.	keine patientenrelevanten Endpunkte.
41	RIDDLE, M. C. & RIDDLE, M. C. 1986. Bedtime administration of intermediate-acting insulin for non-insulin-dependent diabetes mellitus. <i>Practical Cardiology</i> , 12, 65-66.	keine RCT
42	ALVARSSON, M., SUNDKVIST, G., LAGER, I., BERNTORP, K., FERNQVIST-FORBES, E., STEEN, L., ÖRN, T., HOLBERG, M. A., KIRKSÆTHER, N., GRILL, V., ALVARSSON, M., GRILL, V., SUNDKVIST, G., LAGER, I., BERNTORP, K., FERNQVIST-FORBES, E., STEEN, L., ÖRN, T., HOLBERG, M. A. & KIRKSÆTHER, N. 2008. Effects of insulin vs. glibenclamide in recently diagnosed patients with type 2 diabetes: A 4-year follow-up. <i>Diabetes, Obesity and Metabolism</i> , 10, 421-429.	Erhebungszeitpunkte stimmen nicht überein (4 Jahre). Ein indirekter Vergleich ist nicht möglich.
43	BIRKELAND, K. I., RISHAUG, U., HANSEN, K. F. & VAALER, S. 1996. NIDDM: a rapid progressive disease. Results from a long-term, randomised, comparative study of insulin or sulphonylurea treatment. <i>Diabetologia</i> , 39, 1629-33.	Erhebungszeitpunkte stimmen nicht überein (42 Monate). Ein indirekter Vergleich ist nicht möglich.
44	FORST, T., ERIKSSON, J. W., STROTMANN, H. J., BAI, S., BRUNELLE, R., GULLIYA, K. S., GACK, S. & GUDAT, U. 2003. Metabolic effects of mealtime insulin lispro in comparison to glibenclamide in early type 2 diabetes. <i>Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association</i> , 111, 97-103.	falsche Intervention (keine Insulin-Monotherapie)
45	GALIC, E., VRTOVEC, M., BOZIKOV, V., SCHWARZENHOFER, M. & MILICEVIC, Z. 2005. The impact of the timing of Humalog Mix25 injections on blood glucose fluctuations in the postprandial period in elderly patients with type 2 diabetes. <i>Medical science monitor : international medical journal of</i>	zu geringe Studiendauer

	Im Volltext ausgeschlossene Publikationen	Grund
	experimental and clinical research, 11, 87-92.	
46	LI, H., LI, W., GU, Y., HAN, Y., WANG, J., XU, B., LI, F., YU, Y., WU, G., LI, H., LI, W., GU, Y., HAN, Y., WANG, J., XU, B., LI, F., YU, Y. & WU, G. 2009. Comparison of continual insulin or secretagogue treatment in type 2 diabetic patients with alternate insulin-secretagogue administration. <i>Diabetes Research and Clinical Practice</i> , 84, 158-162.	Erhebungszeitpunkte stimmen nicht überein (38 Monate). Ein indirekter Vergleich ist nicht möglich.
47	TOVI, J., SVANBORG, E., NILSSON, B. Y. & ENGFELDT, P. 1998. Diabetic neuropathy in elderly Type 2 diabetic patients: effects of insulin treatment. <i>Acta neurologica Scandinavica</i> , 98, 346-53.	Doppelpublikation zu Tovi and Engfeldt, 1998
48	TOVI, J., INGEMANSSON, S. O. & ENGFELDT, P. 1998. Insulin treatment of elderly type 2 diabetic patients: effects on retinopathy. <i>Diabetes & metabolism</i> , 24, 442-7.	Doppelpublikation zu Tovi and Engfeldt, 1998
49	TOVI, J., THEOBALD, H. & ENGFELDT, P. 1996. Effect of metabolic control on 24-h ambulatory blood pressure in elderly non-insulin-dependent diabetic patients. <i>Journal of Human Hypertension</i> , 10, 589-94.	Doppelpublikation zu Tovi and Engfeldt, 1998
50	HERZ, M., SUN, B., MILICEVIC, Z., ERICKSON, P., FÖVÈNYI, J., GRZYWA, M. & PELIKANOVA, T. 2002. Comparative efficacy of preprandial or postprandial Humalog Mix75/25 versus glyburide in patients 60 to 80 years of age with type 2 diabetes mellitus. <i>Clinical therapeutics</i> , 24, 73-86.	zu geringe Studiendauer
51	ROACH, KOLEDOVA, E., METCALFE, S., HULTMAN, C., MILICEVIC, Z. & ROMANIA/RUSSIA MIX25 STUDY, G. 2001. Glycemic control with Humalog Mix25 in type 2 diabetes inadequately controlled with glyburide. <i>Clinical therapeutics</i> , 23, 1732-44.	zu geringe Studiendauer
52	TURNER, R. C., MANN, J. I., ICETON, G. & ET AL. 1983. UK prospective study of therapies of maturity-onset diabetes. I. Effect of diet, sulphonylurea, insulin or biguanide therapy on fasting plasma glucose and body weight over one year.	Doppelpublikation zu Turner et al. 1998
53	HOLMAN, R. R., CULL, C. A., FOX, C., TURNER, R. C., HOLMAN, R. R., CULL, C. A., FOX, C. & TURNER, R. C. 1995. United Kingdom prospective diabetes study (UKPDS) 13: Relative efficacy of randomly allocated diet, sulphonylurea, insulin, or metformin in patients with newly diagnosed non-insulin dependent diabetes followed for three years. <i>British medical journal</i> , 310, 83-88.	Erhebungszeitpunkte stimmen nicht überein (3 Jahre). Ein indirekter Vergleich ist nicht möglich.
54	1995. U.K. prospective diabetes study 16. Overview of 6 years' therapy of type II diabetes: a progressive disease. U.K. Prospective Diabetes Study Group. <i>Diabetes</i> , 44, 1249-58.	Erhebungszeitpunkte stimmen nicht überein (6 Jahre). Ein indirekter Vergleich ist nicht möglich.
55	TURNER, R., CULL, C. & HOLMAN, R. 1996. United Kingdom Prospective Diabetes Study 17: a 9-year update of a randomized, controlled trial on the effect of improved metabolic control on complications in non-insulin-dependent diabetes mellitus. <i>Annals of internal medicine</i> , 124, 136-45.	Erhebungszeitpunkte stimmen nicht überein (9 Jahre). Ein indirekter Vergleich ist nicht möglich.
56	TURNER, R. & TURNER, R. 1998. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS	Erhebungszeitpunkte stimmen nicht überein (15 Jahre). Ein indirekter Vergleich ist nicht

	Im Volltext ausgeschlossene Publikationen	Grund
	33). Lancet, 352, 837-853.	möglich.
57	TURNER, R. & TURNER, R. 1998. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). Lancet, 352, 854-865.	Erhebungszeitpunkte stimmen nicht überein (15 Jahre). Ein indirekter Vergleich ist nicht möglich.
Literatursuche für indirekten Vergleich Insulin+Sulfonylharnstoff vs. Sulfonylharnstoff		
58	FONSECA, V. A. & FONSECA, V. A. 1981. Sulphonylurea and insulin combination in the treatment of diabetes. British medical journal, 283, 797.	falsche Patientenpopulation
59	MU, P. W., CHEN, Y. M., LU, H. Y., WEN, X. Q., ZHANG, Y. H., XIE, R. Y., SHU, J., WANG, M. M. & ZENG, L. Y. 2012. Effects of a combination of oral anti-diabetes drugs with basal insulin therapy on beta-cell function and glycaemic control in patients with newly diagnosed type 2 diabetes. Diabetes/metabolism research and reviews, 28, 236-40.	falsche Intervention (keine Kombinationstherapie Insulin+Sulfonylharnstoff), falsche Vergleichstherapie
60	PAPA, G., FEDELE, V., CHIAVETTA, A., LORENTI, I., LEOTTA, C., LUCA, S., RABUZZO, A. M., PIRO, S., ALAGONA, C., SPADARO, L., PURRELLO, F. & PEZZINO, V. 2008. Therapeutic options for elderly diabetic subjects: open label, randomized clinical trial of insulin glargine added to oral antidiabetic drugs versus increased dosage of oral antidiabetic drugs. Acta diabetologica, 45, 53-9.	falsche Intervention (keine Kombinationstherapie Insulin+Sulfonylharnstoff), falsche Vergleichstherapie
61	EMANUELE, N., AZAD, N., ABRAIRA, C., HENDERSON, W., COLWELL, J., LEVIN, S., NUTTALL, F., COMSTOCK, J., SAWIN, C., SILBERT, C., MARCOVINA, S., LEE, H. S., EMANUELE, N., SILBERT, C., MARCOVINA, S., LEE, H. S., AZAD, N., ABRAIRA, C., HENDERSON, W., COLWELL, J., LEVIN, S., NUTTALL, F., COMSTOCK, J. & SAWIN, C. 1998. Effect of intensive glycemic control on fibrinogen, lipids, and lipoproteins: Veterans Affairs Cooperative Study in Type II Diabetes Mellitus. Archives of internal medicine, 158, 2485-2490.	falsche Intervention (keine Kombinationstherapie Insulin+Sulfonylharnstoff)
62	CASNER, P. R. 1988. Insulin-glyburide combination therapy for non-insulin-dependent diabetes mellitus: a long-term double-blind, placebo-controlled trial. Clinical pharmacology and therapeutics, 44, 594-603.	falsche Vergleichstherapie
63	SOTANIEMI, E. A., VIERIMAA, E., HUUPPONEN, R., KARVONEN, I., VUOTI, M. J. & RYTÖMAA, K. 1990. Insulin and sulphonylurea in the therapy of type 2 diabetes. Diabetes Research and Clinical Practice, 8, 243-51.	falsche Vergleichstherapie
64	STUART, C. A., GILKISON, C. R., CARLSON, R. F., STUART, C. A., GILKISON, C. R. & CARLSON, R. F. 1997. Effect of adding a sulphonylurea in patients with non-insulin-dependent diabetes mellitus previously well controlled with insulin. Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists, 3, 344-8.	falsche Intervention (keine Kombinationstherapie Insulin+Sulfonylharnstoff)
65	FALKO, J. M. & OSEI, K. 1985. Combination insulin/glyburide therapy in type II diabetes mellitus. Effects on lipoprotein metabolism and glucoregulation. The American journal of medicine, 79, 92-101.	falsche Intervention (keine Kombinationstherapie Insulin+Sulfonylharnstoff)

	Im Volltext ausgeschlossene Publikationen	Grund
66	GARG, A. & GRUNDY, S. M. 1994. Cholestyramine therapy for dyslipidemia in non-insulin-dependent diabetes mellitus. A short-term, double-blind, crossover trial. <i>Annals of internal medicine</i> , 121, 416-22.	falsche Intervention (keine Kombinationstherapie Insulin+Sulfonylharnstoff)
67	VAN DIJK, J. W., MANDERS, R. J. F., TUMMERS, K., BONOMI, A. G., STEHOUWER, C. D. A., HARTGENS, F., VAN LOON, L. J. C., VAN DIJK, J. W., MANDERS, R. J. F., TUMMERS, K., BONOMI, A. G., STEHOUWER, C. D. A., HARTGENS, F. & VAN LOON, L. J. C. 2012. Both resistance- and endurance-type exercise reduce the prevalence of hyperglycaemia in individuals with impaired glucose tolerance and in insulin-treated and non-insulin-treated type 2 diabetic patients. <i>Diabetologia</i> , 55, 1273-1282.	falsche Intervention (keine Kombinationstherapie Insulin+Sulfonylharnstoff)
68	SCHMIDT, F. H., KLUJKO, J., KUHNLE, H. F. & REITER, J. 1986. The mode of action of the combination glibenclamide/insulin in type-II-diabetics. <ORIGINAL> UNTERSUCHUNGEN ZUM WIRKUNGSMECHANISMUS DER KOMBINIERTEN GABE VON GLIBENCLAMID UND INSULIN BEI TYP-II-DIABETIKERN MIT SEKUNDARVERSAGEN AUF DIE ORALE BEHANDLUNG.	falsche Vergleichstherapie
69	TURNER, R. C., CULL, C. A., FRIGHI, V. & HOLMAN, R. R. 1999. Glycemic control with diet, sulfonylurea, metformin, or insulin in patients with type 2 diabetes mellitus. Progressive requirement for multiple therapies (UKPDS 49). <i>Journal of the American Medical Association</i> , 281, 2005-2012.	keine patientenrelevanten Endpunkte
70	PFÜTZNER, A., LORRA, B., ABDOLLAHIA, M. R., KANN, P. H., MATHIEU, D., PEHNERT, C., OLIGSCHLEGER, C., KAISER, M. & FORST, T. 2006. The switch from sulfonylurea to preprandial short- acting insulin analog substitution has an immediate and comprehensive beta-cell protective effect in patients with type 2 diabetes mellitus. <i>Diabetes technology & therapeutics</i> , 8, 375-84.	keine patientenrelevanten Endpunkte
71	BABU, A., MEHTA, A., GUERRERO, P., CHEN, Z., MEYER, P. M., KOH, C. K., ROBERTS, R., SCHAIDER, J. & FOGELFELD, L. 2009. Safe and simple emergency department discharge therapy for patients with type 2 diabetes mellitus and severe hyperglycemia. <i>Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists</i> , 15, 696-704.	keine patientenrelevanten Endpunkte
72	KAWAMORI, R., IWAMOTO, Y., KADOWAKI, T., IWASAKI, M., KIM, S. W., WOO, J. T., BAIK, S. H. & YOON, K. H. 2009. Effects of insulin glulisine as mono- or add-on therapy in patients with type 2 diabetes mellitus. <i>Diabetes, obesity & metabolism</i> , 11, 900-9.	falsche Vergleichstherapie
73	BEBAKAR, W. M. W., CHOW, C. C., KADIR, K. A., SUWANWALAIKORN, S., VAZ, J. A., BECH, O. M., BEBAKAR, W. M. W., CHOW, C. C., KADIR, K. A., SUWANWALAIKORN, S., VAZ, J. A. & BECH, O. M. 2007. Adding biphasic insulin aspart 30 once or twice daily is more efficacious than optimizing oral antidiabetic treatment in patients with type 2 diabetes. <i>Diabetes, Obesity and Metabolism</i> , 9, 724-732.	keine patientenrelevanten Endpunkte
74	LUNDERSHAUSEN, R., ORBAN, S., PISSAREK, D. & PANZRAM, G. 1987. [Long-term effect of combination	keine patientenrelevanten Endpunkte

	Im Volltext ausgeschlossene Publikationen	Grund
	glibenclamide-insulin treatment in the secondary failure of sulfonylurea therapy--results of a one-year double blind study] Langzeiteffekt der Glibenclamid-Insulin-Kombinationsbehandlung beim Sekundärversagen der Sulfonylharnstofftherapie--Ergebnisse einer einjährigen Doppelblindstudie. Wiener klinische Wochenschrift, 99, 603-8.	
75	CASTILLO, M. J., SCHEEN, A. J. & LEFEBVRE, P. J. 1994. Metabolic clearance rate of insulin in type 2 diabetic patients treated with combined insulin and sulfonylurea therapy. Revista española de fisiología, 50, 27-34.	falsche Intervention (keine Kombinationstherapie Insulin+Sulfonylharnstoff)
76	CASTILLO, M., SCHEEN, A. J., PAOLISSO, G., LEFEBVRE, P. J., CASTILLO, M., SCHEEN, A. J., PAOLISSO, G. & LEFEBVRE, P. J. 1987. The addition of glipizide to insulin therapy in Type-II diabetic patients with secondary failure to sulfonylureas is useful in the presence of a significant residual insulin secretion. Acta endocrinologica, 116, 364-372.	falsche Vergleichstherapie
77	HOLMAN, R. R., STEEMSON, J. & TURNER, R. C. 1987. Sulphonylurea failure in type 2 diabetes: treatment with a basal insulin supplement. Diabetic medicine : a journal of the British Diabetic Association, 4, 457-62.	zu geringe Studiendauer

Tabelle 4-114: Im Volltext ausgeschlossene nicht relevante Publikationen für die alternative Vergleichstherapie

	Im Volltext ausgeschlossene Publikationen	Grund
1	DEROSA, G., MAFFIOLI, P., FERRARI, I., MEREU, R., RAGONESI, P. D., QUERCI, F., FRANZETTI, I. G., GADALETA, G., CICCARELLI, L., PICCINNI, M. N., DANDELO, A., SALVADEO, S. A. T., DEROSA, G., PICCINNI, M. N., DANDELO, A., SALVADEO, S. A. T., MAFFIOLI, P., FERRARI, I., MEREU, R., RAGONESI, P. D., QUERCI, F., FRANZETTI, I. G., GADALETA, G. & CICCARELLI, L. 2010. Effects of one year treatment of vildagliptin added to pioglitazone or glimepiride in poorly controlled type 2 diabetic patients. Hormone and Metabolic Research, 42, 663-669.	Keine placebokontrollierte Studie. Die Patienten erhielten im Kontrollarm Pioglitazon.
2	SERRA, D., HE, Y. L., BULLOCK, J., RIVIERE, G. J., BALEZ, S., SCHWARTZ, S., WANG, Y., LIGUEROS-SAYLAN, M., JARUGULA, V. & DOLE, W. P. 2008. Evaluation of pharmacokinetic and pharmacodynamic interaction between the dipeptidyl peptidase IV inhibitor vildagliptin, glyburide and pioglitazone in patients with Type 2 diabetes. International Journal of Clinical Pharmacology and Therapeutics, 46, 349-64.	Zu geringe Behandlungsdauer. In der cross-over Studie pro Phase max. 28 Tage Behandlungsdauer

Anhang 4-D: Liste der abgebrochenen Studien

Benennen Sie in der nachfolgenden Tabelle die abgebrochenen Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-115 (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						
Nicht zutreffend						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Nicht zutreffend						
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 4-116 (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						
D1680C 00003 (SAVO R-TIMI 53)	nein	ja	nein	nein	ja [NCT01107886] (ClinicalTrials.gov , 2013c) [EUCTR2009- 017358-10-HU] (ICTRP Search Portal, 2013b)	nein
CV1811 42	nein	ja	ja	nein	ja [NCT01527747] (ClinicalTrials.gov , 2013d)	nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
--						
<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>						

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 4-117 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 4-117 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-117 (Anhang): Studiendesign und -methodik für Studie CV181040

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p><u>Kurzzeitphase (24 Wochen):</u></p> <p>Primäres Ziel Veränderung des HbA1c-Wertes (ausgehend vom Ausgangswert) bei einer Behandlung mit Saxagliptin 2 mg +Glibenclamid gegenüber Saxagliptin 5 mg+Glibenclamid oder Placebo +Glibenclamid bei Patienten mit Typ-2-Diabetes, die eine unzureichende Blutzuckerkontrolle unter Einnahme einer submaximalen Dosis an Sulfonylharnstoff für mindestens 2 Monate haben.</p> <p>Sekundäre Ziele Veränderung des FPG-Wertes (Nüchternblutzucker, ausgehend vom Ausgangswert) Anteil an Probanden, die eine HbA1c-Wert <7.0% erreicht haben Veränderung des AUC (ausgehend vom Ausgangswert) von 0 bis 180 min des Plasmablutzuckers während des OGTT</p> <p><u>Anschließende Langzeitphase (12 Monate):</u></p> <p>Sicherheit und Verträglichkeit der Saxagliptin +Glibenclamid-Dosen während der gesamten Studie Sicherheit und Verträglichkeit des Saxagliptin+Glibenclamid+Metformin-Dosen während der Langzeitphase (12 Monate)</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Internationale, multizentrische, randomisierte, parallele, doppelblinde, placebokontrollierte, dreiarmlige Phase-III-Studie; Zuteilungsverhältnis 1:1:1
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><u>Einschlusskriterien:</u></p> <p>Patienteneinwilligung Patienten mit Typ-2-Diabetes mellitus, die eine submaximale Dosis an Sulfonylharnstoff für mindestens 2 Monate eingenommen haben HbA1C $\geq 7.5\%$ und $\leq 10\%$ bei der ersten Visite HbA1C $\geq 7\%$ und MFPG ≥ 140 mg/dL oder MFWBG ≥ 131 mg/dL oder FPG ≥ 140 mg/dL</p>

Item ^a	Charakteristikum	Studieninformation
		<p>bei Randomisierung</p> <p>Nüchtern C-Peptid Konzentration ≥ 1.0 ng/mL</p> <p>BMI ≤ 40 kg/m²</p> <p>Männer und Frauen ≥ 18 und ≤ 77</p> <p>Frauen dürfen nicht stillen und nicht schwanger sein, zudem müssen Frauen im gebärfähigen Alter entsprechende Vorkehrungen zur Verhütung treffen</p> <p>72 Stunden vor Studienmedikation muss ein negativer Schwangerschaftstest vorliegen</p> <p><u>Ausschlusskriterien:</u></p> <p>Frauen im gebärfähigen Alter, die keine akzeptablen Verhütungsmittel nehmen können oder möchten, Schwangere, Stillende</p> <p>Schlecht kontrollierte Diabetes</p> <p>Vorerkrankungen an diabetischer Ketoazidose, hypoglykämischem nicht-ketotischem Koma, fortschreitendem Nierenversagen, Alkohol- oder Drogenmissbrauch im letzten Jahr, Hämoglobinopathie, infektiösen Lebererkrankungen,</p> <p>Insulintherapie innerhalb eines Jahres zum Screeningzeitpunkt</p> <p>Signifikante kardiale Vorerkrankung</p> <p>Chronische oder wiederholte periodische Corticoide</p> <p>psychische Probleme</p> <p>Immunsystemgeschwächte</p> <p>Blut- oder Plasmaspende in den letzten 3 Monaten</p> <p>Teilnahme an einer anderen Studie innerhalb der letzten 30 Tage (ausgehend von der 1. Medikation)</p> <p>Verfassung, die nach Meinung des Prüfarztes für die Teilnahme an der Studie ein Risiko darstellt oder eine erfolgreiche Beendigung der Studie nicht gewährleisten kann</p> <p>Lebererkrankungen</p> <p>Serumkreatinin ≥ 2.0 mg/dL</p> <p>Kreatininkinase ≥ 3x größer als die obere Normgrenze</p> <p>Anämie</p> <p>Absolute Lymphozytenanzahl ≤ 1000 Zellen/mm³</p> <p>Blutblättchenanzahl $\leq 140,000$ Zellen/μL</p> <p>Abnormaler TSH- und T4-Wert</p> <p>Kontraindikation bei Saxagliptin, Glibenclamid und Metformin</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Anwendung einer Antihypoglykämie-Therapie innerhalb 12 Wochen vor Screening</p> <p>Behandlung mit Cytochrom P450 3A4-Inhibitor oder Induktor</p> <p>Frühere Behandlung mit Saxagliptin oder einem DPP-4-Inhibitor</p> <p>Vollzug einer Haftstrafe oder in psychischer oder physischer Behandlung befindend</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>Multizentrische, internationale Studie, durchgeführt von 88 Investoren an 86 Zentren in 13 Ländern:</p> <p>USA (28 Zentren), Brasilien (10 Zentren), Mexiko (10 Zentren), Argentinien (6 Zentren), Israel (6 Zentren), Korea (6 Zentren), Südafrika (4 Zentren), Peru (4 Zentren), Philippinen (4 Zentren), Taiwan (3 Zentren), Hong Kong (2 Zentren), Puerto Rico (2 Zentren), Singapur (1 Zentrum)</p>
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p>Saxagliptin 2.5 mg+Glibenclamid (offen) 7.5 mg</p> <p>versus</p> <p>Saxagliptin 5 mg+Glibenclamid (offen) 7.5 mg</p> <p>versus</p> <p>Placebo+Glibenclamid (verblindet) 2.5 mg + Glibenclamid (offen) 7.5 mg</p> <p>Hierbei kann Saxagliptin und Glibenclamid ab-/aufitriert werden</p> <p>Falls eine andauernde Hyperglykämie besteht, kann bis zu 2500 mg Metformin zusätzlich eingenommen werden, danach darf Glibenclamid nicht mehr ab-/aufitriert werden</p>
6	Zielkriterien	
6a	<p>Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten</p>	<p><u>Kuzzeitphase (nach 24 Wochen):</u></p> <p>Primäres Ziel</p> <p>Veränderung des HbA1c-Wertes (ausgehend vom Ausgangswert) bei einer Behandlung mit Saxagliptin 2 mg +Glibenclamid gegenüber Saxagliptin 5 mg+Glibenclamid oder Placebo +Glibenclamid bei Patienten mit Typ-2-Diabetes, die eine unzureichende Blutzuckerkontrolle unter Einnahme einer submaximalen Dosis an Sulfonylharnstoff für mindestens 2 Monate haben.</p> <p>Sekundäre Ziele</p> <p>Veränderung des FPG-Wertes (ausgehend vom Ausgangswert)</p> <p>Anteil an Probanden, die eine HbA1c-Wert</p>

Item ^a	Charakteristikum	Studieninformation
		<p><7.0% erreicht haben</p> <p>Veränderung des AUC (ausgehend vom Ausgangswert) von 0 bis 180 min des PBZ während des OGTT</p> <p>Veränderung der Funktion der β-Zellen (gemessen anhand HOMA-2β)</p> <p>Veränderung der Insulinresistenz (gemessen anhand HOMA-2IR)</p> <p>Veränderung von nüchtern Glukagon, Insulin und C-Peptid, sowie deren Konzentrationen und Konzentrationsdifferenzen nach 0, 30, 60, 120 und 180 min mit einem OGTT</p> <p>Veränderung der Insulinsensitivität und β-Zellfunktion ausgehend von den vorher aufgelisteten Messungen des OGTT</p> <p>Veränderung des BMI, Taillenumfanges und Gewichts</p> <p>Anteil der Probanden unterteilt in Kategorien nach angegebenem glykämischen Wert</p> <p>Veränderung in Prozent des Total-C, LDH-C, HDL-C, TG und FFA</p> <p>Veränderung von hs-CRP, PAI-1, Fibrinogen und IL-6</p> <p>Anteil der Probanden zu den Zeitpunkten 4,6,8,12,16,20 und 24 Wochen, die eine vorher festgelegte Schranke des Glykämiewertes überschreiten und daher Metformin einnehmen</p> <p>Anzahl unerwünschter und schwerwiegender unerwünschter Ereignisse sowie unerwünschter Ereignisse von speziellem Interesse</p> <p>Laborwerte</p> <p>Elektrokardiogramm</p> <p>Vitalwerte</p> <p><u>Langzeitphase (12 Monate):</u></p> <p>Sicherheit und Verträglichkeit der Saxagliptin +Glibenclamid-Dosen während der gesamten Studie</p> <p>Sicherheit und Verträglichkeit des Saxagliptin+Glibenclamid+Metformin-Dosen während der Langzeitphase (12 Monate)</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Überlegenheitsstudie, Paarvergleich der unterschiedlichen Saxagliptin-Dosen mit Placebo-Gruppe über Dunnett's Adjustierungstest, pro Test Signifikanzniveau von 0.027, um das globale Signifikanzniveau

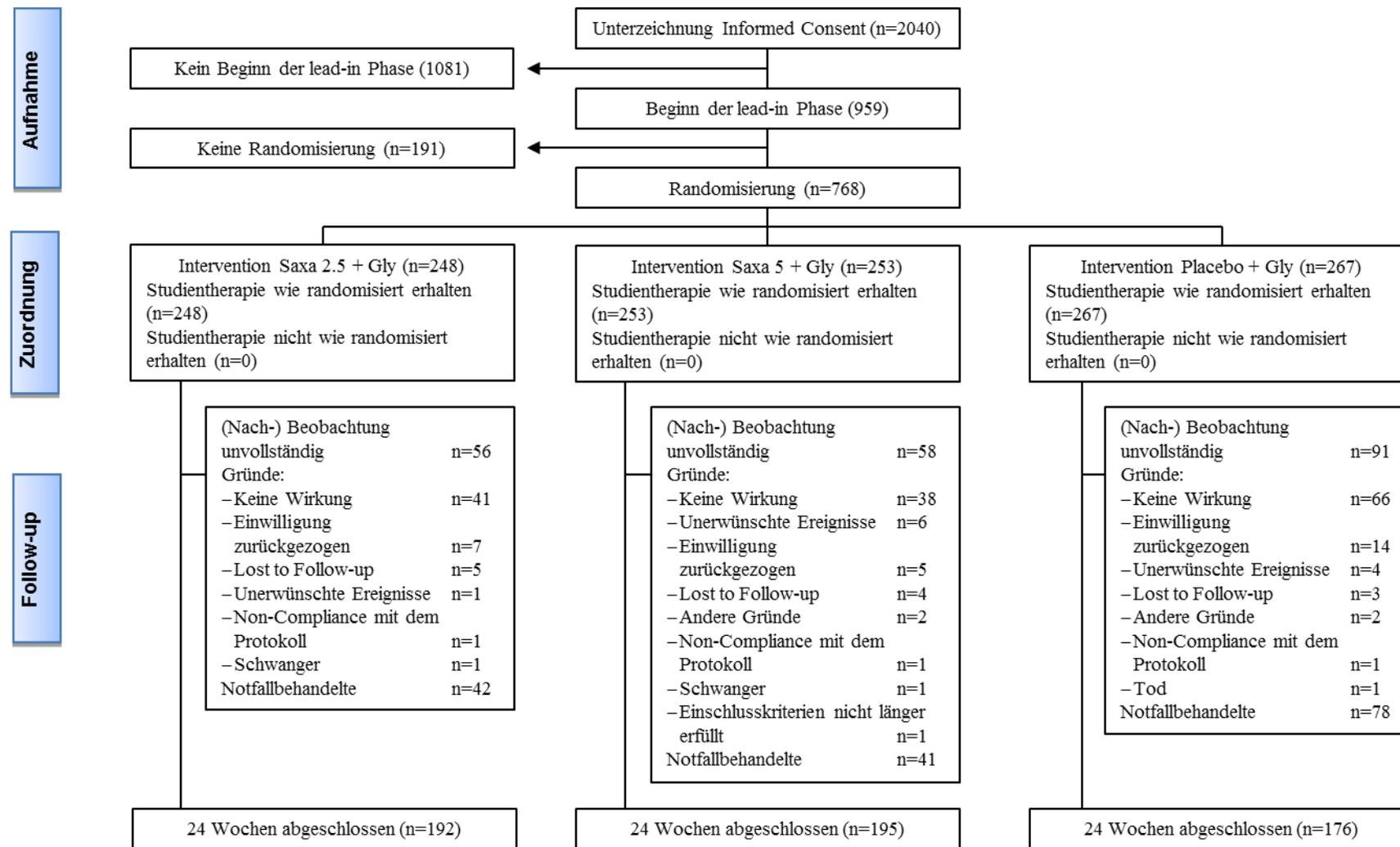
Item ^a	Charakteristikum	Studieninformation
		von 0.05 einzuhalten, Standardabweichung von 1,2% aus Phase II Studie abgeleitet, Unterschied des gemittelten HBA1C von 0.4%, Power von 90%, 15% Drop-out-Rate -> 780 Probanden sind notwendig (260 pro Gruppe)
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Vorzeitiger Studienabbruch: Rückzug der Patienteneinwilligung Schwangerschaft Vollzug einer Haftstrafe, Entdeckung einer psychischen oder physischen Erkrankung MFPG \geq 270 mg/dL oder MFWBG \geq 249 mg/dL nach 2 Wochen Auftreten von unerwünschten Ereignissen
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Patienten erhalten einen E-code (Randomisierungsnummer) nach dem unterschriebenen Patienteneinwilligung.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Blockrandomisierung
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	Zuteilung zur Studienmedikation wird durch das interactive voice response system (IVRS) durchgeführt (unabhängig und zentral).
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	AstraZeneca stellt Randomisierungsnummern sowie deren Kit-Nummern zur Verfügung. Zentren rufen bei Eintritt in Einführungsphase direkt beim IVRS an. IVRS führt Zuteilung durch.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a)verblindet b)verblindet c)verblindet Doppel-dummy-Medikation
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Identisches Design der Darreichungsform, verblindete Verpackung mit Abreisetikett zur Endblindung in Notfällen
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<u>Analysepopulationen:</u> Lead-in-Analyse Set: Alle Patienten, die mindestens eine Dosis der Placebo-Lead-in Medikation eingenommen haben

Item ^a	Charakteristikum	Studieninformation
		<p>Randomisierte Analyse Set: Alle Patienten mit Randomisierungscode, die mindestens eine randomisierte Dosis der Wirksubstanz erhielten.</p> <p>Auswertbare Analyse Set: Teilpopulation des randomisierten Analyse Sets, die keine Protokollverletzungen aufweisen, welche den primären Endpunkt signifikant beeinflussen</p> <p>Behandelte Analyse Set: Alle Patienten, die mindestens eine Dosis der doppelblinden Studienmedikation während der Kurzzeitphase eingenommen haben</p> <p>Rescued Analyse Set: Alle Patienten, die mindestens eine Dosis des Notfallmedikaments während der Kurzzeit und Langzeitphase erhalten haben</p> <p>Responder Analyse Set: Teilpopulation des randomisierten Analyse Sets, die ohne Notfallmedikation 24 Wochen vollendet haben, deren HBA1C-Wert <7% ist und die Langzeitphase begonnen haben</p> <p>Kurzzeit Analyse Set: Teilpopulation des randomisierten Analyse Sets, die ohne Notfallmedikation 24 Wochen vollendet haben und die Langzeitphase begonnen haben</p> <p><u>Methodik:</u></p> <p>Wirksamkeit: Zeitanalyse bis zur ersten Notfallmedikation oder Abbruch wegen fehlender Wirksamkeit Repeated-Measures Analyse über ein gemischtes-Modell um die mittlere Veränderung zu analysieren ANCOVA-Modell mit Behandlungsgruppe als Effekt und Anfangswert als Kovariable für die Analyse der stetigen Endpunkte Punktschätzer und 95%-KI für die Veränderung jeder Saxagliptin-Gruppe gegenüber Placebo Prozentuale mittlere Veränderung zwischen Woche 24 und Woche in der Langzeitphase mit KI aus Fieller's Theorem Häufigkeitstabellen für kategoriale Endpunkte mit exakten 95%-KI Kaplan-Meier Plot für Zeit bis Abbruch wegen Mangel an Wirksamkeit Sicherheit Sensitivitätsanalyse Auswertung der unerwünschten Ereignisse,</p>

Item ^a	Charakteristikum	Studieninformation
		Labortestergebnisse, Elektrokardiogramme und Vitalparameter Kaplan-Meier-Plots
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Unerwünschte Ereignisse: Alter, Geschlecht und Rasse
-	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	a) 248 vs. 253 vs. 267 b) 248 vs. 253 vs. 267 c) 248 vs. 253 vs. 267 davon 42 vs. 41 vs. 78 ohne Notfallbehandlung mit Metformin
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow Chart
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	04/2006 – 09/2008
14b	Informationen, warum die Studie endete oder beendet wurde	Regulär beendet
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

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



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

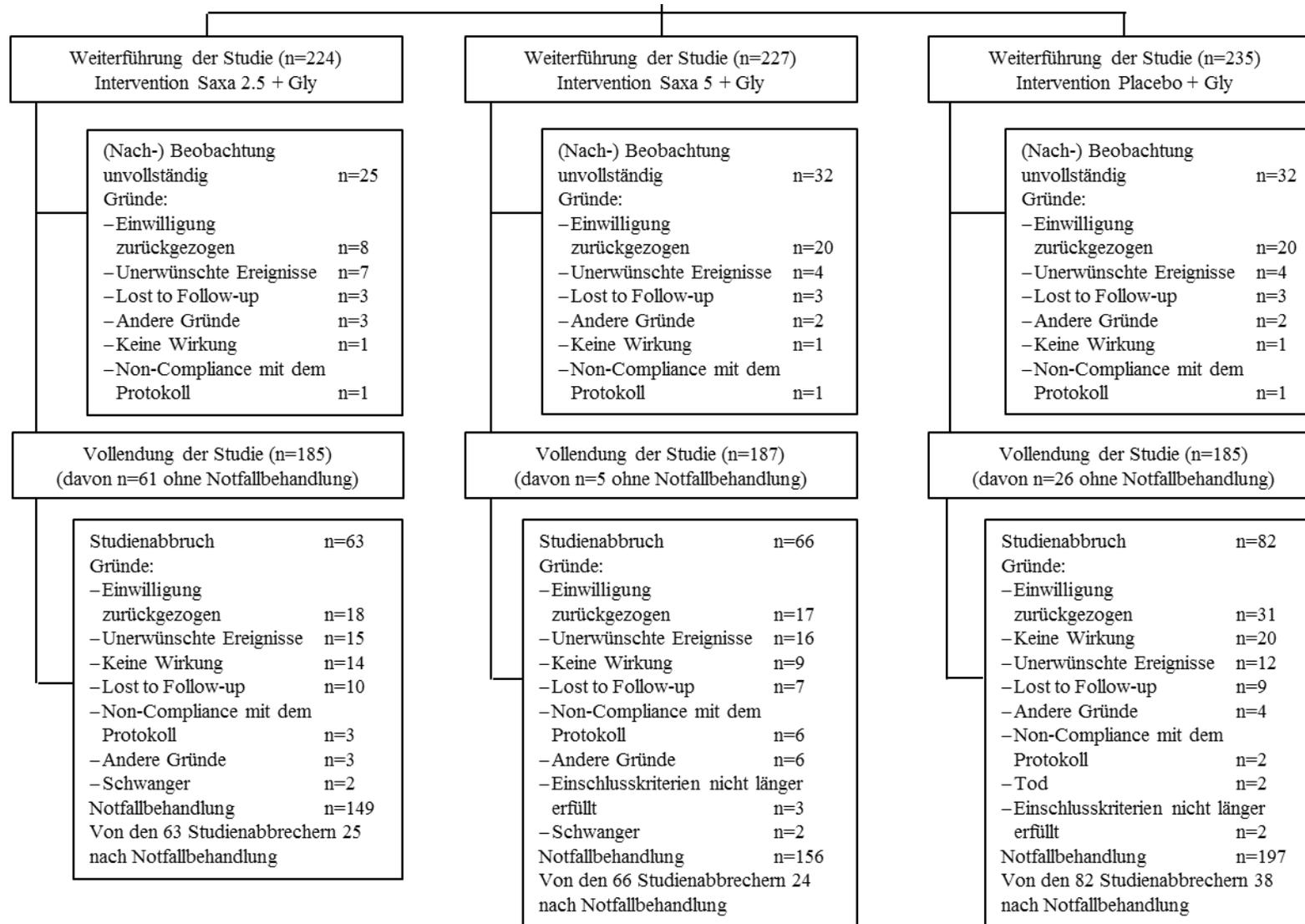


Abbildung 15: Flussdiagramm Studiendesign und-methodik für CV181040

Tabelle 4-118 (Anhang): Studiendesign und -methodik für Studie Shank 1995

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Vergleich des Effekts der Insulingabe vor dem Zubettgehen (BI) in Kombination mit Sulfonylharnstoffen tagsüber (DSU) mit BI alleine oder DSU alleine bei nicht-insulinabhängigen Diabetikern, die unzureichend mit Sulfonylharnstoffen kontrolliert wurden. Bewertung der langfristigen klinischen Wirksamkeit der Kombination aus BI und DSU. Untersuchung des Mechanismus, durch den die Kombinationstherapie aus BI und DSU die glykämische Kontrolle verbessert.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	RCT, doppelblind, placebokontrolliert, mit offener Phase
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht zutreffend
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Nicht-insulinabhängige Diabetiker mit mindestens zweimal FPG <7,77 mmol/L (140 mg/L) und einer Vorbehandlung mit einer maximalen Dosis mit einem anderen Sulfonylharnstoff als Glipizid (Glyburid 18, Chlorpropamid 11, Tolazamid 1) <u>Einschlusskriterien:</u> Patienten zwischen 33 und 70 Jahren Keine andere Medikation, die einen Einfluss auf die Glukosestoffwechsel hat Keine frühere Behandlung mit Insulin Keine andere Haupterkrankung als Diabetes Freiwillige und unterschriebene Einverständniserklärung
4b	Studienorganisation und Ort der Studiendurchführung	Nicht berichtet
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Phase 1 (2 Monate): n=36 Alle Patienten erhielten 20 mg Glipizid (offen) zweimal täglich, oral verabreicht

Item ^a	Charakteristikum	Studieninformation
		<p>Phase 2 (3 Monate): n=30 Randomisierung in 3 Gruppen (je n=10): 1) DSU - kein BI: nur Glipizid 2) BI - kein DSU: NPH Insulin beim Zubettgehen und Placebo Glipizid 3) BI – DSU: NPH Insulin beim Zubettgehen und Glipizid Glipizid/Placebo: 20 mg, zweimal täglich, oral Insulin: Injektionspen (Novolin Pen, Novo Nordisk, Princeton, NJ) mit humanen NPH, Insulin wurde zwischen 22 Uhr und 23:30 Uhr verabreicht; Initialdosis war 5 U/1,73 m² und wurde bis zu 20 U/1,73 m² (low-dose) heraufgesetzt. Es wurde kein Placebo NPH Insulin verwendet, da eine Lösung aus Kortisonacetat das einzige überzeugende Placebo ist. Selbstkontrolle des Blutzuckerspiegels (SMBG) vor Frühstück und Abendessen mit AkkuChek II und Merlin electronic logbook. Alle zwei Wochen wurde die FPG bestimmt um die Zuverlässigkeit der Selbstkontrolle jedes Patienten zu überprüfen.</p> <p>Phase 3 (3 Monate): n=18 (betrifft nicht Gruppe 1) Die Dosen von BI wurden erhöht, um FPG unter 6,66 mmol/L (120 mg/dL) zu halten und um nicht mehr als ein hypoglykämisches Ereignis zu verursachen Alle zwei Wochen fanden eine klinische Untersuchung und die Bestimmung der FPG statt.</p> <p>Phase 4 (6 Monate): n=25 Offene Glipizid- und Insulin-Phase: Gruppe 1 +Insulin, n=7 Gruppe 2 +Glipizid, n=9 Gruppe 3 bleibt, n=9 Monatliches klinisches Follow-up, FPG-Bestimmung und Titration von Insulin</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	FPG, HbA1c, OGTT, 24h- Glukoseprofil, SMBG, Basales HGP <u>Details:</u> Metabolisches Profil nach dem Ende der Phase 1, 2 und 3: - Bestimmung von Glukose, Insulin und der Konzentration der freien Fettsäuren (FFA) aus

Item ^a	Charakteristikum	Studieninformation
		<p>Plasma (alle 30-60 Minuten von 7:30-24:00 Uhr am Tag der Aufnahme und von 7:30 und 8:00 am Tag danach)</p> <ul style="list-style-type: none"> - HbA1c-Bestimmung - Nüchternlipid-Profil im Serum (Gesamt-Cholesterin, HDL-Cholesterin, LDL-Cholesterin, Triglyceride) - 24h-Sammelurin zur Bestimmung von Glukose, Protein, Albumin und Stickstoff <p>Während euglykämischen Insulin -Clamps (am Ende aller 3 Phasen):</p> <ul style="list-style-type: none"> - Gesamt Glukoseaufnahme - Glukose Oxidation - nicht-oxidative Glukose-Abgabe - HPG-Abfall während der Insulin-Infusion <p>Basalwerte und Werte während euklykämischen Insulin Clamp (am Ende aller 3 Phasen):</p> <ul style="list-style-type: none"> - Plasma FFA Konzentration (Plasma und während euglykämischen Insulin Clamps) - Lipidoxidation (Plasma und während euglykämischen Insulin Clamps) <p>Hypoglykämische Episoden pro Woche (für BI-DSU und BI-kein DSU) in Phase 3</p> <p>Gewichtszunahme, Blutdruck und Albumin und Proteinausscheidung im Urin während Phase 2 und 3</p> <p>Phase 4: FPG, HbA1c, SMBG</p>
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 Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Rückzug der Einverständniserklärung aus persönlichen Gründen (Phase 2 und 3) und aufgrund von systematischen Schädigungen (Phase 3)
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Es wurde ein Randomisierungscode benutzt (Roerig)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der	Nicht berichtet

Item ^a	Charakteristikum	Studieninformation
	Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	
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) verblindet (Phase 2) b) verblindet (Phase 2) c) nicht verblindet
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Im postabsorptiven Zustand entspricht die Rate der Anflutung von Glukose (HPG) der Rate der Elimination der Glucose und wurde berechnet als $[3\text{-}^3\text{H}]\text{Glukose Infusion}$ ($\mu\text{Ci}/\text{min}$) geteilt durch $[3\text{-}^3\text{H}]\text{Glukose-spezifische Aktivität}$ ($\mu\text{Ci}/\text{min}$).</p> <p>Während des Insulin Clamps herrschen nicht-stationäre Zustände, und die Raten der Anflutung und Elimination der Glukose wurden mittels Steele's Equation berechnet.</p> <p>Die Rate von HPG wurde durch die Subtraktion der exogenen Rate der Glukoseinfusion von der vom Indikator abgeleiteten Rate der gesamten Anflutung der Glukose im Körper.</p> <p>Die Abgabe der nichtoxidativen Glukose wurde durch Subtraktion der Rate der Glukose-Oxidation (bestimmt durch indirekte Kalorimetrie) von der Rate der Elimination der Glukose im Körper bestimmt.</p> <p>A priori Vergleiche wurden mit Varianzanalysen für wiederholte Messungen durchgeführt. Ein statistisch signifikanter Unterschied lag bei $p < 0,05$ vor. Ergebnisse wurden als Mittelwert\pmSE dargestellt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur	

Item ^a	Charakteristikum	Studieninformation
	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	a) 10 vs. 10 vs. 10 b) 10 vs. 10 vs. 10 c) 7vs. 9 vs 9
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow Chart Phase 1: Drei Patienten zogen Einverständniserklärung aus persönlichen Gründen zurück, zwei wurden aufgrund von Nichteinhaltung ausgeschlossen, ein Patient war lost to follow-up Phase 2: Ein Patient (BI - kein DSU) zeigte symptomatische Schädigungen, ein Patient zog aus persönlichen Gründen, unabhängig von der Therapie, Einverständniserklärung zurück (BI-DSU) Phase 3: keine verlorenen oder ausgeschlossenen Patienten Phase 4: Drei Patienten haben nicht die 6-monatige Phase begonnen (Grund nicht bekannt), 25 Patienten traten in Phase 4 ein
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	Regulär beendet
a: nach CONSORT 2010.		

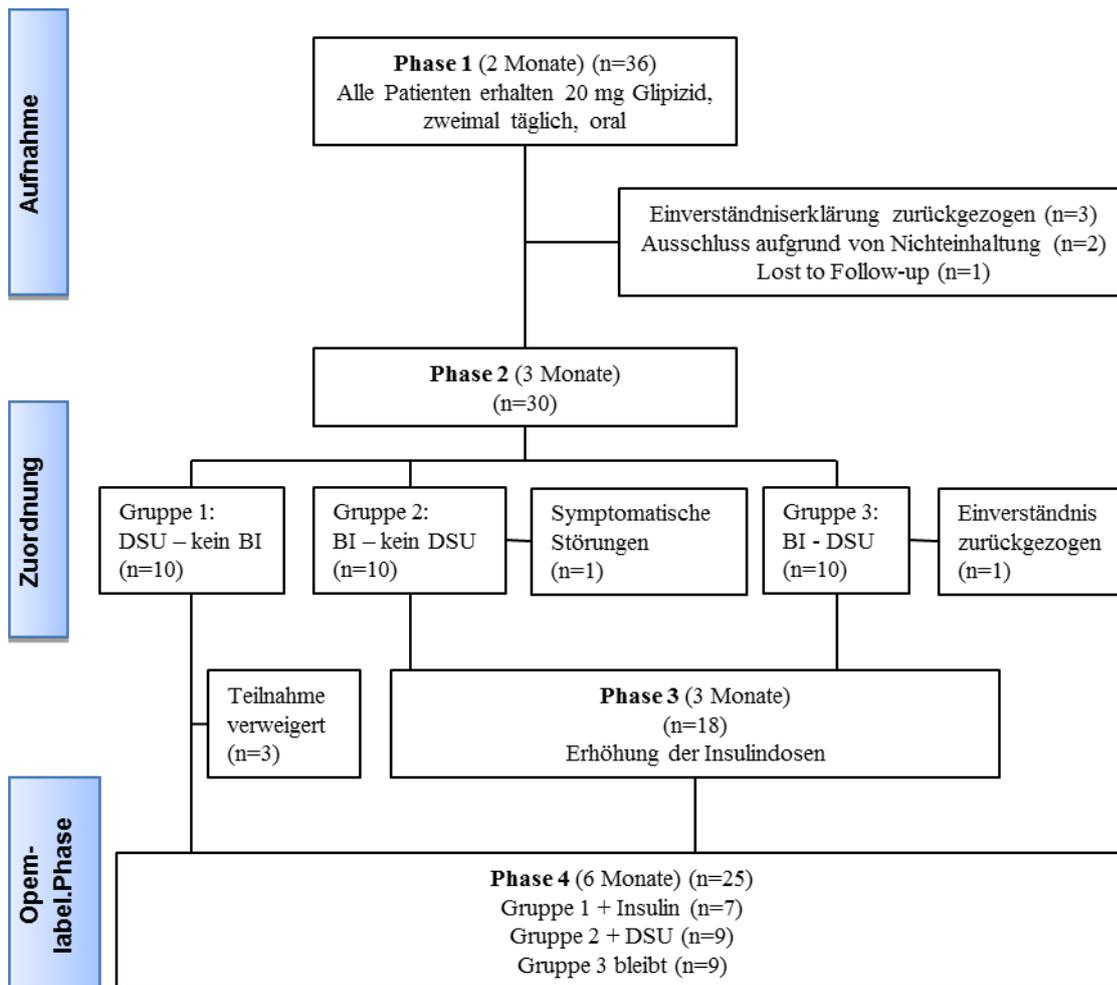


Abbildung 16: Flussdiagramm Studiendesign und-methodik für Shank 1995

Tabelle 4-119 (Anhang): Studiendesign und -methodik für Studie von Birkeland 1994

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Langzeiteffekt von Insulin gegen Sulfonylharnstoff bei Diabetes mellitus Typ-2 Patienten Blutzuckerkontrolle, Herz-Kreislauf Faktoren, Insulinrestistenz, Mikroalbuminurie, Lp(a) Lipoprotein, Triglyceride (TG), gesamt und HDL Cholesterin, diastolischer und systolischer Blutdruck
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisiert, prospektiv, parallel, aktiv-kontrolliert, langzeit, offen, monozentrisch
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht berichtet
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien</u> Diabetes mellitus Typ-2-Patienten nur mit Sulfonylharnstoffbehandlung und Diät Stabiler Blutzuckerwert in den 3 Monaten der Run-in Phase
4b	Studienorganisation und Ort der Studiendurchführung	Diabeteszentrum in Aker, Norwegen
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Run-in drei Monate: nur Diät oder zusätzlich Glibenclamid-Einnahme Insulin (durchschnittlich 0,51 Einheiten b.i.d., 0,06-0,91 Einheiten) um 8 und 22 Uhr vs. Glibenclamid (durchschnittlich 7,88 mg anfangs und 8,88 mg am Ende, 1.75-10.5 mg) Anpassung während der Studie war möglich Zwei Patienten erhielten multiples Dosierungsschema (normales Insulin 30 min vor dem Essen und intermediate-acting Insulin um 22 Uhr) Andere Medikation während der Studie: Antihypertensive Therapie, ACEI (angiotensin converting enzyme inhibitors), Furosemide, Kalziumantagonist,
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden	Langzeiteffekt von Insulin gegen Sulfonylharnstoff bei Diabetes mellitus Typ-2-Patienten

Item ^a	Charakteristikum	Studieninformation
	(z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Blutzuckerkontrolle, Herz-Kreislauf-Faktoren, Insulinresistenz, Mikroalbuminurie, Lp(a) Lipoprotein, Triglyceride (TG), gesamt und HDL Cholesterin, diastolische und systolischer Blutdruck Blutdruck, Gewicht und HbA1c alle 3 Monate, restliche Parameter zu Beginn, nach 6 und 12 Monaten
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht berichtet
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Nicht berichtet
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	Randomisierung
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nicht berichtet
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nicht berichtet
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) b) c) nicht verblindet
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Mittelwert±SE Veränderungen der Variablen über verbundene Student's t-Test, Wilcoxon Vorzeichen-Rang -Test Differenzen der Variablen über Student's t-Test, Mann-Whitney U-Test Pearson Korrelationskoeffizient

Item ^a	Charakteristikum	Studieninformation
		Signifikanzniveau 5%
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	a) 18 vs. 18 b) 18 vs. 18 c) 18 vs. 16
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow Chart
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	Regulär beendet
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar

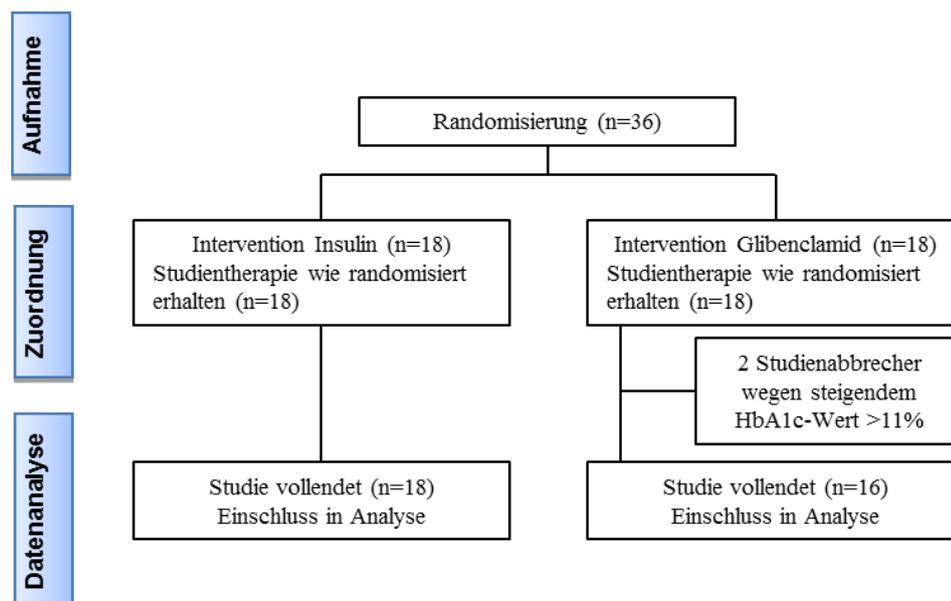


Abbildung 17: Flussdiagramm Studiendesign und-methodik für Birkenland 1994

Tabelle 4-120 (Anhang): Studiendesign und -methodik für Studie von Tovi 1998

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<u>Primäres Ziel</u> Effekt der verbesserten Stoffwechselkontrolle auf das Wohlbefinden und die Symptome von älteren Patienten mit Typ-2 Diabetes mellitus und Insulintherapie Entdecken von Symptomen/Anzeichen, die das Unwohlsein durch hohen Bluthochdruck und nicht andere altersbedingte Symptome verursachen und für den Beginn einer notwendigen Insulintherapie sprechen
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	randomisierte, prospektive, parallele, aktivkontrollierte, monozentrische zweiarmlige Studie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht zutreffend
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien:</u> Patienteneinwilligung Ältere Patienten (>70 Jahre) mit Typ-2 Diabetes mellitus zweimalige erfolglose orale Antidiabetiktherapie ohne Dekompensation schlechte Blutzuckerkontrolle trotz hoher Sulfonylharnstoffdosis
4b	Studienorganisation und Ort der Studiendurchführung	Schweden, monozentrisch
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Insulin 0,53±0,27 Einheiten/kg vs. Sulfonylharnstoff (Glibenclamid 7-10 mg oder Glipizide 10-15 mg) Anpassung bis Blutzuckerwert bei 6-12 mmol/L
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	<u>Primäres Ziel</u> Effekt der verbesserten Stoffwechselkontrolle auf das Wohlbefinden und die Symptome von älteren Patienten mit Typ-2 Diabetes mellitus und Insulintherapie Entdecken von Symptomen/Anzeichen, die

Item ^a	Charakteristikum	Studieninformation
		das Unwohlsein durch hohen Bluthochdruck und nicht andere altersbedingte Symptome verursachen und für den Beginn einer notwendigen Insulintherapie sprechen Veränderung des HbA1c-Wertes über HPLC Technik Veränderung Blutzuckerwert Gewicht und BMI bei jeder Visite Mikroalbuminurie Blutdruck, Achilles Reflex, Fußgeschwüre, Augen Messungen jeweils zu Beginn, nach 6 und 12 Monate Fragebogen zur Lebensqualität („Goteborg quality-of-life instrument“) alle 3 Monate
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 Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht berichtet
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Nicht berichtet
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Randomisierung, offen
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) b) c) nein
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht zutreffend
12	Statistische Methoden	

Item ^a	Charakteristikum	Studieninformation
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Approximativ normalverteilte Variablen – Mittelwert±SD 95%-KI mittels t-Verteilung Vergleiche über zweiseitigen Student's t-Test Nichtparametrische Variablen über χ^2 -Test, Mann-Whitney-Test und verbundenen Wilcoxon-Test p < 0,05 gilt als signifikant
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) 22 vs. 18 b) 22 vs. 18 c) 19 vs. 16
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow Chart
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	Regulär beendet
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie d

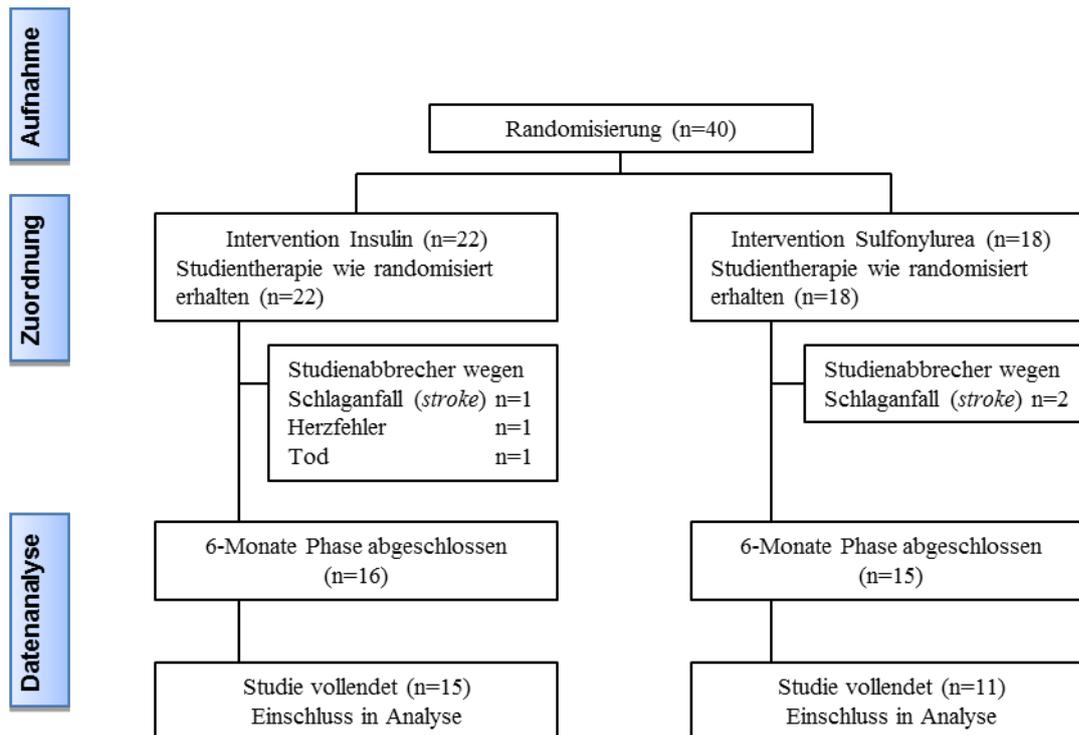


Abbildung 18: Flussdiagramm Studiendesign und-methodik für Tovi 1998

Tabelle 4-121 (Anhang): Studiendesign und -methodik für Studie von Nathan 1988

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Effekt, Risiko und Vorteile einer Insulin-gegen Glibenclamid-Therapie bei Patienten mit Diabetes mellitus Nüchternblutzuckerwert <6,4 mmol/L ohne Hypoglykämie , gut kontrolliert Unterschied erfolgloser Therapie mit Insulin oder oralen Antidiabetika
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	randomisierte, prospektive, parallele, placebo-kontrollierte, monozentrische, doppelblinde, zweiarmige Studie
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht zutreffend
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien</u> Patienten mit Diabetes mellitus 30-70 Jahre alt Nicht insulinabhängig Keine normale Blutzuckerkontrolle allein durch Diät Gewicht >90% vom Idealgewicht <u>Ausschlusskriterien</u> Vor weniger als 6 Monaten Behandlung mit Insulin oder oralen Antidiabetika Ein- oder zweimalige erfolglose orale Antidiabetikatherapie
4b	Studienorganisation und Ort der Studiendurchführung	Diabetes Klinik in Massachusetts
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Insulin (15-30 Einheiten/Tag) vs. Glibenclamid (2,5-10 Einheiten/Tag zweimal täglich) Anpassung der Medikation war möglich Doublel-dummy-Medikation
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	Effekt, Risiko und Vorteile einer Insulin-gegen Glibenclamid-Therapie bei Patienten mit Diabetes mellitus Nüchternblutzuckerwert <6,4 mmol/L ohne Hypoglykämie , gut kontrolliert Unterschied erfolgloser Therapie mit Insulin

Item ^a	Charakteristikum	Studieninformation
		oder oralen Antidiabetika Run-in 1 Monat: Diät Nur falls Nüchternblutzuckerspiegel <7,8 mmol/L darf Studie fortgeführt werden Messungen jeden Monat (bis 9 Monate)
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 Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Nicht berichtet
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Randomisierungsliste mit Computer erstellt
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Nicht berichtet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Randomisierungsliste
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) nein
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht zutreffend
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Verbundene und unverbundene Stundt's t-Tests Pearson-Korrelation Auswertung mittels SPSS Mittelwert±SD ITT-Analyse
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht berichtet
-	Resultate	

Item^a	Charakteristikum	Studieninformation
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	a) 15 vs. 16 b) 16 vs 17 c) 15 vs. 16 (ITT-Analyse)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow Chart
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	Regulär beendet
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar

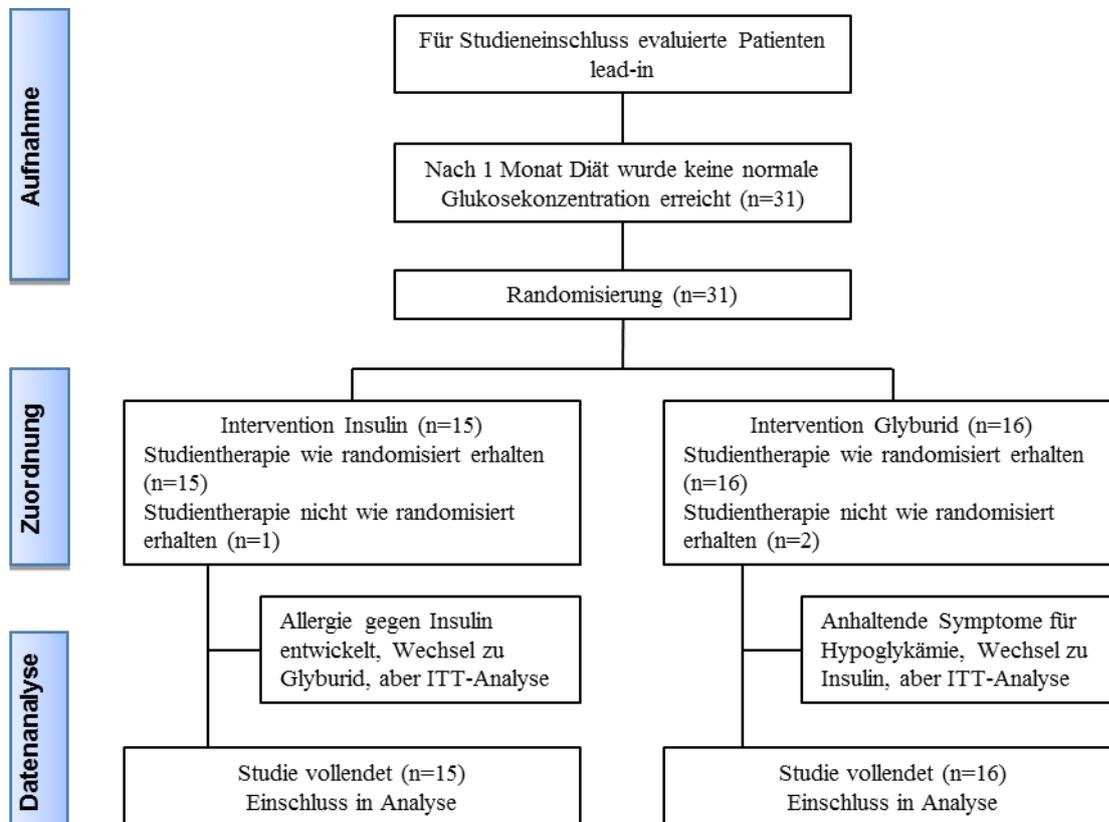


Abbildung 19: Flussdiagramm Studiendesign und -methodik für die Studie Nathan et al., 1988

Tabelle 4-122 (Anhang): Studiendesign und -methodik für die Turner 1998 (UKPDS 24)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Untersuchung und Vergleich des Ansprechens auf die Therapie mit Sulfonylharnstoffe, Insulin oder Metformin bei Typ-2-Diabetes-Patienten deren Erkrankung durch Diät alleine in den ersten 3 Studienmonaten nicht ausreichend kontrolliert werden konnte. Zudem Vergleich dieser nicht auf Diät ansprechenden Gruppe (Primäre Diätversager, Primary diet failure Group) mit den Daten der Hauptrandomisierungsgruppe (Main randomisation Group).
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multizentrische, prospektive, randomisierte, kontrollierte Interventionsstudie. Zuteilungsverhältnis (Primäre Diätversager): 2:2:4:1 für Chlorpropamid:Glibenclamid:Insulin: Metformin (adipöse Patienten) Zuteilungsverhältnis (Hauptrandomisierungsgruppe): wird in der Publikation nicht näher berichtet (werden zu denselben medikamentösen Interventionen wie die primären Diätversager randomisiert und ein Teil der Patienten zudem zu Diät allein)
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Nicht zutreffend.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: 25-65 Jahre alt neu diagnostiziertem Typ-2-Diabetes (FPG >6 mmol/L bei zwei Kontrollen) Ausschlusskriterien: Ketonurie >3 mmol/L Vorgeschichtlich Myokardinfarkt im vorangegangenen Jahr, akute Angina pectoris oder Herzversagen, mehr als eine schwere kardiale Episode Serumkreatinin >175µmol/L Maligne Hypertonie schwere Retinopathie, die mit

Item ^a	Charakteristikum	Studieninformation
		Photokoagulation behandelt werden muss eine nicht korrigierte endokrine Erkrankung eine schwere Begleiterkrankung, die lebensbedrohlich ist oder eine extensive systemische Therapie erforderlich macht
4b	Studienorganisation und Ort der Studiendurchführung	15 verschiedene Studienzentren in Großbritannien
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Run-in-Phase: Die Patienten wurden einmal monatlich untersucht und sollten eine Diät (50% Kohlenhydrate, wenig gesättigte Fettsäuren, gemäßigt hohen Faseranteil, reduzierte Gesamtkalorienzahl bei übergewichtigen Patienten) einhalten.</p> <p><u>Primäre Diätversager</u>: falls die Patienten in der initialen Beobachtungsphase unter Diät Nüchtern glukose-Werte >15 mmol/L hatten oder persistente hyperglykämische Symptome entwickelten, wurden sie randomisiert zur medikamentösen Intervention (Insulin, Sulfonylharnstoff oder Metformin)</p> <p>Stratifizierung nach Übergewicht: Nicht-adipöse Patienten: 56% erhielten Sulfonylharnstoff (Chlorpropamid oder Glibenclamid) und 44% Insulin. Adipöse Patienten: 42% erhielten Sulfonylharnstoff, 32% Insulin und 26% Metformin</p> <p><u>Hauptrandomisierungsgruppe</u>: Patienten, die innerhalb von 3 Monaten Nüchtern glukose-Werte von 6-15 mmol/L hatten, wurden zur medikamentösen Intervention (Insulin, Sulfonylharnstoff oder Metformin) oder Diät alleine randomisiert.</p> <p>746 Patienten zeigten Nüchtern glukose-Werte von <6 mmol/L nach 3 Monaten Diättherapie (diet satisfactory) und wurden weiter beobachtet.</p> <p>Die Sulfonylharnstoff-Dosis wurde insoweit angepasst, dass Nüchtern glukose-Werte von <6 mmol/L ohne hypoglykämische Episoden erreicht werden konnten. Falls unter maximaler Sulfonylharnstoff-Dosis die Nüchtern glukose-Werte bei zwei</p>

Item ^a	Charakteristikum	Studieninformation
		<p>aufeinanderfolgenden Visiten ≥ 15 mmol/L waren oder die Patienten hyperglykämische Symptome entwickelten, wurde Metformin zum Therapieregime hinzugefügt. Umgekehrt wurde bei Patienten, die diese Symptome oder Nüchtern glukose-Werte unter maximaler Metformin-Therapie erreichten, Glibenclamid dazu gegeben. Patienten, die trotz der Kombination von Sulfonylharnstoff und Metformin keine verbesserte glykämische Kontrolle hatten, wurden auf Insulin umgestellt.</p> <p>Alle Patienten mit Insulintherapie erhielten ultralentes Insulin. Diese Patienten überwachten ihre Glukosewerte zuhause falls sie mehr als 14 Einheiten ultralentes Insulin pro Tag benötigten. Falls die Blutglukosewerte vor den Mahlzeiten oder vor dem zu Bett gehen höher waren als 7 mmol/L, wurde zwei bis dreimal täglich lösliches Insulin oder zweimal täglich ein Gemisch aus mittellang wirksamen und kurzwirksamen Insulinen gegeben.</p> <p>Alle 3 Monate der Behandlung wurden die Patienten angeschrieben, und die Compliance der Patienten erfragt.</p> <p>Bei jedem Arztbesuch (alle 3 Monate) wurden die Patienten nach hypoglykämischen Symptomen gefragt. Die Einteilung in leichte bzw. schwere Hypoglykämie richtete sich danach, ob der Patient die Symptome ohne Hilfe beheben konnte oder ob eine Hilfe Dritter in Anspruch genommen werden musste bzw. eine medizinische Intervention notwendig wurde.</p> <p>Das Zentrallabor der Kliniken untersuchte Nüchtern glukose-Werte mit einem monatlichen Qualitätssicherungsschema mit dem Varianzkoeffizient von $<4\%$.</p> <p>Der HbA1c wurde mittels High-performance liquid chromatography (HPLC) (Säule mit Normbereich von 4,5-6%), Insulin mittels Radioimmunoassay mit einer 100%-igen Kreuzreaktion zu Proinsulin überprüft. HDL-Cholesterin und Gesamtcholesterin wurde über Präzipitationsmethoden gemessen.</p> <p>Homeostasis Model Assessment, ein strukturelles Modell der Glukose-Insulin-Interaktionen, wurde benutzt, um den prozentualen Anteil der β-Zell-Funktion und der Insulinsensitivität, die die gemessenen Nüchterninsulin- und Nüchtern glukose-</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Spiegel jedes Patienten induzieren, abzuschätzen. Die Prozentangaben wurden in Relation zu einer Referenzgruppe von 40 normoglykämischen Personen (Alter: 18 bis 25 Jahre, BMI: 18-25 kg/m²) ausgedrückt und wurden bei allen Patienten bei Randomisierung abgeschätzt. Sie wurden auch bei Personen bestimmt, die nach 1 Jahr weiterhin die Therapie mit dem Sulfonylharnstoff oder eine Metformintherapie erhielten.</p> <p>Eine Retinopathie wurde mit einer modifizierten Wisconsin-Einstufung der farbigen Netzhautphotographie von 4 horizontalen, 30-Grad Felder pro Auge bestimmt. Bei einem Grad von 20-20 oder schlechter wurde von einer Retinopathie gesprochen (das entspricht einem Mikroaneurysma in beiden Augen oder einer schweren Retinopathie in einem Auge).</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>Endpunkte:</p> <p>Ansprechen nach einem Jahr und nach 6 Jahren bei Patienten die mit Sulfonylharnstoff, Insulin oder Metformin behandelt wurden:</p> <ul style="list-style-type: none"> - mediane Nüchternplasmaglukose - mediane HbA1c-Konzentration - Körpergewicht (Mittelwert) <p>Zudem Vergleich dieser nicht auf Diät ansprechenden Gruppe (primäre Diätversager) mit den Daten der Hauptrandomisierungsgruppe.</p> <p>Erhalt von zusätzlicher Therapie (Sulfonylharnstoffe oder Metformin nach 6 Jahren)</p> <p>Veränderungen im Körpergewicht über 6 Jahre hinweg in der Gruppe der primären Diätversager verglichen mit der Hauptrandomisierungsgruppe (Vergleich adipöser und nicht-adipöser Patienten)</p> <p>Plasma Insulinspiegel über 6 Jahre hinweg, β-Zellfunktion und Insulinsensitivität nach einem Jahr (primäre Diätversager und Hauptrandomisierungsgruppe)</p> <p>Jährliche Rate der hypoglykämischen</p>

Item ^a	Charakteristikum	Studieninformation
		Episoden (primäre Diätversager und Hauptrandomisierungsgruppe)
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 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)	<p>Primäre Diätversager: Stratifizierung nach Übergewicht (adipös, nicht-adipös), im Anschluss Randomisierung</p> <p>Zuteilungsverhältnis (primäre Diätversager): 2:2:4:1 für Chlorpropamid:Glibenclamid:Insulin:Metformin (adipöse Patienten)</p> <p>Zuteilungsverhältnis (Hauptrandomisierungsgruppe): wird in der Publikation nicht näher berichtet (werden zu denselben medikamentösen Interventionen wie die primären Diätversager randomisiert und ein Teil der Patienten zudem zu Diät allein)</p>
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?	<p>a) nein</p> <p>b) nein</p> <p>c) nein</p>
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht berichtet.
12	Statistische Methoden	

Item ^a	Charakteristikum	Studieninformation
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Alle Analysen wurden als ITT-Analyse vorgenommen (außer wenn anders berichtet).</p> <p>Für die Beschreibung der Gruppen wurden der Mittelwert±SD, geometrischer Mittelwert (1-SD Intervall), oder der Median (Interquartilsabstand) berechnet.</p> <p>Für Unterschiede über die Zeit wurden die Daten als Mittelwert mit einem 95% Konfidenzintervall ausgedrückt.</p> <p>Prozentuale Angaben wurden mit einem 95% Konfidenzintervall notiert, außer die tatsächlichen Zahlen der Patienten waren gegeben.</p> <p>Der Mann-Whitney U-Test, der T-Test oder die Varianzanalyse mit gepaarten Vergleichen und der Anwendung der Bonferroni-Korrektur wurde für Gruppenvergleiche verwendet. Anteile wurden mit dem χ^2-Test verglichen.</p> <p>Varianzanalysen von wiederholten Messungen wurden benutzt, um Daten zwischen den Gruppen über die Zeit zu vergleichen. Zur Adjustierung von Ausgangswerten wurden Kovarianzanalysen benutzt. Die Product-Limit (Kaplan-Meier) Methode wurde benutzt, um Schätzer der Überlebensfunktionen für die Zeit bis zum Bedarf von zusätzlicher Therapie zu berechnen. Für diese Analyse wurden die Patienten in 3 Gruppen aufgeteilt, gemäß den Tertilen der Verteilungen von Alter oder BMI bei Randomisierung. Für alle Analysen wurde ein p-Wert <0,05 als signifikant angesehen.</p> <p>Die jährlichen Daten zum Körpergewicht und der Nüchternplasmaglukose wurden als Median von 3 aufeinanderfolgenden Untersuchungen für jeden Patienten berechnet (die jährliche Untersuchung und die vorausgehende und anschließende 3-Monats-Untersuchung). HbA1c-Konzentrationen und Plasmaspiegel von Insulin und Lipiden wurden jährlich bestimmt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	-
-	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) 560 (Primäre Diätversager), 2769 (Hauptrandomisierungsgruppe)

Item ^a	Charakteristikum	Studieninformation
	<p>a) randomisiert wurden,</p> <p>b) tatsächlich die geplante Behandlung/Intervention erhalten haben,</p> <p>c) in der Analyse des primären Zielkriteriums berücksichtigt wurden</p>	<p>b) 458 (Follow up für 6 Jahre, primäre Diätversager), 1620 (Follow up für 6 Jahre, Hauptrandomisierungsgruppe)</p> <p>Von den 458 Patienten in der Gruppe der primären Diätversager erhielten bei den adipösen:</p> <p>71 Sulfonylharnstoff, 51 Insulin und 49 Metformin</p> <p>und bei den nicht-adipösen:</p> <p>160 Sulfonylharnstoff, 127 Insulin und keiner Metformin</p> <p>c) 458 (Follow up für 6 Jahre, primäre Diätversager)</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow chart
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	1977-1991
14b	Informationen, warum die Studie endete oder beendet wurde	Studie endet regulär.
a: nach CONSORT 2010.		

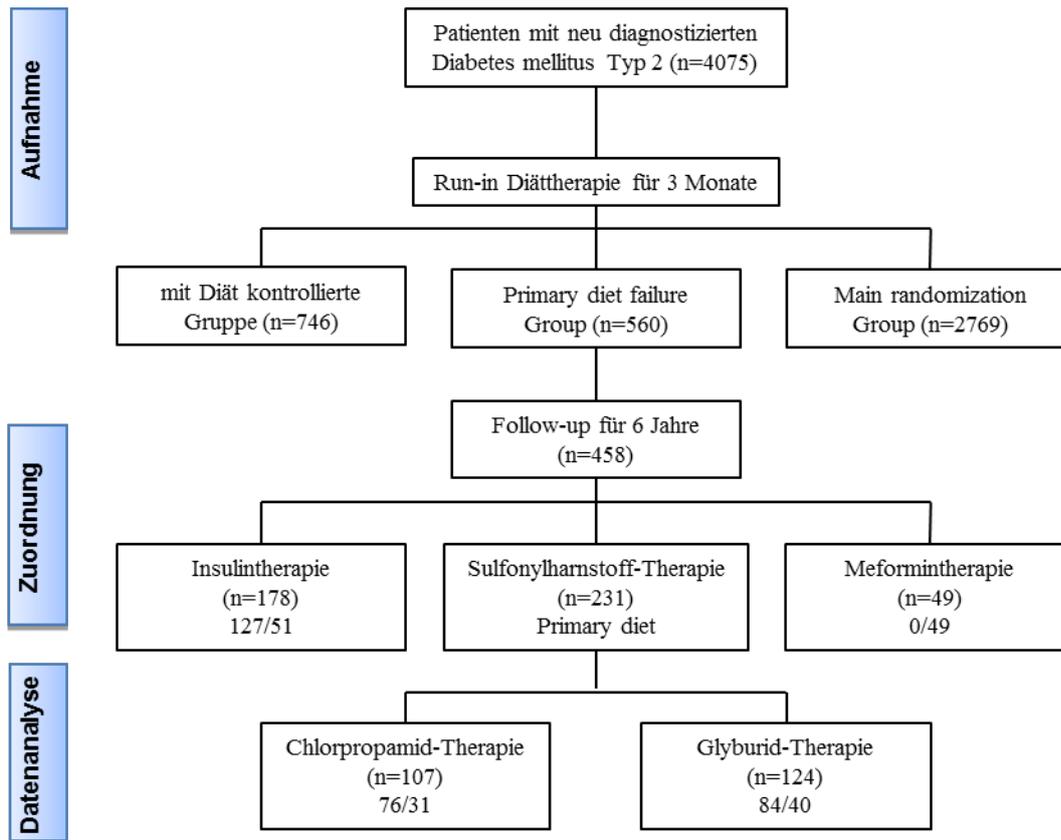


Abbildung 20: Flussdiagramm Studiendesign und-methodik für Turner 1998 (UKPDS 24)

Tabelle 4-123 (Anhang): Studiendesign und -methodik für Studie Hermansen et al. 2007

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Bestimmung von Wirksamkeit und Sicherheit des selektiven oralen DDP-4-Inhibitors 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 einfachblinder 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</p> <p>Typ-2-Diabetes</p> <p>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 einfachblinden Placebo Run-in Phase</p> <p>Compliance ≥75 % in der Placebo Run-in Phase</p> <p>Ausschlusskriterien</p> <p>Typ-1-Diabetes</p> <p>Insulin innerhalb 8 Wochen vor der Screening-Untersuchung</p> <p>Kreatininclearance <45ml/min (<60 ml/min bei Metforminmedikation)</p> <p>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+100 mg Sitagliptin 1 x tgl. - Glimepirid+Metformin+100 mg Sitagliptin

Item ^a	Charakteristikum	Studieninformation
		<p>1 x tgl.</p> <p>Die Dosierung von Glimperid 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 Glimperid±Metformin und anschließend eine bis zu 10-wöchige Dosisstabilisierungsphase 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</p> <p>Prozentuale Veränderung von</p> <ul style="list-style-type: none"> - Gesamtcholesterin - LDL-Cholesterin - Triglyzeriden - HDL-Cholesterin - Non HDL-Cholesterin <p>zu Woche 24</p> <p>Weitere Zielvariablen 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</p> <p>441 Patienten randomisiert:</p> <ul style="list-style-type: none"> - 106 Placebo +Glimperid, 113 Placebo+Glimperid+Metformin (n=219) - 106 Sitagliptin+Glimperid, 116 Sitagliptin+Glimperid+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

Item ^a	Charakteristikum	Studieninformation
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung und die Stratifizierung wurden mittels eines interaktiven Voice Response Systems (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 100 mg 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

Item ^a	Charakteristikum	Studieninformation
		Geschlecht, Ethnizität, Alter (\geq vs. $>$ Median) 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, Lost to Follow-up 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, Lost to Follow-up 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	keine Angaben

Item ^a	Charakteristikum	Studieninformation
	beendet wurde	
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

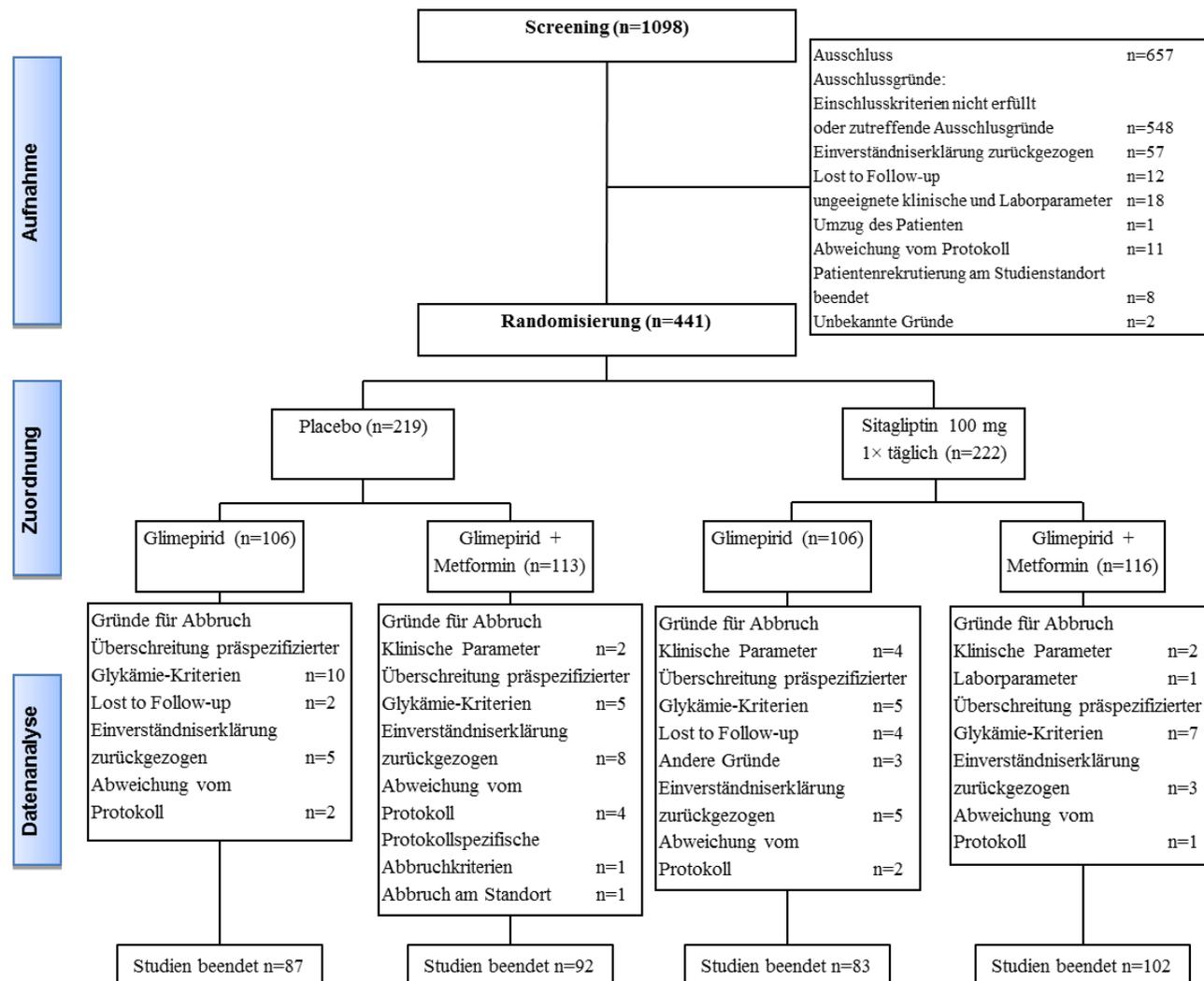


Abbildung 21: Flussdiagramm Studiendesign und-methodik für Hermansen 2007

Tabelle 4-124 (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 4 mg Glimepirid 1 x täglich
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multinationale, randomisierte, doppelblinde, placebokontrollierte Studie mit 4-wöchiger 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 \geq7,5 mg, \geq2 mg Glimepirid oder einer äquipotenten Dosis Glipizid seit mindestens 3 Monaten stabile Dosis OAD seit mind. 4 Wochen vor Screening-Visite HbA1c 7,5 – 11% Nüchternblutzucker <15 mmol/L</p> <p>Ausschlusskriterien</p> <p>Typ-1 Diabetes oder sekundäre Diabetesform Myokardinfarkt, Instabile Angina oder CABG in den letzten 6 Monaten Herzinsuffizienz NYHA III/IV Schwere Lebererkrankung (Zirrhose, chronisch aktive Hepatitis) Andere OAD als Sulfonylharnstoff in den letzten 2 Monaten erhöhte Werte für ALT oder AST (>3 x ULN) direktes Bilirubin (>1,3 x ULN) Serumkreatinin (>220 μmol/L) TSH Nüchtertriglyzeride (>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

Item ^a	Charakteristikum	Studieninformation
		0,5 Prozentpunkten mit einer Power von mindestens 90% bei einer Drop-out-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 und 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 Kovariante: Ausgangswert Zweiseitige Tests, Signifikanzniveau 0,05 Für HbA1c und Nüchternblutzucker Mehrfachtestung: Step-up-Prozedur nach Hochberg Vergleich der Anteile Patienten mit

Item ^a	Charakteristikum	Studieninformation
		HbA1c < 7% mit χ^2 -Test
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Präspezifizierte Subgruppen: - Patienten mit Ausgangs-HbA1c ≤ 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: Unerwünschte Ereignisse: 4 Abweichende Laborwerte: 0 Mangelnder Therapieeffekt: 6 Protokollverletzung: 2 Einwilligung widerrufen: 4 Lost to Follow up: 2 Administrativ: 1 Vildagliptin 100 mg: Unerwünschte Ereignisse: 4 Abweichende Laborwerte: 1 Mangelnder Therapieeffekt: 5 Protokollverletzung: 0 Einwilligung widerrufen: 4 Lost to Follow up: 26 Administrativ: 1 Placebo: Unerwünschte Ereignisse: 2 Abweichende Laborwerte: 0 Mangelnder Therapieeffekt: 16 Protokollverletzung: 2 Einwilligung widerrufen: 12 Lost to Follow up: 4 Administrativ: 0
14	Aufnahme / Rekrutierung	

Item ^a	Charakteristikum	Studieninformation
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Mai 2004 bis Oktober 2005
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.

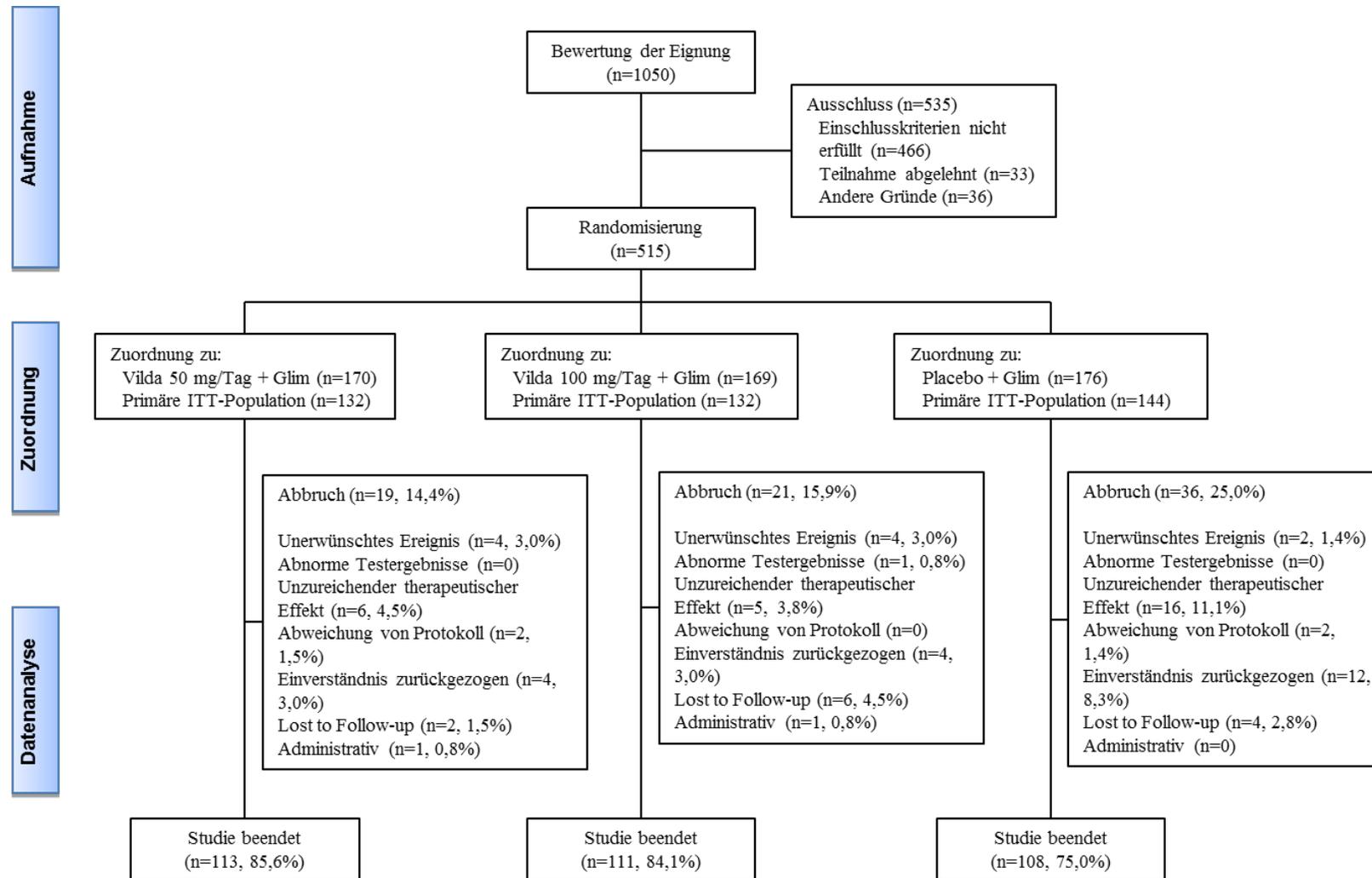


Abbildung 22: Flussdiagramm Studiendesign und-methodik für Garber et al. 2008

Tabelle 4-125 (Anhang): Studiendesign und -methodik für Studie Lewin 2012

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere 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 Run-in
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 6 Monaten Behandlung mit Rosiglitazon, Pioglitazon, GLP-1-Agonisten, 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 Sulfonylharnstoff-Medikation in konstanter

Item ^a	Charakteristikum	Studieninformation
		Dosis
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-Measures Analysis zur HbA1c-Änderung bis Woche 6 Woche 12 Woche 18 Änderung des Nüchternblutzuckers 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

Item ^a	Charakteristikum	Studieninformation
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 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 Unerwünschte Ereignisse: 5 Protokollverletzung: 1 Lost to Follow up: 0 Patiententenwunsch: 2 keine Angaben zum Abbruchgrund 2 Placebo 7 Unerwünschte Ereignisse: 3 Protokollverletzung: 0 Lost to Follow-up: 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

Item ^a	Charakteristikum	Studieninformation
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

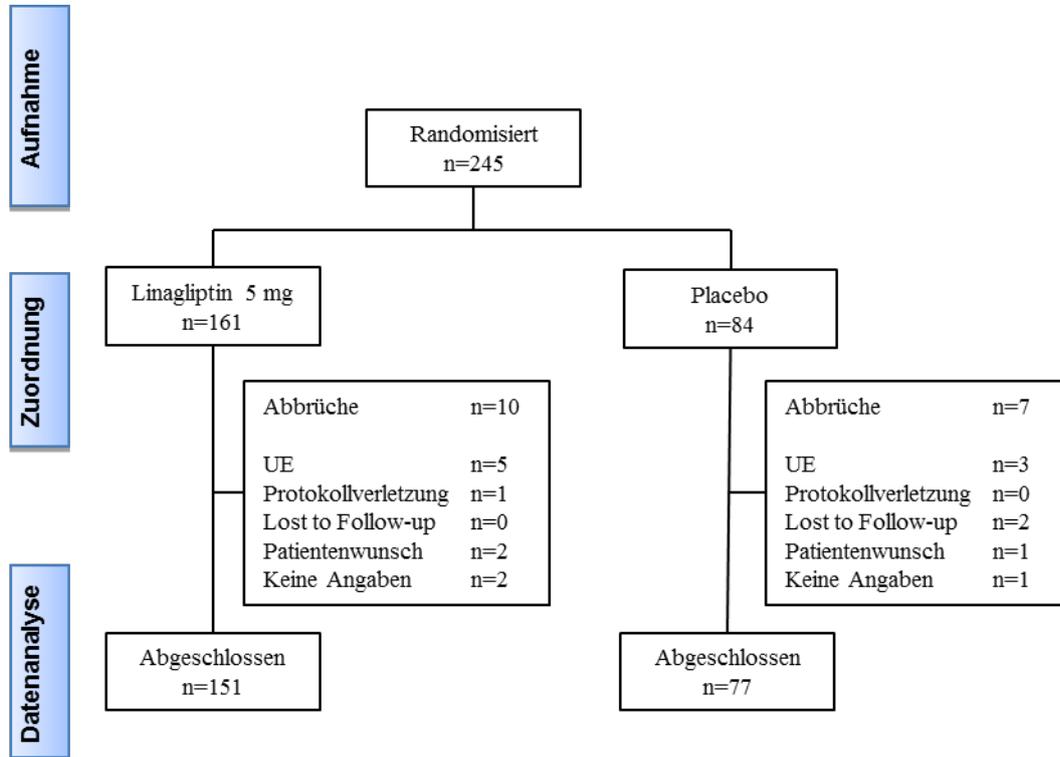


Abbildung 23: Flussdiagramm Studiendesign und-methodik für die Studie Lewin 2012

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 4-126 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie CV181040

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: CV181040

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Studienbericht CV181040 (Bristol-Myers Squibb, 2008)	CSR40

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 (CSR40)

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 (CSR40)

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 (CSR40)

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 (CSR40)

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: 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 (CSR40)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine*Beachte die Hinweise zu Punkt A.4!*

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Gewichtsveränderung

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Erhebung und Beurteilung des Endpunktes erfolgte verblindet (CSR40)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung
Erhebung und Beurteilung des Endpunktes erfolgte verblindet (CSR40)

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-127 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Shank 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: Shank 1995

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Publikation (Shank et al., 1995)	Pub_Shank_1995

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 (PubShank1995)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es gab kein Placebo-Insulin (PubShank1995)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere

Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wird als doppelblind beschrieben, aber es gibt kein Insulin-Placebo (PubShank1995)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wird als doppelblind beschrieben, aber es gibt kein Insulin-Placebo (PubShank1995)

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. (PubShank1995)

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:

Aufgrund des Studiendesigns (PubShank1995)

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
Es gab kein Placebo-Insulin (PubShank1995)

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:

Aufgrund der geringen Fallzahl in der Studie (PubShank1995)

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. (PubShank1995)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund des Studiendesigns und der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden (PubShank1995)

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
Es gab kein Placebo-Insulin (PubShank1995)

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:

Aufgrund der geringen Fallzahl in der Studie (PubShank1995)

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. (PubShank1995)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund des Studiendesigns und der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden (PubShank1995)

Tabelle 4-128 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Birkeland 1994

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: Birkeland 1994

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Publikation (Birkeland et al., 1994)	Pub_Birkeland_1994

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 (PubBirkeland1994)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genauen Angabe zur Erzeugung der Randomisierungssequenz
(PubBirkeland1994)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es fand keine Verblindung statt (PubBirkeland1994)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es fand keine Verblindung statt (PubBirkeland1994)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es fand keine Verblindung statt (PubBirkeland1994)

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. (PubBirkeland1994)

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:

Aufgrund des Studiendesigns (PubBirkeland1994)

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
Es fand keine Verblindung statt (PubBirkeland1994)

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:

Aufgrund der geringen Fallzahl in der Studie, kann eine Verzerrung nicht ausgeschlossen werden (PubBirkeland1994)

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. (PubBirkeland1994)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund des Studiendesigns und der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden (PubBirkeland1994)

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
Es fand keine Verblindung statt (PubBirkeland1994)

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:

Aufgrund der geringen Fallzahl in der Studie, kann eine Verzerrung nicht ausgeschlossen werden (PubBirkeland1994)

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. (PubBirkeland1994)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund des Studiendesigns und der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden (PubBirkeland1994)

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
Es fand keine Verblindung statt (PubBirkeland1994)

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:

Aufgrund der geringen Fallzahl in der Studie, kann eine Verzerrung nicht

ausgeschlossen werden (PubBirkeland1994)

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. (PubBirkeland1994)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund des Studiendesigns und der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden (PubBirkeland1994)

Tabelle 4-129 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Tovi 1998

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: Tovi 1998

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
-----------------------------	--------

Publikation (Tovi et al., 1998)	Pub_Tovi_1998

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 (PubTovi1998)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genauen Angabe zur Erzeugung der Randomisierungssequenz (PubTovi1998)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es fand keine Verblindung statt (PubTovi1998)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es fand keine Verblindung statt (PubTovi1998)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es fand keine Verblindung statt (PubTovi1998)

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. (PubTovi1998)

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:

Aufgrund des Studiendesigns (PubTovi1998)

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
Es fand keine Verblindung statt (PubTovi1998)

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:

Aufgrund der geringen Fallzahl in der Studie, kann eine Verzerrung nicht ausgeschlossen werden (PubTovi1998)

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. (PubTovi1998)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund des Studiendesigns und der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden

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
Es fand keine Verblindung statt (PubTovi1998)

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 Gruppeneuteilung 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:

Aufgrund der geringen Fallzahl in der Studie, kann eine Verzerrung nicht ausgeschlossen werden (PubTovi1998)

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. (PubTovi1998)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund des Studiendesigns und der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden (PubTovi1998)

Tabelle 4-130 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Nathan 1988

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: Nathan 1988

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Publikation (Nathan et al., 1988)	Pub_Nathan_1988

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 (PubNathan1988)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genauen Angabe zur Erzeugung der Randomisierungssequenz (PubNathan1988)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es fand eine Verblindung statt (PubNathan1988)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es fand eine Verblindung statt (PubNathan1988)

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. (PubNathan1988)

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
Es fand eine Verblindung statt (PubNathan1988)

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:

Aufgrund der geringen Fallzahl in der Studie, kann eine Verzerrung nicht ausgeschlossen werden (PubNathan1988)

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. (PubNathan1988)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden
(PubNathan1988)

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
Es fand eine Verblindung statt (PubNathan1988)

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:

Aufgrund der geringen Fallzahl in der Studie, kann eine Verzerrung nicht ausgeschlossen werden (PubNathan1988)

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. (PubNathan1988)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden

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

Es fand eine Verblindung statt (PubNathan1988)

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:

Aufgrund der geringen Fallzahl in der Studie, kann eine Verzerrung nicht

ausgeschlossen werden (PubNathan1988)

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. (PubNathan1988)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

aufgrund der geringen Fallzahl kann eine Verzerrung nicht ausgeschlossen werden
(PubNathan1988)

Tabelle 4-131 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Turner 1998 (UKPDS 24)

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: Turner 1998 (UKPDS 24)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Publikation (Turner et al., 1998)	Pub_Turner_1998

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 (PubTurner1998)

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genauen Angabe zur Erzeugung der Randomisierungssequenz (PubTurner1998)

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

- ja:** Die Gruppen wurden zeitlich parallel verfolgt.
- unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.
- nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

- ja:** Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
 - Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
 - Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet
- unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.
- nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es fand keine Verblindung statt (PubTurner1998)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

- ja:** Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.

- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es fand keine Verblindung statt (PubTurner1998)

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Es fand keine Verblindung statt (PubTurner1998)

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. (PubTurner1998)

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:

Aufgrund des Studiendesigns (PubTurner1998)

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
Es fand keine Verblindung statt (PubTurner1998)

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. (PubTurner1998)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund des hohen Verzerrungspotential auf Studienebene (PubTurner1998)

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
Es fand keine Verblindung statt (PubTurner1998)

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. (PubTurner1998)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund des hohen Verzerrungspotential auf Studienebene (PubTurner1998)

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
Es fand keine Verblindung statt (PubTurner1998)

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. (PubTurner1998)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:
aufgrund des hohen Verzerrungspotential auf Studienebene (PubTurner1998)

Tabelle 4-132 (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)	Pub_Hermansen_2007

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 (Pub_Hermansen_2007)

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. (Pub_Hermansen_2007)

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
(Pub_Hermansen_2007)

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. (Pub_Hermansen_2007)

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
(Pub_Hermansen_2007)

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. (Pub_Hermansen_2007)

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
(Pub_Hermansen_2007)

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. (Pub_Hermansen_2007)

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
(Pub_Hermansen_2007)

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. (Pub_Hermansen_2007)

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 4-133 (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)	Pub_Garber_2008

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunkübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie (PubGarber2008)

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

**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

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 Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

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 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: 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 4-134 (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
Publikation (Lewin et al., 2012)	Pub_Lewin_2012

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 (PubLewin2012)

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:

**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:

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 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. (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 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. (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: