

Dossier zur Nutzenbewertung gemäß § 35a SGB V

*Elvitegravir/Cobicistat/Emtricitabin/Tenofoviridisoproxil
(Stribild[®])*

Gilead Sciences GmbH

Modul 4 B

*Zur Behandlung der Infektion mit dem Humanen
Immundefizienzvirus 1 (HIV-1) bei Erwachsenen im
Alter von 18 Jahren und darüber, bei denen HIV-1 keine
Mutationen aufweist, die bekanntermaßen mit
Resistenzen gegen einen der drei antiretroviralen
Wirkstoffe von Stribild[®] assoziiert sind.*

**Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen**

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	15
Abkürzungsverzeichnis	17
4.1 Zusammenfassung der Inhalte von Modul 4.....	21
4.2 Methodik	41
4.2.1 Fragestellung	41
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	46
4.2.3 Informationsbeschaffung	49
4.2.3.1 Studien des pharmazeutischen Unternehmers	50
4.2.3.2 Bibliografische Literaturrecherche	50
4.2.3.3 Suche in Studienregistern	51
4.2.3.4 Selektion relevanter Studien	52
4.2.4 Bewertung der Aussagekraft der Nachweise.....	53
4.2.5 Informationssynthese und -analyse	55
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	55
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	56
4.2.5.3 Meta-Analysen.....	65
4.2.5.4 Sensitivitätsanalysen	67
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	68
4.2.5.6 Indirekte Vergleiche	71
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	73
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	74
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	74
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	74
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	79
4.3.1.1.3 Studien aus der Suche in Studienregistern	85
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	85
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	98
4.3.1.2.1 Studiendesign und Studienpopulationen	98
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	106
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	108
4.3.1.3.1 Endpunkte– RCT.....	109
4.3.1.3.1.1 Virologisches Ansprechen– RCT	110
4.3.1.3.1.2 Zeit bis Verlust des virologischen Ansprechens – RCT.....	117
4.3.1.3.1.3 Veränderung der CD4-Zellzahl– RCT	120
4.3.1.3.1.4 Unerwünschte Ereignisse– RCT.....	124
4.3.1.3.2 Subgruppenanalysen – RCT.....	136

4.3.1.3.2.1	Virologisches Ansprechen (Snapshot-Algorithmus; präspezifizierte Subgruppenanalysen).....	156
4.3.1.3.2.2	Virologisches Ansprechen (TLOVR-Algorithmus)	162
4.3.1.3.2.3	Zeit bis zum Verlust des virologischen Ansprechens.....	163
4.3.1.3.2.4	Zeit bis zum Verlust des virologischen Ansprechens zu Woche 96	163
4.3.1.3.2.5	Veränderung der CD4-Zellzahl	164
4.3.1.3.2.6	Gesamtrate unerwünschter Ereignisse.....	166
4.3.1.3.2.7	Gesamtrate schwerwiegender unerwünschter Ereignisse.....	167
4.3.1.3.2.8	Therapieabbrüche aufgrund unerwünschter Ereignisse.....	168
4.3.1.3.2.9	Unerwünschte Ereignisse Grad 3-4	169
4.3.1.3.2.10	Neurologische & psychiatrische Ereignisse	170
4.3.1.3.2.11	Hautausschläge	172
4.3.1.3.2.12	Knochenfrakturen	174
4.3.1.3.2.13	Erkrankungen des Gastrointestinaltrakts (SOC).....	176
4.3.1.3.2.14	Diarrhoe	178
4.3.1.3.2.15	Übelkeit	180
4.3.1.3.2.16	Fatigue	182
4.3.1.3.2.17	Erkrankungen des Nervensystems (SOC)	183
4.3.1.3.2.18	Kopfschmerzen	185
4.3.1.3.2.19	Psychiatrische Erkrankungen (SOC)	186
4.3.1.3.2.20	Abnorme Träume.....	188
4.3.1.3.2.21	Schlaflosigkeit	188
4.3.1.3.2.22	Depression	189
4.3.1.3.2.23	Erkrankungen der Haut und des Unterhautzellgewebes (SOC)	190
4.3.1.3.2.24	Hautausschlag.....	190
4.3.1.3.2.25	Augenerkrankungen (SOC)	191
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	194
4.3.2	Weitere Unterlagen.....	206
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	206
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	206
4.3.2.1.1.1	Studien des pharmazeutischen Unternehmers	207
4.3.2.1.1.2	Studien aus der bibliografischen Literaturrecherche	210
4.3.2.1.1.3	Studien aus der Suche in Studienregistern	212
4.3.2.1.1.4	Resultierender Studienpool aus den einzelnen Suchschritten	215
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	217
4.3.2.1.2.1	Studiendesign und Studienpopulationen	217
4.3.2.1.2.2	Verzerrungspotenzial auf Studienebene	232
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	234
4.3.2.1.3.1	Endpunkte – indirekte Vergleiche aus RCT	236
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	274
4.3.2.2	Nicht randomisierte vergleichende Studien.....	275
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	275
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	275
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	276

4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	276
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	277
4.3.2.3	Weitere Untersuchungen.....	277
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	277
4.3.2.3.1.1	Studien des pharmazeutischen Unternehmers	278
4.3.2.3.1.2	Studien aus der bibliografischen Literaturrecherche	280
4.3.2.3.1.3	Studien aus der Suche in Studienregistern	282
4.3.2.3.1.4	Resultierender Studienpool: weitere Untersuchungen	282
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	283
4.3.2.3.2.1	Studiendesign und Studienpopulationen	284
4.3.2.3.2.2	Verzerrungspotenzial auf Studienebene	290
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	291
4.3.2.3.3.1	Virologisches Ansprechen – weitere Untersuchungen	291
4.3.2.3.3.2	Veränderung der CD4-Zellzahl– weitere Untersuchungen	292
4.3.2.3.3.3	Unerwünschte Ereignisse– weitere Untersuchungen	293
4.3.2.3.3.4	Subgruppenanalysen – weitere Untersuchungen	297
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	297
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	303
4.4.1	Beurteilung der Aussagekraft der Nachweise	303
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	305
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	341
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	342
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	343
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	343
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	343
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	344
4.5.4	Verwendung von Surrogatendpunkten	347
4.6	Liste der eingeschlossenen Studien.....	352
4.7	Referenzliste.....	355
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		366
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		382
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund		387
Anhang 4-D : Liste der abgebrochenen Studien		389
Anhang 4-E : Liste der laufenden Studien.....		390
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....		391
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		443

Tabellenverzeichnis

	Seite
Tabelle 4-1: Gegenüberstellung von Zusatznutzen und Schaden von EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF	28
Tabelle 4-2: Übersicht der Kriterien für den Studieneinschluss für den Vergleich von EVG/COBI/FTC/TDF gegen EFV/FTC/TDF, ATV/r+FTC/TDF und RAL+FTC/TDF	47
Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	74
Tabelle 4-4: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	76
Tabelle 4-5: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	85
Tabelle 4-6: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	86
Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel (nicht-vorbehandelte Patienten).....	92
Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	94
Tabelle 4-9: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	96
Tabelle 4-10: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	97
Tabelle 4-11: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	99
Tabelle 4-12: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	100
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Alter, Geschlecht, BMI; Safety Analyse-Set).....	100
Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Rasse, Safety Analyse-Set)	100
Tabelle 4-15: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Ethnische Abstammung, Safety Analyse-Set).....	101
Tabelle 4-16: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (HIV-1-RNA, Safety Analyse-Set)	101
Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (CD4-Zellzahl, Safety Analyse-Set)	101
Tabelle 4-18: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (HIV-Status, Safety Analyse-Set)	101
Tabelle 4-19: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Adhärenz, Safety Analyse-Set).....	102
Tabelle 4-20: Charakterisierung der Studienpopulationen (Patientendisposition) – RCT mit dem zu bewertenden Arzneimittel	102

Tabelle 4-21: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	107
Tabelle 4-22: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	108
Tabelle 4-23: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	108
Tabelle 4-24: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	109
Tabelle 4-25: Operationalisierung von virologisches Ansprechen	110
Tabelle 4-26: Bewertung des Verzerrungspotenzials für virologisches Ansprechen in RCT mit dem zu bewertenden Arzneimittel	110
Tabelle 4-27: Ergebnisse für virologisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set)	114
Tabelle 4-28: Ergebnisse für virologisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set)	115
Tabelle 4-29: Ergebnisse für virologisches Versagen (Wirksamkeit) aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 48)	115
Tabelle 4-30: Operationalisierung von Zeit bis Verlust des virologischen Ansprechens	117
Tabelle 4-31: Bewertung des Verzerrungspotenzials für Zeit bis Verlust des virologischen Ansprechens in RCT mit dem zu bewertenden Arzneimittel.....	118
Tabelle 4-32: Ergebnisse für Zeit bis Verlust des virologischen Ansprechens aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set)	118
Tabelle 4-33: Operationalisierung von Veränderung der CD4-Zellzahl.....	120
Tabelle 4-34: Bewertung des Verzerrungspotenzials für Veränderung der CD4-Zellzahl in RCT mit dem zu bewertenden Arzneimittel	121
Tabelle 4-35: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set).....	122
Tabelle 4-36: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set).....	123
Tabelle 4-37: Operationalisierung von Unerwünschte Ereignisse	124
Tabelle 4-38: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel	125
Tabelle 4-39: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48).....	128
Tabelle 4-40: Ergebnisse für wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)	128
Tabelle 4-41: Ergebnisse für unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)	129
Tabelle 4-42: Ergebnisse für die häufigsten unerwünschten Ereignisse (SOC, Grenze: $\geq 10\%$) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48) ...	129

Tabelle 4-43: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)	132
Tabelle 4-44: Ergebnisse für Wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)	132
Tabelle 4-45: Ergebnisse für Unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)	132
Tabelle 4-46: Ergebnisse für die häufigsten unerwünschten Ereignisse (SOC, Grenze: $\geq 10\%$) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96) ...	133
Tabelle 4-47: Übersicht über alle präspezifizierten (fett gedruckt) und post-hoc durchgeführten Interaktionstests	138
Tabelle 4-48: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Alter (Woche 48)	156
Tabelle 4-49: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Geschlecht (Woche 48).....	157
Tabelle 4-50: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Rasse (Woche 48)	157
Tabelle 4-51: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Ausgangsviruslast (Woche 48)	158
Tabelle 4-52: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)	158
Tabelle 4-53: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Therapieadhärenz (Woche 48).....	159
Tabelle 4-54: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Alter (Woche 96)	159
Tabelle 4-55: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Geschlecht (Woche 96).....	160
Tabelle 4-56: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Rasse (Woche 96)	160
Tabelle 4-57: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Ausgangsviruslast (Woche 96)	161
Tabelle 4-58: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)	161
Tabelle 4-59: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Therapieadhärenz (Woche 96).....	162
Tabelle 4-60: Subgruppenanalyse für virologisches Ansprechen (TLOVR), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)	162
Tabelle 4-61: Subgruppenanalyse für virologisches Ansprechen (TLOVR), stratifiziert nach Therapieadhärenz (Woche 96).....	163
Tabelle 4-62: Subgruppenanalyse für Zeit bis zum Verlust des virologischen Ansprechens, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96).....	163

Tabelle 4-63: Subgruppenanalyse für Zeit bis zum Verlust des virologischen Ansprechens, stratifiziert nach Therapieadhärenz (Woche 96).....	164
Tabelle 4-64: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)	164
Tabelle 4-65: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach Therapieadhärenz (Woche 48)	165
Tabelle 4-66: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach Geschlecht (Woche 96)	165
Tabelle 4-67: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)	166
Tabelle 4-68: Subgruppenanalyse für Gesamtrate unerwünschter Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 48)	166
Tabelle 4-69: Subgruppenanalyse für Gesamtrate schwerwiegender unerwünschter Ereignisse, stratifiziert nach Alter (Woche 48)	167
Tabelle 4-70: Subgruppenanalyse für Gesamtrate schwerwiegender unerwünschter Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 48)	167
Tabelle 4-71: Subgruppenanalyse für Gesamtrate schwerwiegender unerwünschter Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 96)	168
Tabelle 4-72: Subgruppenanalyse für Therapieabbrüche aufgrund unerwünschter Ereignisse, stratifiziert nach Rasse (Woche 48)	168
Tabelle 4-73: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Ausgangsviruslast (Woche 48)	169
Tabelle 4-74: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Therapieadhärenz (Woche 48)	169
Tabelle 4-75: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Ausgangsviruslast (Woche 96)	170
Tabelle 4-76: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Geschlecht (Woche 48)	170
Tabelle 4-77: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Therapieadhärenz (Woche 48)	171
Tabelle 4-78: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Geschlecht (Woche 96)	171
Tabelle 4-79: Subgruppenanalyse für Hautausschläge, stratifiziert nach Alter (Woche 48) .	172
Tabelle 4-80: Subgruppenanalyse für Hautausschläge, stratifiziert nach Geschlecht (Woche 48)	172
Tabelle 4-81: Subgruppenanalyse für Hautausschläge, stratifiziert nach Rasse (Woche 48)	173
Tabelle 4-82: Subgruppenanalyse für Hautausschläge, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)	173
Tabelle 4-83: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach Alter (Woche 48)	174

Tabelle 4-84: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach Geschlecht (Woche 48).....	174
Tabelle 4-85: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)	175
Tabelle 4-86: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach Geschlecht (Woche 96).....	175
Tabelle 4-87: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Rasse (Woche 48)	176
Tabelle 4-88: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)	176
Tabelle 4-89: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Rasse (Woche 96)	177
Tabelle 4-90: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Ausgangsviruslast (Woche 96)	177
Tabelle 4-91: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)	178
Tabelle 4-92: Subgruppenanalyse für Diarrhoe, stratifiziert nach Ausgangsviruslast (Woche 48).....	178
Tabelle 4-93: Subgruppenanalyse für Diarrhoe, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)	179
Tabelle 4-94: Subgruppenanalyse für Diarrhoe, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)	179
Tabelle 4-95: Subgruppenanalyse für Übelkeit, stratifiziert nach Alter (Woche 48)	180
Tabelle 4-96: Subgruppenanalyse für Übelkeit, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)	180
Tabelle 4-97: Subgruppenanalyse für Übelkeit, stratifiziert nach Alter (Woche 60)	181
Tabelle 4-98: Subgruppenanalyse für Übelkeit, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)	181
Tabelle 4-99: Subgruppenanalyse für Fatigue, stratifiziert nach Alter (Woche 48)	182
Tabelle 4-100: Subgruppenanalyse für Fatigue, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)	182
Tabelle 4-101: Subgruppenanalyse für Fatigue, stratifiziert nach Alter (Woche 96)	183
Tabelle 4-102: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Alter (Woche 48).....	183
Tabelle 4-103: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Rasse (Woche 48)	184
Tabelle 4-104: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Rasse (Woche 96)	184
Tabelle 4-105: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Rasse (Woche 48).....	185

Tabelle 4-106: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)	185
Tabelle 4-107: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Geschlecht (Woche 48).....	186
Tabelle 4-108: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Ausgangsviruslast (Woche 48)	186
Tabelle 4-109: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Therapieadhärenz (Woche 48).....	187
Tabelle 4-110: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Alter (Woche 96)	187
Tabelle 4-111: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Geschlecht (Woche 96).....	188
Tabelle 4-112: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Therapieadhärenz (Woche 96).....	188
Tabelle 4-113: Subgruppenanalyse für Depression, stratifiziert nach Alter (Woche 96)	189
Tabelle 4-114: Subgruppenanalyse für Depression, stratifiziert nach Geschlecht (Woche 96).....	189
Tabelle 4-115: Subgruppenanalyse für Depression, stratifiziert nach Rasse (Woche 96)	190
Tabelle 4-116: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Alter (Woche 48).....	190
Tabelle 4-117: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Alter (Woche 48).....	191
Tabelle 4-118: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Rasse (Woche 48).....	191
Tabelle 4-119: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Ausgangsviruslast (Woche 48).....	192
Tabelle 4-120: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Therapieadhärenz (Woche 48)	192
Tabelle 4-121: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Alter (Woche 96).....	193
Tabelle 4-122: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Rasse (Woche 96).....	193
Tabelle 4-123: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Therapieadhärenz (Woche 96)	194
Tabelle 4-124: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel für indirekte Vergleiche	208
Tabelle 4-125: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – Studien für indirekte Vergleiche.....	209
Tabelle 4-126: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel für indirekte Vergleiche.....	213

Tabelle 4-127: Relevante Studien aus der Suche in Studienregistern – RCT mit der zweckmäßigen Vergleichstherapie für indirekte Vergleiche	214
Tabelle 4-128: Studienpool – Studien für indirekte Vergleiche (zu bewertendes Arzneimittel und zweckmäßige Vergleichstherapie)	216
Tabelle 4-129: Charakterisierung der eingeschlossenen Studien – Studien für den indirekten Vergleich	219
Tabelle 4-130: Charakterisierung der Interventionen – Studien für indirekte Vergleiche.....	221
Tabelle 4-131: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.).....	222
Tabelle 4-132: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.).....	222
Tabelle 4-133: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.).....	223
Tabelle 4-134: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.).....	223
Tabelle 4-135: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.).....	224
Tabelle 4-136: Charakterisierung der Studienpopulationen (Patientendisposition) – Studien für indirekte Vergleiche	224
Tabelle 4-137: Verzerrungspotenzial auf Studienebene – Studien für indirekte Vergleiche.	233
Tabelle 4-138: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	235
Tabelle 4-139: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	235
Tabelle 4-140: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	236
Tabelle 4-141: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	236
Tabelle 4-142: Operationalisierung von virologisches Ansprechen	237
Tabelle 4-143: Bewertung des Verzerrungspotenzials für virologisches Ansprechen in RCT für indirekte Vergleiche.....	238
Tabelle 4-144: Ergebnisse für virologisches Ansprechen (M=F) aus RCT für indirekte Vergleiche (ITT Analyse-Set).....	240
Tabelle 4-145: Ergebnisse für virologisches Ansprechen (M=F) aus RCT für indirekte Vergleiche	242
Tabelle 4-146: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	243
Tabelle 4-147: Operationalisierung von Veränderung der CD4-Zellzahl.....	244
Tabelle 4-148: Bewertung des Verzerrungspotenzials für Veränderung der CD4-Zellzahl in RCT für indirekte Vergleiche.....	244

Tabelle 4-149: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT für indirekte Vergleiche	246
Tabelle 4-150: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT für indirekte Vergleiche (Woche 48)	248
Tabelle 4-151: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	250
Tabelle 4-152: Operationalisierung von unerwünschten Ereignissen.....	251
Tabelle 4-153: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse in RCT für indirekte Vergleiche.....	253
Tabelle 4-154: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 48, Safety Set).....	256
Tabelle 4-155: Ergebnisse für wichtige unerwünschte Ereignisse (S102: neurologische & psychiatrische Ereignisse; STARTMRK: neuropsychiatrische Ereignisse) aus RCT für indirekte Vergleiche (Woche 48, Safety Set).....	257
Tabelle 4-156: Ergebnisse für die häufigsten schwerwiegenden unerwünschten Ereignisse (SOC) aus RCT für indirekte Vergleiche (Woche 48, Safety Set).....	258
Tabelle 4-157: Ergebnisse für die häufigsten mittel- bis schwerwiegenden Therapie-assoziierten unerwünschten Ereignisse (PTs) aus RCT für indirekte Vergleiche (Woche 48, Safety Set)	258
Tabelle 4-158: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 96, Safety Set).....	262
Tabelle 4-159: Ergebnisse für wichtige unerwünschte Ereignisse (S102: neurologische & psychiatrische Ereignisse; STARTMRK: neuropsychiatrische Ereignisse) aus RCT für indirekte Vergleiche (Woche 96, Safety Set).....	263
Tabelle 4-160: Ergebnisse für häufigste unerwünschte Ereignisse (SOC, Grenze: $\geq 10\%$) aus RCT für indirekte Vergleiche (Woche 96, Safety Set)	263
Tabelle 4-161: Definition und Überschneidungen der neurologischen & psychiatrischen Ereignisse bzw. neuropsychiatrischen Ereignisse	270
Tabelle 4-162: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 48).....	272
Tabelle 4-163: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 96).....	272
Tabelle 4-164: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 96, Fortsetzung I).....	273
Tabelle 4-165: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 96, Fortsetzung II)	274
Tabelle 4-166: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	276
Tabelle 4-167: Operationalisierung von <Endpunkt xxx>.....	276
Tabelle 4-168: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	277

Tabelle 4-169: Liste der Studien des pharmazeutischen Unternehmers – weitere Untersuchungen.....	279
Tabelle 4-170: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – weitere Untersuchungen	280
Tabelle 4-171: Relevante Studien aus der Suche in Studienregistern – weitere Untersuchungen.....	282
Tabelle 4-172: Studienpool – weitere Untersuchungen	283
Tabelle 4-173: Charakterisierung der eingeschlossenen Studien – weitere Untersuchungen	285
Tabelle 4-174: Charakterisierung der Interventionen – weitere Untersuchungen	286
Tabelle 4-175: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)	286
Tabelle 4-176: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)	286
Tabelle 4-177: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)	286
Tabelle 4-178: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)	287
Tabelle 4-179: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)	287
Tabelle 4-180: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)	287
Tabelle 4-181: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)	288
Tabelle 4-182: Charakterisierung der Studienpopulationen (Patientendisposition) – weitere Untersuchungen	288
Tabelle 4-183: Verzerrungspotenzial auf Studienebene – weitere Untersuchungen	290
Tabelle 4-184: Operationalisierung von virologisches Ansprechen	291
Tabelle 4-185: Bewertung des Verzerrungspotenzials für virologisches Ansprechen in weitere Untersuchungen.....	291
Tabelle 4-186: Operationalisierung von Veränderung der CD4-Zellzahl.....	292
Tabelle 4-187: Bewertung des Verzerrungspotenzials für Veränderung der CD4-Zellzahl in weitere Untersuchungen.....	292
Tabelle 4-188: Operationalisierung von unerwünschten Ereignissen.....	293
Tabelle 4-189: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse in weitere Untersuchungen.....	294
Tabelle 4-190: Ergebnisse für unerwünschte Ereignisse aus weiteren Untersuchungen (Safety Analyse-Set)	296
Tabelle 4-191: Ergebnisse für häufigste unerwünschte Ereignisse (SOC, Grenze: $\geq 10\%$) aus weiteren Untersuchungen (Safety Analyse-Set)	296

Tabelle 4-192: Ausmaß des Zusatznutzen von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF auf Endpunktebene	310
Tabelle 4-193: Gegenüberstellung von Zusatznutzen und Schaden von EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF	330
Tabelle 4-194: Ausmaß des Zusatznutzen von EVG/COBI/FTC/TDF gegenüber RAL+FTC/TDF auf Endpunktebene.....	331
Tabelle 4-195: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzen	342
Tabelle 4-196: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	343
Tabelle 4-197: Ovid: Suchstrategie EMBASE 1947 to 2013 Week 14	367
Tabelle 4-198: Ovid: Suchstrategie MEDLINE(R) without Revisions to March Week 4 2013.....	368
Tabelle 4-199: Ovid: Suchstrategie Cochrane 1991 to March 2013	369
Tabelle 4-200: Ovid: Suchstrategie EMBASE 1947 to 2013 Week 14	371
Tabelle 4-201: Ovid: Suchstrategie MEDLINE(R) without Revisions to March Week 4 2013	372
Tabelle 4-202: Ovid: Suchstrategie Cochrane 1991 to March 2013	373
Tabelle 4-203: Ovid: Suchstrategie EMBASE 1947 to 2013 Week 14	374
Tabelle 4-204: Ovid: Suchstrategie MEDLINE(R) without Revisions to March Week 4 2013	375
Tabelle 4-205: Ovid: Suchstrategie Cochrane 1991 to March 2013	376
Tabelle 4-206: Ovid: Suchstrategie EMBASE 1947 bis heute: am 08. April 2013.....	377
Tabelle 4-207: Ovid: Suchstrategie MEDLINE(R) without Revisions bis heute: am 08. April 2013	378
Tabelle 4-208: Ovid: Suchstrategie Cochrane 1991 bis heute: am 08. März 2013.....	380
Tabelle 4-209: (Anhang): im Volltext gesichtete und ausgeschlossene Studien/Publicationen aus der Recherche zu RAL+FTC/TDF	387
Tabelle 4-210: (Anhang): im Volltext gesichtete und ausgeschlossene Studien/Publicationen aus dem Screening der Publikationen auf clinicaltrials.gov zu STARTMRK	387
Tabelle 4-211 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel.....	389
Tabelle 4-212 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	390
Tabelle 4-213 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0102.....	392
Tabelle 4-214 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0103.....	406
Tabelle 4-215 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0104.....	419
Tabelle 4-216 (Anhang): Studiendesign und -methodik für STARTMRK-Studie	432

Tabelle 4-217 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0123 (nach TREND-Statement)	438
Tabelle 4-218 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GS-US-236-0102	444
Tabelle 4-219 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GS-US-236-0103	459
Tabelle 4-220 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GS-US-236-0104	477
Tabelle 4-221 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie STARTMRK	493

Abbildungsverzeichnis

	Seite
Abbildung 1: Dauerhaftigkeit (Persistenz) einer Initialtherapie als Funktion der Dosierungskomplexität [51].....	36
Abbildung 2: Forestplot des Effekts einer Einmal- im Vergleich zu einer Zweimal-Gabe auf die Adhärenzrate [52].....	37
Abbildung 3: Nicht-Adhärenz bei verschiedenen ART-Regimen [53].....	38
Abbildung 4: Relation zwischen Veränderung der CD4-Zellzahl und der Hazard für klinische Ereignisse (Aids/Tod) [73]	62
Abbildung 5: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	80
Abbildung 6: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie EFV/FTC/TDF im Vergleich mit dem zu bewertenden Arzneimittel.....	82
Abbildung 7: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie ATV/r+FTC/TDF im Vergleich mit dem zu bewertenden Arzneimittel	83
Abbildung 8: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie RAL+FTC/TDF im Vergleich mit dem zu bewertenden Arzneimittel	84
Abbildung 9: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	95
Abbildung 10: Kaplan-Meier Kurven für Zeit bis zum Verlust des virologischen Ansprechens, aus RCT; EVG/COBI/FTC/TDF versus ATV/r+FTC/TDF.....	120
Abbildung 11: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie	212
Abbildung 12: Netzwerkstruktur des indirekten Vergleichs zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF.....	237
Abbildung 13: Meta-Analyse für virologisches Ansprechen (M=F) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, ITT-Set).....	241
Abbildung 14: Netzwerkstruktur des indirekten Vergleichs zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF.....	243
Abbildung 15: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, MWD, ITT).....	248
Abbildung 16: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, Hedges' g, ITT).....	248
Abbildung 17: Netzwerkstruktur des indirekten Vergleichs zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF.....	250
Abbildung 18: Meta-Analyse für die Gesamtrate unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	268

Abbildung 19: Meta-Analyse für die Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	268
Abbildung 20: Meta-Analyse für Therapieabbrüche aufgrund unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	269
Abbildung 21: Meta-Analyse für neurologische & psychiatrische Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	269
Abbildung 22: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	281
Abbildung 23: Dauerhaftigkeit (Persistenz) einer Initialtherapie als Funktion der Dosierungskomplexität [51].....	337
Abbildung 24: Forestplot des Effekts einer Einmal- im Vergleich zu einer Zweimal-Gabe auf die Adhärenzrate [52].....	338
Abbildung 25: Nicht-Adhärenz bei verschiedenen ART-Regimen [53].....	339
Abbildung 26: Relation zwischen Veränderung der CD4-Zellzahl und der Hazard für klinische Ereignisse (Aids/Tod) [73]	352
Abbildung 27: Flow-Chart für Studie GS-US-236-0102	405
Abbildung 28: Flow-Chart für Studie GS-US-236-0103	418
Abbildung 29: Flow-Chart für Studie GS-US-236-0104	431
Abbildung 30: Flow Chart der Studie Lennox (2009)	437
Abbildung 31: Flow-Chart für Studie GS-US-236-0123	442

Abkürzungsverzeichnis

Abkürzung	Bedeutung
/r	Geboostert mit Ritonavir
Aids	Erworbenes Immundefektsyndrom (Acquired immunodeficiency syndrome)
ALT	Alanin-Aminotransferase
ANOVA	Varianzanalyse (Analysis of variance)
ART	Antiretrovirale Therapie
AST	Aspartat-Aminotransferase
ATV	Atazanavir
BMI	Body Mass Index
CD4	Cluster of Differentiation 4 (als Einheit der CD4-Zellzahl wird Zellen/ μ L oder Zellen/mm ³ als Äquivalent verwendet)
CDC	Center for Disease Control and Prevention
CG	Cockcroft-Gault
COBI	Cobicistat
CONSORT	Consolidated Standards of Reporting Trials
DIMDI	Deutsches Institut für Medizinische Dokumentation
DHHS	Department of Health and Human Services
EFV	Efavirenz
EG	Europäische Gemeinschaft
eGFR	geschätzte glomeruläre Filtrationsrate (estimated glomerular filtration rate)
EKG	Elektrokardiogramm
EMA	European Medicines Agency
EVG	Elvitegravir
FDA	U.S. Food and Drug Administration
FEM	Modell mit festen Effekten (Fixed effect model)
FSH	Serum-Follikel-stimulierendes Hormonlevel
FTC	Emtricitabin
G-BA	Gemeinsamer Bundesausschuss
HAART	Highly Active Antiretroviral Therapy
HCV	Hepatitis C-Virus

HIV-1	Humanes Immundefizienz-Virus (Typ 1) (human immunodeficiency virus (type 1))
HIVAN	HIV-assoziierte Nephropathie
HIVTSQs	HIV Treatment Satisfaction Questionnaire Status
HLGT	High level group term
HR	Hazard Ratio
ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
ICTRP	International Clinical Trials Registry Platform Search Portal
IDMC	Independent Data-Monitoring Committee
INI	Integrase-Inhibitor (alternative Abkürzungen, z. B. in Leitlinien: InSTI, INSTI, ITI)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
MedDRA	Medical Dictionary for Regulatory Activities
MTC	Mixed Treatment Comparison
MW	Mittelwert
MWD	Mittelwertdifferenz
n.b.	Nicht berechenbar
n.d.	Nicht durchführbar
n.z.	Nicht zutreffend
NDA	New drug application
NNRTI	Nicht-Nukleosidaler Reverse-Transkriptase-Inhibitoren
NRTI	Nukleosidaler/nukleotidaler Reverse-Transkriptase-Inhibitor
OR	Odds Ratio
PI	Protease-Inhibitor
PP	Per Protokoll
PT	Preferred term
RAL	Raltegravir
RAS	Randomisiertes Analyse-Set
RCT	Randomized Controlled Trial
RD	Risikodifferenz

REM	Modell mit zufälligen Effekten (Random effects model)
RNA	Ribonukleinsäure (ribonucleic acid)
SAS	Safety Analyse-Set
SD	Standardabweichung
SF-36	Short form 36
SGB	Sozialgesetzbuch
SOC	System Organ Class
STE	Surrogate Threshold Effects
STR	Single-Tablet Regimen
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
TAF	Tenofoviralfenamid
TDF	Tenofovirdisoproxilfumarat (300 mg Tenofovirdisoproxilfumarat entsprechen 245 mg Tenofovirdisoproxil)
TLOVR	Time to Loss Of Virologic Response
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschtes Ereignis
UK	United Kingdom
ULN	Obergrenze des Normbereichs (upper limit of normal)
USA	Vereinigte Staaten von Amerika
WHO	World Health Organization
ZN	Zusatznutzen

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

Ziel der vorliegenden Untersuchung ist die Bewertung des medizinischen Nutzens und Zusatznutzens des Single-Tablet-Regimen (STR) Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil (im Folgenden: EVG/COBI/FTC/TDF) im Vergleich zur zweckmäßigen Vergleichstherapie – diese ist eine individuelle antiretrovirale Therapie in Abhängigkeit der Vortherapie(en) und unter Berücksichtigung des Grundes für den Therapiewechsel, insbesondere Therapieversagen aufgrund eines virologischen Versagens und etwaig einhergehender Resistenzbildung oder aufgrund von Nebenwirkungen. Da es sich bei der festgelegten zweckmäßigen Vergleichstherapie um eine patientenindividuelle und entsprechend nicht substanzspezifizierte Therapie mit einer nicht überschaubaren Anzahl möglicher Regime handelt, wurden nach folgenden Kriterien drei für den größten Teil der Umstellungs-Population repräsentative Regime festgelegt:

1. EVG/COBI/FTC/TDF ist eine Dreifachtherapie, bestehend aus einem NRTI-Backbone und einer Substanz aus der Substanzklasse der 3. Partner, die als komplettes Regime nicht in Kombination mit anderen antiretroviralen Arzneimitteln eingesetzt werden darf [1]. Es ist davon auszugehen, dass die zweckmäßige Vergleichstherapie primär ebenfalls aus der Kombination eines Backbones mit einem 3. Partner besteht.
2. Substanzklassenauswahl: Ebenso wie in der Initialtherapie erhält der mit deutlichem Abstand größte Teil der Patienten auch nach Umstellung eine individualisierte Dreifachtherapie - primär bestehend aus einem NRTI-Backbone in Kombination mit einem 3. Partner aus der Substanzklasse der PI, der NNRTI oder der INI.
3. Substanzauswahl: Der mit großem Abstand am häufigsten eingesetzte Backbone ist FTC/TDF. EFV (NNRTI) und ATV/r (PI) gehören jeweils zu den am häufigsten verwendeten Vertretern ihrer Klasse; RAL ist der einzige Vertreter der INI. Diese Substanzen werden insofern stellvertretend für ihre jeweilige Substanzklasse untersucht.

Vor diesem Hintergrund, wird EVG/COBI/FTC/TDF mit den folgenden, für die zweckmäßige Vergleichstherapie repräsentativen Regimen verglichen:

- Efavirenz in Kombination mit Emtricitabin und Tenofoviridisoproxil (im Folgenden: EFV/FTC/TDF),
- Ritonavir-geboostertes Atazanavir in Kombination mit Emtricitabin und Tenofoviridisoproxil (im Folgenden: ATV/r+FTC/TDF) und
- Raltegravir in Kombination mit Emtricitabin und Tenofoviridisoproxil (im Folgenden: RAL+FTC/TDF).

Diese Regime repräsentieren die individualisierte zweckmäßige Vergleichstherapie und decken entsprechend den größten Teil der hier relevanten Patientenpopulation ab.

Die gemäß Indikation zugelassene Patientenpopulation umfasst zum einen nicht-vorbehandelte, erwachsene Patienten (≥ 18 Jahre), und zum anderen erwachsene Patienten (≥ 18 Jahre), bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, unabhängig von ihrem Behandlungsstatus.

Die Patientenpopulation der nicht-vorbehandelten Patienten (≥ 18 Jahre) umfasst auch die nicht-vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind. Der **nicht-vorbehandelte** Anteil der Gesamt-Population der Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, ist durch die Gesamt-Population der nicht-vorbehandelten Patienten abgedeckt. Somit werden in Modul 4B **vorbehandelte** Patienten (≥ 18 Jahre), bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, untersucht.

Die Untersuchung der Patientenpopulation, definiert als erwachsene Patienten (≥ 18 Jahre) mit einer HIV-1 Infektion, die antiretroviral nicht-vorbehandelt sind, ist in Modul 4A zu finden.

Datenquellen

Es wurde nach randomisierten kontrollierten Studien gesucht, da diese mit der größten Ergebnissicherheit behaftet sind. Allerdings stehen zum Zeitpunkt der Dossiereinreichung keine Daten zu vorbehandelten HIV-1-infizierten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, aus RCT zur Verfügung.

Es laufen jedoch aktuell zwei randomisierte kontrollierte Switch-Studien, in welchen die Patienten von Ihrer jeweiligen bisherigen Behandlung (GS-US-236-0115: Vorbehandlung mit Ritonavir-geboostertem Protease-Inhibitor (PI) plus FTC/TDF, GS-US-236-0121: Vorbehandlung mit nicht-nukleosidalen Reverse-Transkriptase-Inhibitor (NNRTI) plus FTC/TDF) auf EVG/COBI/FTC/TDF oder die weitere Einnahme der bisherigen Behandlung randomisiert wurden. Erste Ergebnisse aus diesen Studien werden Anfang des Jahres 2014 erwartet.

Unterstützend für die Übertragbarkeit der Daten der nicht-vorbehandelten Patienten auf die vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, werden 12-Wochen-Ergebnisse einer einarmigen Studie (GS-US-236-0123) präsentiert. Diese Daten werden jedoch aufgrund des kurzen Beobachtungszeitraums und des einarmigen Studiendesigns nicht zur Nutzen- bzw. Zusatznutzenableitung herangezogen (siehe Ein-/ Ausschlusskriterien).

Die auf die vorbehandelten Patienten zu übertragenden Daten der nicht-vorbehandelten Patienten stützen sich auf insgesamt vier RCT, die wie folgt zur Nutzenbewertung gegenüber der jeweiligen zweckmäßigen Vergleichstherapie herangezogen wurden.

Die Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF im Vergleich zur zweckmäßigen Vergleichstherapie EFV/FTC/TDF basiert auf zwei aktiv-kontrollierten randomisierten Studien des pharmazeutischen Unternehmers, die das zu bewertende Arzneimittel EVG/COBI/FTC/TDF mit EFV/FTC/TDF direkt vergleichen. Es werden Daten zu Woche 48, 60 und 96 präsentiert. Zusätzlich zu den Informationen aus dem Studienbericht werden post-hoc durchgeführte Subgruppenanalysen dargestellt. Die Ergebnisse aus diesen Studien werden in Modul 4A präsentiert.

Die Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF im Vergleich zur zweckmäßigen Vergleichstherapie ATV/r+FTC/TDF basiert auf einer aktiv-kontrollierten randomisierten Studie des pharmazeutischen Unternehmers, die das zu bewertende Arzneimittel EVG/COBI/FTC/TDF mit ATV/r+FTC/TDF direkt vergleicht. Es werden Daten zu Woche 48 und 96 präsentiert. Zusätzlich zu den Informationen aus dem Studienbericht wurden post-hoc durchgeführte Subgruppenanalysen dargestellt.

Die Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF im Vergleich zur zweckmäßigen Vergleichstherapie RAL+FTC/TDF basiert auf zwei aktiv-kontrollierten randomisierten Studien des pharmazeutischen Unternehmers, die das zu bewertende Arzneimittel EVG/COBI/FTC/TDF mit EFV/FTC/TDF direkt vergleichen, und außerdem einer randomisierten aktiv-kontrollierten Studie die RAL+FTC/TDF mit EFV+FTC/TDF direkt vergleicht. Es wird also ein adjustierter indirekter Vergleich über EFV+FTC/TDF als Brückenkompator durchgeführt. Es werden Daten zu Woche 48 und 96 präsentiert.

Ein-/Ausschlusskriterien für Studien

Die Einschlusskriterien für die einzuschließenden Studien mit dem zu bewertenden Arzneimittel EVG/COBI/FTC/TDF wurden wie folgt definiert:

- Population: Erwachsene vorbehandelte Patienten mit einer HIV-1 Infektion, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind.
- Intervention: Einmal täglich 150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (EVG/COBI/FTC/TDF)
- Vergleichstherapie:
 - Einmal täglich 600 mg Efavirenz, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (EFV/FTC/TDF)
 - Einmal täglich 300 mg Atazanavir, 100 mg Ritonavir, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (ATV/r+FTC/TDF)
 - Zweimal täglich 400 mg Raltegravir, einmal täglich 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (RAL+FTC/TDF)

- Endpunkte: Virologisches Ansprechen (Absenkung der Viruslast unter die Nachweisgrenze), Zeit bis zum Verlust des virologischen Ansprechens, Veränderung der CD4-Zellzahl, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, Therapieabbrüche aufgrund unerwünschter Ereignisse, unerwünschte Ereignisse Grad 3-4, wichtige unerwünschte Ereignisse (Hautausschläge, neurologische & psychiatrische Ereignisse), unerwünschte Ereignisse von Interesse (Renale Ereignisse, Knochenfrakturen), häufigste unerwünschte Ereignisse (Erkrankungen des Gastrointestinaltrakts [System Organ Class (SOC)], allgemeine Erkrankungen und Beschwerden am Verabreichungsort [SOC], Erkrankungen des Nervensystems [SOC], psychiatrische Erkrankungen [SOC], Erkrankungen der Haut und des Unterhautzellgewebes [SOC], Augenerkrankungen [SOC]). Der Endpunkt Augenerkrankungen (SOC) ist dabei nur in Zusammenhang mit ATV/r+FTC/TDF relevant.
- Studientyp: RCT
- Studiendauer: ≥ 48 Wochen
- Publikationstyp: Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht

Die Ausschlusskriterien für Studien mit dem zu bewertenden Arzneimittel EVG/COBI/FTC/TDF wurden wie folgt festgelegt:

- Population: Patienten mit anderen Infektionen; Kinder und Jugendliche
- Intervention: Andere Intervention oder abweichende Dosierungen
- Vergleichstherapie: Andere Vergleichstherapie oder abweichende Dosierungen
- Endpunkte: Keiner der unter den Einschlusskriterien definierten Endpunkte wird berichtet
- Studientyp: Nicht randomisierte Studien
- Studiendauer: < 48 Wochen
- Publikationstyp: Review-Artikel oder Case Reports; Keine Vollpublikation oder Bericht verfügbar oder es werden keine zusätzlichen Informationen dargestellt.

Für den Einschluss von Studien mit RAL+FTC/TDF (für den indirekten Vergleich) wurden folgende Kriterien definiert:

- Population: Erwachsene vorbehandelte Patienten mit einer HIV-1 Infektion, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind.
- Intervention: Zweimal täglich 400 mg Raltegravir, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (EVG/COBI/FTC/TDF)

- Vergleichstherapie: Aktiver Komparator, der als Brückenkomparator in Frage kommt (Es liegen nur RCT mit EVG/COBI/FTC/TDF im direkten Vergleich zu Efavirenz bzw. Ritonavir-geboostertem Atazanavir jeweils in Kombination mit TDF und FTC vor): Efavirenz oder Ritonavir-geboostertes Atazanavir jeweils in Kombination mit 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil
- Endpunkte: Virologisches Ansprechen (Absenkung der Viruslast unter die Nachweisgrenze), Zeit bis zum Verlust des virologischen Ansprechens, Veränderung der CD4-Zellzahl, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, Therapieabbrüche aufgrund unerwünschter Ereignisse, unerwünschte Ereignisse Grad 3-4, wichtige unerwünschte Ereignisse (Hautausschläge, neurologische & psychiatrische Ereignisse), unerwünschte Ereignisse von Interesse (Renale Ereignisse, Knochenfrakturen), häufigste unerwünschte Ereignisse (Erkrankungen des Gastrointestinaltrakts [SOC], allgemeine Erkrankungen und Beschwerden am Verabreichungsort [SOC], Erkrankungen des Nervensystems [SOC], psychiatrische Erkrankungen [SOC], Erkrankungen der Haut und des Unterhautzellgewebes [SOC])
- Studientyp: RCT
- Studiendauer: ≥ 48 Wochen, Daten zu Woche 48 oder 96
- Publikationstyp: Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht

Die Ausschlusskriterien für Studien mit RAL+FTC/TDF wurden wie folgt definiert:

- Population: Patienten mit anderen Infektionen, vorbehandelte Patienten oder Kinder und Jugendliche
- Intervention: Andere Intervention oder abweichende Dosierungen
- Vergleichstherapie: Andere Vergleichstherapie oder abweichende Dosierungen
- Endpunkte: Keiner der unter den Einschlusskriterien definierten Endpunkte wird berichtet
- Studientyp: Nicht randomisierte Studien
- Studiendauer: < 48 Wochen, Daten nur zu anderen Zeitpunkten als Woche 48 oder 96
- Publikationstyp: Review-Artikel oder Case Reports. Keine Vollpublikation oder Bericht verfügbar oder es werden keine zusätzlichen Informationen dargestellt.

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

Die eingeschlossenen Studien wurden mittels CONSORT-Kriterien beschrieben und das Verzerrungspotenzial auf Studien- und Endpunktebene wurde gemäß den Vorgaben aus Anhang 4-G bewertet.

Zur Synthese von Ergebnissen wurden, soweit sinnvoll und möglich, Meta-Analysen und Sensitivitätsanalysen durchgeführt. Des Weiteren wurden Subgruppenanalysen für alle eingeschlossenen Endpunkte durchgeführt.

Die nicht-randomisierte einarmige Studie GS-US-236-0123 wurde mittels TREND-Statement beschrieben und das Verzerrungspotenzial wurde auf Studien- und Endpunktebene gemäß den Vorgaben aus der Dossiervorlage bewertet.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Generell finden hier aufgrund der Übertragbarkeit der Daten die Ergebnisse aus dem Vergleich von EVG/COBI/FTC/TDF mit EFV/FTC/TDF bei nicht-vorbehandelten HIV-infizierten Patienten Anwendung – diese Daten wurden vollumfänglich im Modul 4A dargestellt (siehe Modul 4A). Auf der gleichen Basis der Übertragbarkeit der Daten finden die im vorliegenden Modul 4B vorgestellten Ergebnisse aus dem Vergleich von EVG/COBI/FTC/TDF mit ATV/r+FTC/TDF und mit RAL+FTC/TDF bei nicht-vorbehandelten HIV-infizierten Patienten Anwendung.

Lediglich zur Unterstützung der Übertragbarkeit der Daten werden im Folgenden die Ergebnisse der einarmigen Studie GS-US-236-0123 dargestellt:

Studie GS-US-236-0123 ist eine einarmige, offene Phase 3B-Pilotstudie zur Untersuchung der Umstellung eines Regimes aus Raltegravir (RAL) in Kombination mit einer Fixkombination aus Emtricitabin/Tenofovirdisoproxil (FTC/TDF) auf das Single-Tablet-Regimen (STR) Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofovirdisoproxil (EVG/COBI/FTC/TDF) in virologisch supprimierten HIV-1-infizierten Patienten.

Das virologische Ansprechen lag nach Umstellung auf und 12 Wochen Behandlung mit EVG/COBI/FTC/TDF bei 100%: Bei allen Patienten konnte die Viruslast unter der Nachweisgrenze gehalten werden.

Zu Baseline wiesen die Patienten aufgrund der mindestens 6-monatigen Vorbehandlung bereits eine hohe CD4-Zellzahl von 711 Zellen/ μ L auf. Diese stieg nach 12 Wochen Behandlung mit EVG/COBI/FTC/TDF nochmals um durchschnittlich 33 Zellen/ μ L an. Allerdings ist der Unterschied zur Baseline nicht statistisch signifikant ($p=0,25$).

In der Studie traten keine unerwünschten Ereignisse, höher als Grad 2, auf. 36 (75,0%) Patienten hatten mindestens ein unerwünschtes Ereignis, davon wurde es bei elf (22,9%) Patienten als Grad 2 eingestuft. Es gab keine Todesfälle und es traten weder schwerwiegende

unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4 noch Therapieabbrüche aufgrund unerwünschter Ereignisse auf.

Bei 12 (25,0%) Patienten traten gastrointestinale Nebenwirkungen auf; davon bei fünf (10,4%) Patienten Diarrhoe.

Psychiatrische unerwünschte Ereignisse traten bei elf (22,9%) der Patienten auf; das häufigste Ereignis war Schlaflosigkeit (fünf [10,4%] Patienten).

Insgesamt weist EVG/COBI/FTC/TDF nach Umstellung von RAL+FTC/TDF zu Woche 12 somit eine gute Wirksamkeit und Verträglichkeit auf.

Im Folgenden werden die Daten der Vergleiche zu nicht-vorbehandelten Patienten, die auf die Population der vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, übertragen werden, dargestellt. Alle in diese Vergleiche eingeschlossenen Studien waren randomisierte kontrollierte Studien und entsprechen demnach der Evidenzstufe Ib.

EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Die Ergebnisse aus dem Vergleich von EVG/COBI/FTC/TDF mit EFV/FTC/TDF finden sich in Modul 4A– diese sind vollständig auf die hier vorliegende Patientenpopulation zu übertragen.

EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF

Im Folgenden (Tabelle 4-1) wird zusammenfassend für die Vergleiche zwischen EVG/COBI/FTC/TDF und ATV/r+FTC/TDF die Ableitung des Zusatznutzens auf Endpunktebene dargestellt. Dabei werden nur die Endpunkte berücksichtigt, für die ein Zusatznutzen bzw. Schaden durch EVG/COBI/FTC/TDF festgestellt werden konnte.

Tabelle 4-1: Gegenüberstellung von Zusatznutzen und Schaden von EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF

Zusatznutzen durch EVG/COBI/FTC/TDF <u>Format:</u> Endpunkt Population: Wahrscheinlichkeit und Ausmaß des ZN Behandlungseffekt (Zeitpunkt)	Schaden durch EVG/COBI/FTC/TDF <u>Format:</u> Endpunkt Population: Wahrscheinlichkeit und Ausmaß des Schadens Behandlungseffekt (Zeitpunkt)
<ul style="list-style-type: none"> • Gesamtrate SUEs >100.000 Kopien/mL: Anhaltspunkt für geringen ZN OR: 0,45 [0,21;0,94], p=0,0325 (Woche 96) • UEs Grad 3-4 >100.000 Kopien/mL: Hinweis auf erheblichen ZN OR: 0,32 [0,15;0,67], p=0,0025 (Woche 48) OR: 0,37 [0,18;0,74], p=0,0050 (Woche 96) • Knochenfrakturen Gesamt: Anhaltspunkt für geringen ZN OR: 0,28 [0,09;0,86], p=0,029 (Woche 96) • Erkrankungen des Gastrointestinaltrakts (SOC) Nicht-kaukasisch: Hinweis auf beträchtlichen ZN^a OR: 0,46 [0,25;0,84], p=0,0113 (Woche 48) OR: 0,42 [0,23;0,78], p=0,0057 (Woche 96) ≤100.000 Kopien/mL: Anhaltspunkt für geringen ZN OR: 0,65 [0,44;0,97], p=0,0328 (Woche 96) >350 Zellen/μL: Anhaltspunkt für geringen ZN OR: 0,61 [0,40;0,92], p=0,0202 (Woche 96) Diarrhoe >350 Zellen/μL: Hinweis auf geringen ZN OR: 0,49 [0,30;0,80], p=0,0042 (Woche 48) OR: 0,59 [0,37;0,93], p=0,0240 (Woche 96) • Augenerkrankungen (SOC) Gesamt: Hinweis auf beträchtlichen ZN OR: 0,19 [0,11;0,33], p<0,00001 (Woche 48) OR: 0,27 [0,17;0,43], p<0,00001 (Woche 96) Skleraler Ikterus Gesamt: Hinweis auf beträchtlichen ZN OR: 0,03 [0,01;0,14], p<0,00001 (Woche 	<ul style="list-style-type: none"> • Veränderung der CD4-Zellzahl ≤350 Zellen/μL: Anhaltspunkt für geringen Schaden MWD: -40,00 [-77,16;-2,84], p=0,0350 (Woche 96) • UEs Grad 3-4 ≤100.000 Kopien/mL: Anhaltspunkt für geringen Schaden OR: 1,95 [1,08;3,52], p=0,0262 (Woche 48)

Zusatznutzen durch EVG/COBI/FTC/TDF <u>Format:</u> Endpunkt Population: Wahrscheinlichkeit und Ausmaß des ZN Behandlungseffekt (Zeitpunkt)	Schaden durch EVG/COBI/FTC/TDF <u>Format:</u> Endpunkt Population: Wahrscheinlichkeit und Ausmaß des Schadens Behandlungseffekt (Zeitpunkt)
48) OR: 0,03 [0,01;0,14], p<0,00001 (Woche 96)	
ZN = Zusatznutzen ^a Der in der nicht-kaucasischen Subgruppe gefundene Unterschied bei den Erkrankungen des Gastrointestinaltrakts (SOC) zugunsten von EVG/COBI/FTC/TDF lässt sich insbesondere auch auf Basis der vorliegenden Evidenz (eher höhere Nebenwirkungsraten unter Nicht-Kaukasiern) nicht schlüssig begründen und wird für die Nutzenableitung nicht weiter berücksichtigt.	

EVG/COBI/FTC/TDF vs. RAL+FTC/TDF

Im Vergleich von EVG/COBI/FTC/TDF mit RAL+FTC/TDF wurde bei keinem der betrachteten Endpunkte ein statistisch signifikanter Unterschied zwischen den Behandlungen gefunden. Lediglich hinsichtlich der Gesamtrate unerwünschter Ereignisse zu Woche 48 wurde ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF gefunden (OR [95%-KI]: 2,79 [1,08;7,20]). Da die Ausprägung des Effektes allerdings nicht stark genug war, kann daraus kein Schaden abgeleitet werden.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Bei den Therapieprinzipien der antiretroviralen Therapie (ART) gilt es als generelles Ziel der Therapie, durch die Hemmung der HIV-Replikation infektionsbedingte Symptome zu unterdrücken, Krankheitsprogression zu vermindern, eine immunologische Rekonstitution zu erreichen und die chronische Immunaktivierung zu reduzieren. Nur eine hohe virologische Wirksamkeit der ART und eine dauerhafte Suppression der Viruslast auf unter 50 RNA-Kopien/mL können Resistenzentwicklung und konsekutives Therapieversagen nachhaltig verhindern [2].

Grundsätzlich unterscheidet man in der ART zwei Patientenpopulationen, für die die übergeordneten Therapieprinzipien (dauerhafte Hemmung der Replikation und virale Suppression auf <50 Kopien/mL) gleichermaßen gültig sind:

1. Patienten, die noch keine antiretrovirale Therapie erhalten (nicht-vorbehandelte Patienten):
2. Patienten, die antiretroviral behandelt werden (vorbehandelte Patienten)

In den Empfehlungen zur Umstellung antiretroviral vorbehandelter Patienten wird in allen nationalen und internationalen Leitlinien zwischen zwei Subpopulationen, resp. Indikationen zur Umstellung unterschieden [2-6]:

- Supprimierte Patienten: Umstellung aufgrund von Nebenwirkungen oder zur Therapievereinfachung und Unterstützung, resp. Verbesserung der Adhärenz
- Patienten von mit virologischem Versagen: Umstellung aufgrund des virologischen Versagens und konsekutiver Virämie

Diese Unterscheidung ergibt sich aus dem entscheidenden Kriterium, das die Grundvoraussetzung für die virologische Wirksamkeit einer ART darstellt: dem Vorliegen Resistenz-assoziiierter Mutationen gegen die einzusetzenden Substanzen.

Die Unterscheidung zwischen Patienten mit und ohne Resistenz-assoziierten Mutationen ist von elementarer Bedeutung: Ein vorbehandelter Patient, dessen Virus keine Resistenz-assoziierten Mutationen gegen antiretrovirale Substanzen aufweist, entspricht hinsichtlich des Resistenz-Status einem nicht-vorbehandelten Patienten. Damit ist dieser Patient im Hinblick auf die entsprechenden antiretroviralen Substanzen als funktionell nicht-vorbehandelt zu definieren - womit beide Populationen (nicht-vorbehandelt und vorbehandelt ohne Resistenzen) als gleichgestellt zu betrachten sind. Vor dem Hintergrund der Gleichstellung der Population können klinische Daten zu nicht-vorbehandelten Patienten somit auf vorbehandelte Patienten ohne Resistenzen extrapoliert und damit übertragen werden [7].

Die Rationale der Übertragbarkeit klinischer Daten von nicht-vorbehandelten Patienten auf vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, lässt sich zudem wie folgt belegen:

1. Die DHHS-Leitlinien führen in den Empfehlungen zur Umstellung supprimierter Patienten explizit aus: „Für Regime, die zur Therapievereinfachung eingesetzt werden, sollte grundsätzlich eine hohe Wirksamkeit bei antiretroviral nicht-vorbehandelten Patienten nachgewiesen worden sein, oder es sollte für den Patienten anhand seiner individuellen Vorbehandlung und seines Resistenzprofils eine hohe Wirksamkeit vorhergesagt werden können“ [4]. Gemäß der DHHS-Leitlinie ist somit davon auszugehen, dass die nachgewiesene Wirksamkeit eines antiretroviralen Regimes bei nicht-vorbehandelten Patienten auf supprimierte vorbehandelte Patienten extrapoliert und damit übertragen werden kann.
2. Die Übertragbarkeit der Daten ist hinsichtlich der Wirksamkeit und des Verträglichkeitsprofils auch rein mechanistisch und zwar sowohl pharmakologisch als auch biologisch nachvollziehbar:

Eine Initialtherapie muss ausreichend virologisch wirksam sein, um das Virus erfolgreich zu kontrollieren und die Ausgangsviruslast unabhängig von ihrer Höhe (etwa 50% der Patienten weisen eine Ausgangsviruslast von >100.000 Kopien/mL auf) unter die Nachweisgrenze zu senken. Antiretrovirale Therapieregime, die zur Umstellung supprimierter Patienten eingesetzt werden, hingegen müssen lediglich ausreichend virologisch wirksam sein, um die Viruslast unter der Nachweisgrenze zu halten. Entsprechend muss die virologische Potenz einer antiretroviralen Initialtherapie mindestens so hoch sein wie die eines Regimes, auf das ein supprimierter Patient umgestellt wird. Vor diesem Hintergrund ist davon auszugehen, dass klinische Daten zur virologischen Wirksamkeit, die bei nicht-vorbehandelten Patienten erhoben werden, direkt auf die Umstellung vorbehandelter Patienten extrapolierbar sind, sofern diese keine Resistenzen gegen die Substanzen aufweisen.

Unter Therapie auftretende Nebenwirkungen sind primär Substanz- und nicht Virus-assoziiert, woraus schlussgefolgert werden kann, dass die Höhe der Viruslast zu Beginn einer bestimmten Therapie keinen Einfluss auf die auftretenden Nebenwirkungen hat. Entsprechend wird auch im Bericht zum Sicherheitsprofil in den Fachinformationen [8-35] zu den verschiedenen antiretroviralen Substanzen nicht zwischen vorbehandelten und nicht-vorbehandelten Patienten unterschieden – ganz im Gegenteil beruht das in den Fachinformationen berichtete Sicherheitsprofil stets auf Basis des gesamten klinischen Entwicklungsprogramms.

3. Die European Medicines Agency (EMA) hat für Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil sowohl die Zulassung für nicht-vorbehandelte Patienten, als auch die Zulassung für die Patienten erteilt, die keine Resistenzen gegen einen der drei enthaltenen antiretroviralen Substanzen aufweisen. Da diese Zulassung ausschließlich auf Basis von klinischen Daten zu nicht-vorbehandelten Patienten eingereicht wurde, muss geschlussgefolgert werden, dass die EMA die beiden Population auf Basis der dargestellten wissenschaftlichen Rationale funktionell gleichgestellt und die Daten entsprechend extrapoliert hat. Dies wird auch durch ein von der EMA erstelltes Konzeptpapier zur Extrapolation von klinischen Daten klar unterstützt: Die angestrebte Extrapolation ist definiert als Deduktion von Studienergebnissen in bestimmten Subpopulation (Quellenpopulation) auf andere Subpopulationen (Zielpopulation) – analog gilt dies für verwandte Erkrankungen oder Substanzen. Zielsetzung der Extrapolation von Daten ist zum einen die Reduktion des Bedarfs der zusätzlichen Datengenerierung für die Zielpopulation, und zum anderen die Vermeidung unnötiger Studien und Allokation von Ressourcen in die klinischen Studien, die tatsächlich notwendig sind. Grundvoraussetzung für die Übertragbarkeit ist eine pharmakologische und biologische Rationale, die für Quellen- und Zielpopulation anhand von Erkrankung, Medikation (Wirkmechanismus, Pharmakokinetik und Pharmakodynamik) sowie Anwendbarkeit von Wirksamkeits- und Sicherheitsendpunkten bewertet werden sollte [7]. Für die beiden hier vorliegenden Populationen der nicht-vorbehandelten Quellenpopulation und der Zielpopulation von Patienten ohne Resistenz-assoziierte Mutationen, zeigt sich die geforderte pharmakologische und biologische Rationale deutlich bei der vollständigen

Übereinstimmung der genannten Parameter. Vor diesem Hintergrund ist mit Erteilung der Zulassung also davon auszugehen, dass die Europäische Zulassungsbehörde die Übertragbarkeit der Daten von den nicht-vorbehandelten Patienten auf die vorbehandelten Patienten ohne Resistenz-assoziierte Mutationen als gegeben ansieht.

In logischer Konsequenz bedeutet die dargestellte grundsätzliche Übertragbarkeit der Daten, dass auch ein in den klinischen Studien zu nicht-vorbehandelten Patienten gefundener Zusatznutzen für auf die vorbehandelte Patientenpopulation, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, zu übertragen ist.

EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Für die Population der erwachsenen nicht-vorbehandelten Patienten wurde in der Gesamtschau für EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF ein positiver Zusatznutzen auf Endpunktebene abgeleitet, dessen Ausmaß als gering eingestuft wird. Da zwei randomisierte, aktiv-kontrollierte Studien zugrunde liegen, wird dieser Zusatznutzen mit der Wahrscheinlichkeit eines Belegs eingestuft (siehe Modul 4A).

Entsprechend der Rationale zur Übertragbarkeit der Daten besteht somit auch ein **Beleg** für einen **geringen** Zusatznutzen durch EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF, für die Population der vorbehandelten erwachsenen mit HIV-1 infizierten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind. Für weitere Ausführungen und Details zur Zusatznutzenableitung gegenüber EFV/FTC/TDF siehe Modul 4A.

EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF

Zusammenfassend wurde im Vergleich von EVG/COBI/FTC/TDF mit ATV/r+FTC/TDF bei nicht-vorbehandelten Patienten, dessen Ergebnisse auf die in diesem Modul 4B dargestellte Patientenpopulation zu übertragen sind, anhand von patientenrelevanten und validen Endpunkten gezeigt, dass EVG/COBI/FTC/TDF, bei gleichwertiger Wirksamkeit wie ATV/r+FTC/TDF (Nichtunterlegenheit zur 12%-Grenze wurde gezeigt), eine vom Nebenwirkungsprofil verbesserte Behandlungsoption darstellt als die zweckmäßige Vergleichstherapie ATV/r+FTC/TDF. Durch EVG/COBI/FTC/TDF kann der durch die zweckmäßige Vergleichstherapie ausgelöste patientenrelevante Schaden vermindert werden.

Es konnte gezeigt werden, dass EVG/COBI/FTC/TDF eine gleichwertige bzw. nichtunterlegene Wirksamkeit wie ATV/r+FTC/TDF besitzt. Dies zeigt sich anhand der Endpunkte virologisches Ansprechen, Zeit bis zum Verlust des virologischen Ansprechens und Veränderung der CD4-Zellzahl.

Bei sicherheitsbezogenen Endpunkten zeigten sich unter EVG/COBI/FTC/TDF deutliche patientenrelevante Vorteile hinsichtlich der Vermeidung von schwerwiegenden unerwünschten Ereignissen, der Vermeidung unerwünschter Ereignisse Grad 3-4, von Knochenfrakturen, Erkrankungen des Gastrointestinaltrakts (darunter Diarrhoe) und von

Augenerkrankungen (darunter skleraler Ikterus). Bei der Ableitung eines Gesamtfazits zum Zusatznutzen werden die Gruppen von PTs (neurologische & psychiatrische Ereignisse und Hautausschläge) nicht als Zusatzinformation betrachtet, da sie durch die standardisierten SOC Erkrankungen des Nervensystem, psychiatrische Erkrankungen und Erkrankungen der Haut und des Unterhautzellgewebes abgedeckt werden. Sie tragen somit kein Gewicht zur Ermittlung eines gesamten Zusatznutzens bei.

Bei der Verträglichkeit ist nicht nur der rein medizinische Aspekt von Bedeutung, sondern bei bestimmten Nebenwirkungen im gleichen Ausmaß auch der psychologische: Die für den Wirkstoff ATV typische Nebenwirkung skleraler Ikterus beispielsweise, ist ein äußerlich sichtbarer Effekt, der zu Fragen hinsichtlich der Ursachen und somit zur unbeabsichtigten Bekanntgabe einer HIV-Infektion führen kann. Dies ist besonders in der homosexuellen Community, die gut vernetzt und zum Thema HIV gut informiert ist, von hoher Relevanz. Da unerwünschte Ereignisse wie skleraler Ikterus – zumal dieser sich nicht verbergen lässt - im Kreis der HIV-infizierten Bevölkerung direkt stigmatisierend wirken können, ist es von besonderem Interesse, Ereignisse dieser Art weitgehend zu vermeiden. Die Stigmatisierung eines HIV-infizierten Patienten kann auch weitere Erkrankungen, so beispielsweise Depressionen, zur Folge haben. Gerade für Depressionen wiederum konnte in vielen Studien gezeigt werden, dass diese einen negativen Einfluss auf die Adhärenz haben, was in der Folge zu einem Wiederanstieg der Viruslast mit konsekutivem virologischen Therapieversagen, immunologischer Verschlechterung und erhöhter Morbidität und Mortalität führen kann [36, 37].

Der hier beobachtete Vorteil durch EVG/COBI/FTC/TDF hinsichtlich der Augenerkrankungen ist als weitest gehende Vermeidung solch stigmatisierender Nebenwirkungen zu verstehen. Dies hat nicht nur einen Einfluss auf die Lebensqualität der Patienten, sondern auch auf den Behandlungserfolg, da ein wichtiger die Therapieadhärenz negativ beeinflussender Faktor unter der Behandlung mit EVG/COBI/FTC/TDF nicht besteht, wodurch das Risiko eines Anstiegs der Viruslast mit nachfolgendem immunologischem Versagen [38-41], das vermehrte Risiko für Resistenzentwicklung und schließlich klinischer Progression zu Aids oder Tod [42] reduziert wird. Dies ist angesichts einer Zahl von 25-30% der Patienten, die im ersten Jahr ihre ART wechseln, ein erstrebenswertes Ziel [43, 44]. Durch diese bessere Verträglichkeit sinkt insgesamt das Risiko eines durch den Patienten gewünschten Therapiewechsels, wodurch die Dauerhaftigkeit des Erstregimes positiv beeinflusst wird. Das langfristige Verbleiben auf einem Regime erhält die bei einem nötigen Therapiewechsel weiterhin zur Verfügung stehenden Behandlungsoptionen. Zudem zeigen Daten, dass die Dauer der viralen Suppression einen Prädiktor für eine dauerhafte virale Suppression darstellt, d. h. mit der Dauer der viralen Suppression sinkt die Wahrscheinlichkeit für einen virologischen Rebound [45, 46].

Daten zeigen, dass nicht nur die antiretrovirale Medikation – insbesondere geboosterte Protease-Inhibitoren - sondern auch die HIV-Infektion selbst und der immunologische Status des Patienten einen Einfluss auf den Gastrointestinaltrakt haben [47]. Das könnte zur Folge haben, dass die Verträglichkeitsunterschiede zwischen den Behandlungen bei den kränkeren

Patienten (Ausgangsviruslast >100.000 Kopien/mL, resp. CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL) durch den zusätzlichen Einfluss der HIV-Infektion überlagert werden und sich die tatsächlichen Unterschiede der Medikation erst bei einem besseren immunologischen Status zeigen, da der additive negative Effekt der HIV-Infektion gewissermaßen wegfällt. Dies korreliert mit den gefundenen Ergebnissen: Bei den Erkrankungen des Gastrointestinaltrakts zeigt sich ein Anhaltspunkt für einen geringen Zusatznutzen für EVG/COBI/FTC/TDF bei Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL, resp. CD4-Zellzahl zu Baseline >350 Zellen/ μL . Darüber hinaus zeigt eine integrierte Analyse der Studien S102 und S103, dass die Inzidenz der gastrointestinalen Nebenwirkungen, darunter Nausea, unter EVG/COBI/FTC/TDF zu Beginn der Therapie am höchsten ist und dann über die Zeit deutlich abnimmt, wohingegen sie unter ATV/r+FTC/TDF zwar ebenfalls abfällt, sich jedoch auf einem deutlichen höheren Niveau einpendelt [48].

Insgesamt zeigen verschiedene Daten, dass beeinträchtigende Nebenwirkungen generell einen negativen Einfluss auf die Therapieadhärenz und damit Persistenz haben, was wiederum das Risiko der Bildung von Resistenzen und virologischem Versagen mit konsekutiver klinischer Progression erhöht. Da jedoch gerade in der Behandlung der HIV-Infektion eine hohe Therapieadhärenz unabdingbar für einen Behandlungserfolg ist und einen direkten Einfluss auf Krankheitsprogression, Morbidität und Mortalität haben, spielt die Vermeidung von unerwünschten Ereignissen, welche die Therapieadhärenz negativ beeinflussen könnten, eine zentrale Rolle [4, 49, 50].

Neben den positiven sicherheitsbezogenen Aspekten konnten auch einige Nachteile von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF beobachtet werden. Dies zeigte sich bei der Veränderung der CD4-Zellzahl bei Patienten mit CD4-Zellzahl bei Baseline ≤ 350 Zellen/ μL und unerwünschten Ereignissen Grad 3-4 bei Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL.

Der Anhaltspunkt für einen geringen Schaden hinsichtlich der Veränderung der CD4-Zellzahl durch EVG/COBI/FTC/TDF besteht nur zu Woche 96 in der Subgruppe der Patienten, die zu Baseline eine CD4-Zellzahl ≤ 350 Zellen/ μL aufwiesen. Dieser Schaden wirkt sich aus folgenden Gründen nicht auf die allgemeine Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF aus, die als äquivalent bzw. nichtunterlegen zu ATV/r+FTC/TDF eingeschätzt wird: Zum einen besteht der Schaden lediglich in einer Subgruppe und wird auch nur mit der Wahrscheinlichkeit eines Anhaltspunkts bewertet; zum anderen bedeuten die absoluten CD4-Zellzahlen unter EVG/COBI/FTC/TDF zu Woche 96 (488 CD4-Zellen/ μL unter EVG/COBI/FTC/TDF versus 539 CD4-Zellen/ μL unter ATV/r+FTC/TDF) einen guten immunologischen Status, so dass der Unterschied zwischen den Behandlungen zwar statistisch signifikant ist, nicht jedoch klinisch relevant.

Dem Anhaltspunkt für einen geringen Schaden hinsichtlich der Veränderung der CD4-Zellzahl in der Subgruppe mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL stehen ein Anhaltspunkt für einen geringen Zusatznutzen in der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL hinsichtlich der Vermeidung schwerwiegender unerwünschter Ereignisse, und darüber hinaus ein Anhaltspunkt für einen geringen

Zusatznutzen hinsichtlich der Vermeidung gastrointestinaler Erkrankungen in der weniger stark erkrankten Population (Ausgangsviruslast ≤ 100.000 Kopien/mL bzw. CD4-Zellzahl zu Baseline >350 Zellen/ μ L) gegenüber. Da die CD4-Zellzahl zu Woche 96 wie beschrieben oberhalb des kritischen Bereichs von <200 Zellen/ μ L lag, stellt die Verminderung schwerwiegender unerwünschter Ereignisse in der Patientengruppe mit Ausgangsviruslast >100.000 Kopien/mL sowie die Verminderung gastrointestinaler Ereignisse in der weniger stark erkrankten Population (Ausgangsviruslast ≤ 100.000 Kopien/mL bzw. CD4-Zellzahl zu Baseline >350 Zellen/ μ L) einen wesentlich patientenrelevanteren Aspekt dar.

Der Anhaltspunkt für einen geringen Schaden hinsichtlich unerwünschter Ereignisse Grad 3-4 durch EVG/COBI/FTC/TDF in der Subgruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL zu Woche 48 steht einem Hinweis auf einen erheblichen Zusatznutzen durch EVG/COBI/FTC/TDF bei diesem Endpunkt in der Subgruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL zu Woche 48 und 96 gegenüber. Die Gegenläufigkeit der Effekte in den beiden Viruslast-Strata ist medizinisch nicht plausibel. Allerdings besteht der Anhaltspunkt für einen geringen Schaden bei Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL lediglich zu Woche 48 und der statistisch signifikante Unterschied geht zu Woche 96 verloren, während sich der erhebliche Zusatznutzen bei den unerwünschten Ereignissen Grad 3-4 bei Patienten mit Ausgangsviruslast >100.000 Kopien/mL dauerhaft sowohl in Woche 48 als auch in Woche 96 zeigt. Insgesamt steht dem transienten geringen Schaden somit ein dauerhaft erheblicher Zusatznutzen durch EVG/COBI/FTC/TDF gegenüber.

In der Gesamtschau des auf Endpunktebene abgeleiteten Zusatznutzens verbleibt nach der Nutzen-Schaden-Abwägung ein positiver Zusatznutzen für EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF, dessen Ausmaß als **gering** eingestuft wird. Die Wahrscheinlichkeit dieses Zusatznutzens wird als **Hinweis** eingestuft.

EVG/COBI/FTC/TDF vs. RAL+FTC/TDF

Der auf die vorliegende Patientenpopulation zu übertragende Zusatznutzen aus dem Vergleich von EVG/COBI/FTC/TDF mit RAL+FTC/TDF bei nicht-vorbehandelten Patienten ergibt sich folgendermaßen: Obschon EVG/COBI/FTC/TDF im indirekten Vergleich sowohl eine gleichwertige Wirksamkeit als auch ein äquivalentes Sicherheitsprofil zeigt wie RAL+FTC/TDF, weist EVG/COBI/FTC/TDF durch die Formulierung erhebliche patientenrelevante Vorteile auf:

Mit Raltegravir wurde im Jahr 2008 der erste Vertreter der Klasse der Integrase-Inhibitoren (INI) eingeführt. Von besonderer Bedeutung ist der neue Wirkmechanismus dieser Klasse, der zur Folge hat, dass auch Patienten mit (Primär)-Resistenzen gegenüber verfügbaren antiretroviralen Substanzen erfolgreich behandelt werden können. Darüber hinaus zeichnen sich INI durch eine sehr gute antiretrovirale Wirksamkeit mit distinkt rascher Absenkung der Viruslast bei gleichzeitig sehr guter Verträglichkeit aus.

Allerdings zeigte ist die vorgesehene Einmal-Gabe von Raltegravir aus pharmakokinetischen Gründen nicht möglich – die zweimal tägliche Gabe ist zum einen ein hochrelevanter Aspekt für die Adhärenz und führt zum anderen zwangsläufig zu einem Mehrtablettenregime.

Die Einmalgabe stellt einen wesentlichen Einflussfaktor auf die Adhärenz dar: Willig et al. [51] untersuchten Daten von 542 Patienten, die zwischen 2000 und 2007 ihre erste ART begonnen haben und diese Therapie mindestens 14 Tage lang einnahmen. Die Daten zeigen, dass die QD-Gabe der BID-Gabe im Hinblick auf die Persistenz der Therapie über die Zeit statistisch signifikant überlegen ist ($p=0,0002$ Abbildung 1).

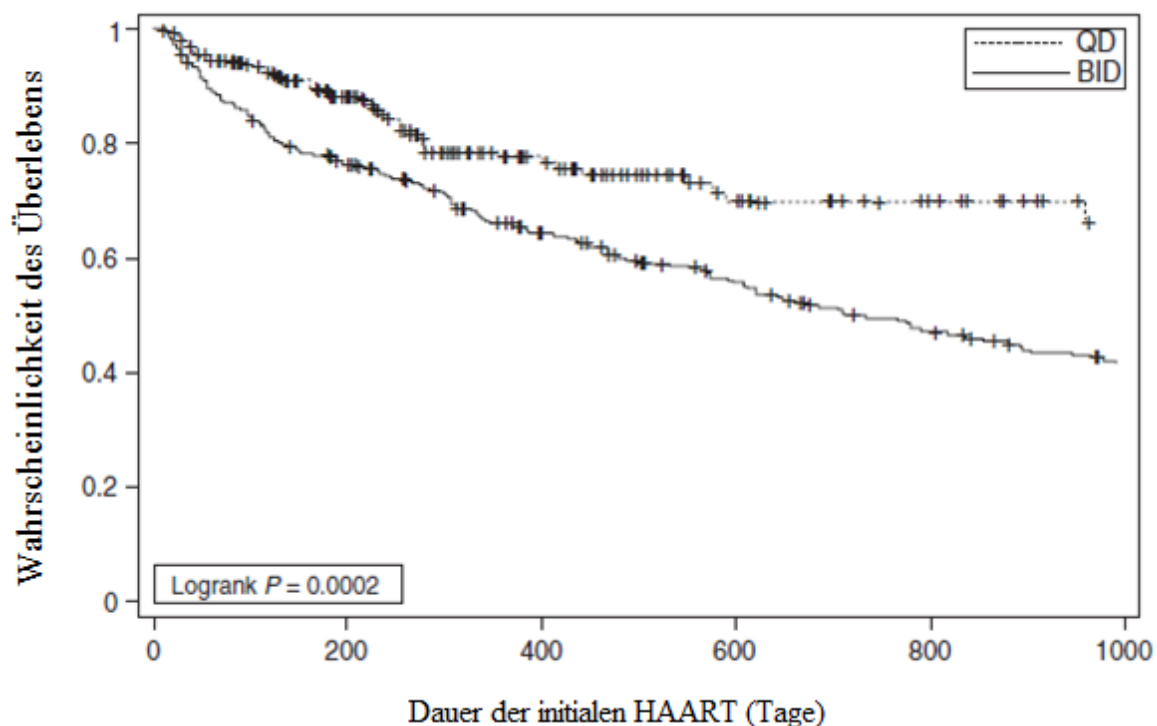


Abbildung 1: Dauerhaftigkeit (Persistenz) einer Initialtherapie als Funktion der Dosierungskomplexität [51]

Zu dem gleichen Schluss kommen Parienti et al. [52] in einer Meta-Analyse mit insgesamt elf klinischen Studien ($n=3.029$): Die Adhärenz-Rate unter einem einmal täglichen Regime war signifikant höher als unter einem zweimal täglichen Regime (+2,9%, 95%-KI: [1,0%;4,8%], $p<0,003$). Bei den fünf Studien mit nicht-vorbehandelten Patienten zeigte sich ($n=1.927$) unter einmal täglichen Regimen sowohl eine signifikant bessere Adhärenz (+4,4% 95%-KI: [1,8%;7,0%], $p<0,001$) als auch ein besseres virologisches Outcome (+5,7%, 95%-KI: [0,7%;10,8%], $p<0,027$) (Abbildung 2).

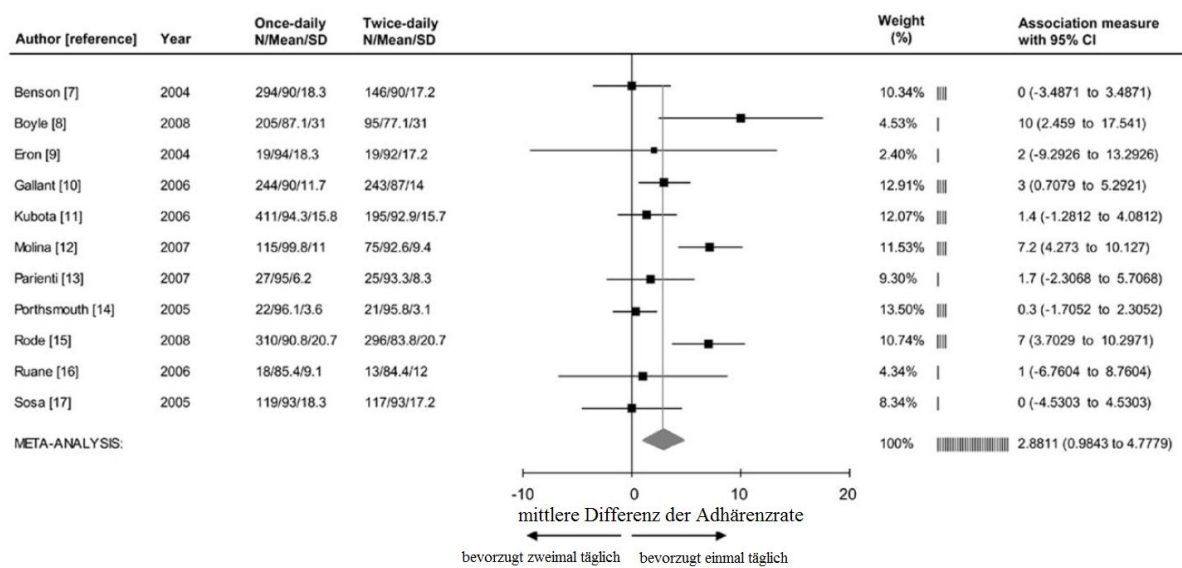


Abbildung 2: Forestplot des Effekts einer Einmal- im Vergleich zu einer Zweimal-Gabe auf die Adhärenzrate [52]

Darüber hinaus birgt die zweimal tägliche Gabe ebenso wie die Gabe von Mehrtabletten-Regimen das Risiko der selektiven Nicht-Adhärenz, das heißt der selektiven Einnahme, resp. Nicht-Einnahme bestimmter Substanzen des antiretroviralen Regimes – sei es unbeabsichtigt durch Vergessen oder beabsichtigt zur Vermeidung von Nebenwirkungen. Der Einsatz eines Single-Tablet-Regimen hat zur Folge, dass durch die Formulierung automatisch tatsächlich alle notwendigen Komponenten seines Regimes eingenommen werden - das heißt, die selektive Einnahme, resp. Nicht-Einnahme wird wirksam verhindert. Dies führt, wie im Folgenden dargestellt, zu einem verbesserten klinischen Outcome hinsichtlich der virologischen Suppression und reduzierter Hospitalisierungsraten und ist damit hochrelevant für den Patienten.

Eine Analyse klinischer Daten von 1.604 Patienten in Italien ergab in der COMPACT-Studie eine Adhärenz von $80,4 \pm 14,7\%$ für Patienten unter Single-Tablet-Regimen - unter Mehrtablettenregimen war die Adhärenz von Patienten aufgrund der beobachteten selektiven Nicht-Adhärenz (Tage ohne NRTI-Backbone oder 3. Partner) teils deutlich niedriger: Die Rate an selektiver Nicht-Adhärenz betrug 6% unter PI-basierten, 3% unter NNRTI-basierten und 13% unter Raltegravir-basierten Regimen (Abbildung 3). Die klinische Relevanz der gefundenen Ergebnisse zeigt sich in der Korrelation zwischen Adhärenz- und therapeutischer Erfolgsrate, gemessen anhand von CD4-Zellzahl (% Patienten mit CD4-Zellzahl >500 Zellen/ μL) und Viruslast (% Patienten mit <50 RNA-Kopien/mL). Der Anteil der Patienten mit Viruslast <50 Kopien/mL lag bei Patienten mit Single-Tablet-Regimen bei 96% - im Vergleich dazu bei 78% unter PI-basierten Regimen, 88% unter NNRTI-basierten Regimen und 87% unter INI (Raltegravir)-basierten Regimen. Auch die immunologischen Erfolgsraten korrelierten analog: Der Anteil der Patienten mit CD4-Zellzahl >500 Zellen/ μL lag bei Patienten mit Single-Tablet-Regimen bei 61% - im Vergleich dazu bei 44% unter PI-

basierten Regimen, 48% unter NNRTI-basierten Regimen und 42% unter INI (Raltegravir)-basierten Regimen.

Dies hatte bei Patienten mit einer Adhärenz-Rate von $\leq 80\%$ klinische Folgen: Patienten mit selektiver Nicht-Adhärenz zeigten je nach Ausmaß der selektiven Nicht-Adhärenz ein um 24%, resp. 39% erhöhtes Risiko für eine Hospitalisierung [53].

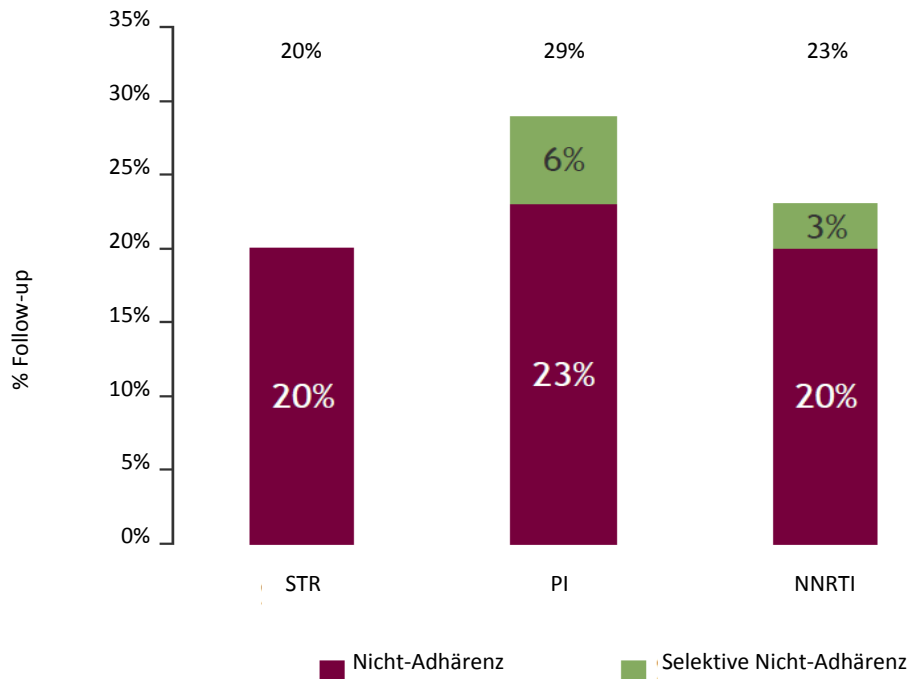


Abbildung 3: Nicht-Adhärenz bei verschiedenen ART-Regimen [53]

Bestätigende Daten liefert eine weitere italienische Kohorte: Hier war die (selbstberichtete) Nicht-Adhärenz bei Patienten unter einem Single-Tablet-Regimen deutlich niedriger als bei Patienten, die Mehrtablettentregime einnahmen [54].

Der Einfluss der selektiven Nicht-Adhärenz auf das klinische Outcome wurde auch auf Basis von Daten von 6.938 Patienten in der US-amerikanischen Medicaid Claims Datenbank (Zeitraum: Juni 2009 – Dezember 2011) untersucht: Die Daten-Analyse zeigte eine höhere Gesamt-Adhärenz für Patienten unter einem Single-Tablet-Regime (85,7%) im Vergleich zu Patienten unter einem Raltegravir-basierten (73,9%), einem PI/r-basierten (77,9%) oder einem NNRTI-basierten (79,6%) Mehrtablettentregime (>2 Tabletten/Tag) [55]. Während eine komplette Nicht-Adhärenz mit 14% bis 16% in allen vier Gruppen ähnlich war, zeigten sich erhebliche Unterschiede in der selektiven Nicht-Adhärenz: Diese lag für PI/r-basierte Regime bei 7%, bei NNRTI-basierten Regimen bei 7% und bei Raltegravir-basierten Regimen mit 13% deutlich höher. Zusammengefasst zeigten sich somit Adhärenz-Probleme bei 26% der Patienten unter RAL-basierter ART, 23% der Patienten unter PI-basierter ART und 21% der

Patienten unter NNRTI-basierter ART – demgegenüber bei lediglich 14% der Patienten unter einem STR (jeweils $p < 0,0001$ versus andere Regime). Dies zeigt sich auch klinisch: Selektive Nicht-Adhärenz war assoziiert mit einem statistisch signifikant erhöhten Risiko (43-54%) für eine Hospitalisierung [55].

In einer portugiesischen Kohorte mit 1.436 HIV-Patienten (Zeitraum: 2001 – 2011) wurden folgende Regime miteinander verglichen: QD, BID, STR (FTC/TDF/EFV), FTC/TDF+EFV, Einsatz von Fixkombinationen und keine Fixkombinationen. Während bereits die vollständige Nicht-Adhärenz bei Patienten, die keine Fixkombinationen, resp. ein zweimal tägliches Regime einnahmen mit 23%, resp. 22% % am schlechtesten war, zeigte sich bereits unter der Einnahme von Fixkombinationen, resp. einem einmal täglichen Regime eine statistisch signifikante deutliche Verbesserung auf 15%, resp. 13% vollständiger Nicht-Adhärenz; diese war unter einem Single-Tablet-Regimen mit 11% am geringsten. Fasst man vollständige Nicht-Adhärenz und selektive Nicht-Adhärenz zusammen, ergeben sich klare, statistisch signifikante Vorteile für Single-Tablet-Regimen: Während die Gesamtrate von Patienten mit Adhärenzproblemen unter Einzelsubstanz-Regimen oder Regimen, bei denen Fixkombinationen eingesetzt wurden, bei 27%, resp. 18% lag, zeigten nur 11% von Patienten unter einem STR Probleme mit der Adhärenz ($p < 0,001$, resp. $p < 0,02$). Auch im Vergleich zu zweimal täglichem Mehrtablettenregimen (26%) zeigten sich statistisch signifikante Vorteile für ein Single-Tablet-Regimen ($p < 0,001$). Sogar der direkte Vergleich zwischen den substanzgleichen Formulierungen aus TDF, FTC und EFV ergab Vorteile für das STR aus FTC/TDF/EFV (Gesamt-Adhärenzrate: 89%) im Vergleich zu EFV in Kombination mit FTC/TDF (Gesamt-Adhärenzrate: 83%) ($p = 0,003$). Sowohl komplette als auch selektive Nicht-Adhärenz waren Prädiktoren für virologisches Versagen ($p < 0,001$ bzw. $p = 0,020$); eine Erhöhung der Nicht-Adhärenz um 1% resultierte in einem bis zu 4% höheren Risiko eines virologischen Versagens. Auch die Wahrscheinlichkeit für eine Hospitalisierung war bei einem Single-Tablet-Regime statistisch signifikant geringer im Vergleich zu den anderen Regimen (3,8% versus 10,3%; $p < 0,001$) [56].

Auch die Zufriedenheit der Patienten war im Falle eines Single-Tablet-Regimens am höchsten. Dabei zeigte sich, dass die Einfachheit eines Regimes – sowohl hinsichtlich der Dosisfrequenz als auch hinsichtlich der Tablettenzahl – die Patienten-Akzeptanz des Regimes, die eine Grundvoraussetzung für Adhärenz darstellt, deutlich erhöht [57].

Die Ergebnisse der oben dargestellten Analysen zeigen einen Vorteil von einmal täglichem und vor allem Single-Tablet-Regimen hinsichtlich des Therapieerfolgs und möglicher klinischer Folgen (Hospitalisierung). Dieser Vorteil resultiert insbesondere auch aus der geringeren Rate an selektiver Nicht-Adhärenz, die insgesamt zu einer verbesserten Adhärenz gegenüber Mehrtablettenregimen führt.

Insgesamt wird für EVG/COBI/FTC/TDF im Vergleich zur zweckmäßigen Vergleichstherapie RAL+FTC/TDF auf Basis der Gleichwertigkeit im indirekten Vergleich und auf Basis der patientenrelevanten Vorteile, die sich durch die Formulierung ergeben, ein **Hinweis** auf einen **geringen** Zusatznutzen abgeleitet.

Zusammenfassung - Gesamtschau des Zusatznutzen von EVG/COBI/FTC/TDF gegenüber der zweckmäßigen Vergleichstherapie, einer individuellen antiretroviralen Therapie in Abhängigkeit der Vortherapie(en) und unter Berücksichtigung des Grundes für den Therapiewechsel, insbesondere Therapieversagen aufgrund eines virologischen Versagens und etwaig einhergehender Resistenzbildung oder aufgrund von Nebenwirkungen

Die zweckmäßige Vergleichstherapie ist eine individuelle antiretrovirale Therapie in Abhängigkeit der Vortherapie(en) und unter Berücksichtigung des Grundes für den Therapiewechsel, insbesondere Therapieversagen aufgrund eines virologischen Versagens und etwaig einhergehender Resistenzbildung oder aufgrund von Nebenwirkungen. Diese drei Vergleichstherapien stellen jeweils eine Kombination eines NRTI-Backbone, bestehend aus FTC/TDF, mit einem 3. Partner aus einer der drei am häufigsten eingesetzten Substanzklassen NNRTI, PI und INI, dar. Sie repräsentieren somit einen beträchtlichen Teil der individualisierten zweckmäßigen Vergleichstherapie und decken entsprechend einen bedeutenden Teil der hier relevanten Patientenpopulation ab.

Aufgrund der Übertragbarkeit der Daten der nicht-vorbehandelten Patienten auf die vorbehandelten Patienten kann in der Gesamtschau für vorbehandelte HIV-1-Infizierte, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von Stribild assoziiert sind, für EVG/COBI/FTC/TDF ein geringer Zusatznutzen gegenüber jeder der drei gewählten (Auswahlkriterien siehe Abschnitt 4.2.1) Vergleichstherapien EFV+FTC/TDF, ATV/r+FTC/TDF und RAL+FTC/TDF abgeleitet werden.

Aufgrund der Studienanzahl, die in den Vergleich von EVG/COBI/FTC/TDF mit den ausgewählten Vertretern der zweckmäßigen Vergleichstherapie (individuelle antiretrovirale Therapie) und dem Beleg für einen geringen Zusatznutzen im Vergleich zu EFV/FTC/TDF bzw. den Hinweisen für einen geringen Zusatznutzen im Vergleich zu ATV/r+FTC/TDF und zu RAL+FTC/TDF, ist im Vergleich zur zweckmäßigen Vergleichstherapie insgesamt ein **geringer** Zusatznutzen für EVG/COBI/FTC/TDF abzuleiten; die Wahrscheinlichkeit wird als **Beleg** eingestuft.

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.

Ziel der vorliegenden Untersuchung ist die Bewertung des medizinischen Nutzens und Zusatznutzens des Single-Tablet-Regimen (STR) Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil (im Folgenden: EVG/COBI/FTC/TDF) im Vergleich zu einer individuellen antiretroviralen Therapie in Abhängigkeit der Vortherapie(en) und unter Berücksichtigung des Grundes für den Therapiewechsel, insbesondere Therapieversagen aufgrund eines virologischen Versagens und etwaig einhergehender Resistenzbildung oder aufgrund von Nebenwirkungen.

Patientenpopulation

Die gemäß Indikation zugelassene Patientenpopulation umfasst zum einen nicht-vorbehandelte, erwachsene Patienten (≥ 18 Jahre), und zum anderen erwachsene Patienten (≥ 18 Jahre), bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind.

Die Patientenpopulation der nicht-vorbehandelten Patienten (≥ 18 Jahre) umfasst auch die nicht-vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind. Der **nicht-vorbehandelte** Anteil der Gesamt-Population der Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF

assoziiert sind, ist durch die Gesamt-Population der nicht-vorbehandelten Patienten abgedeckt. Somit werden im vorliegenden Modul 4B **vorbehandelte** Patienten (≥ 18 Jahre), bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, untersucht.

Die Untersuchung der Patientenpopulation, definiert als erwachsene Patienten (≥ 18 Jahre) mit einer HIV-1 Infektion, die antiretroviral nicht-vorbehandelt sind, ist in Modul 4A zu finden.

Intervention

Die zu bewertende Intervention ist die Fixkombination EVG/COBI/FTC/TDF. EVG/COBI/FTC/TDF wird einmal täglich als Filmtablette zu einer Mahlzeit oral eingenommen und enthält 150 mg des Integrase-Inhibitors (INI) Elvitegravir, 150 mg des Boosters Cobicistat, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (nukleosidale Reverse-Transkriptase-Inhibitoren [NRTI]).

Vergleichstherapie

Für das Anwendungsgebiet „Behandlung der Infektion mit HIV-1 bei vorbehandelten Erwachsenen im Alter von 18 Jahren und darüber, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von Stribild® assoziiert sind“ fand das Beratungsgespräch zur Festlegung der zweckmäßigen Vergleichstherapie am 29.05.2013 statt.

Als zweckmäßige Vergleichstherapie wurde folgende festgelegt [58]:

Eine individuelle antiretrovirale Therapie in Abhängigkeit der Vortherapie(en) und unter Berücksichtigung des Grundes für den Therapiewechsel, insbesondere Therapieversagen aufgrund eines virologischen Versagens und etwaig einhergehender Resistenzbildung oder aufgrund von Nebenwirkungen. Die jeweilige Zulassung der Präparate ist zu beachten

Der G-BA stimmt hier den Ausführungen des pharmazeutischen Unternehmers zu, dass eine einzige zweckmäßige Vergleichstherapie nicht dem allgemein anerkannten Stand der medizinischen Evidenz entspricht.

Da es sich bei der festgelegten zweckmäßigen Vergleichstherapie um eine patientenindividuelle und entsprechend nicht substanzspezifizierte Therapie handelt, wurden nach folgenden Kriterien drei für einen großen Teil der Umstellungs-Population repräsentative Regime festgelegt:

1. EVG/COBI/FTC/TDF ist eine Dreifachtherapie, bestehend aus einem NRTI-Backbone und einer Substanz aus der Substanzklasse der 3. Partner, die als komplettes Regime nicht in Kombination mit anderen antiretroviralen Arzneimitteln eingesetzt werden darf [1]. Es ist davon auszugehen, dass die zweckmäßige Vergleichstherapie primär ebenfalls aus der Kombination eines Backbones mit einem 3. Partner besteht.

2. Substanzklassenauswahl: Ebenso wie in der Initialtherapie erhält der mit deutlichem Abstand größte Teil der Patienten auch nach Umstellung eine individualisierte Dreifachtherapie - primär bestehend aus einem NRTI-Backbone in Kombination mit einem 3. Partner aus der Substanzklasse der PI, der NNRTI oder der INI.
3. Substanzauswahl: Der mit großem Abstand am häufigsten eingesetzte Backbone ist FTC/TDF. EFV (NNRTI) und ATV/r (PI) gehören jeweils zu den am häufigsten verwendeten Vertretern ihrer Klasse; RAL ist der einzige Vertreter der INI. Diese Substanzen werden insofern stellvertretend für ihre jeweilige Substanzklasse untersucht.

Vor diesem Hintergrund wird EVG/COBI/FTC/TDF mit den folgenden, für die zweckmäßige Vergleichstherapie repräsentativen Regimen verglichen:

- Efavirenz in Kombination mit Emtricitabin und Tenofoviridisoproxil (im Folgenden: EFV/FTC/TDF),
- Ritonavir-geboostertem Atazanavir in Kombination mit Emtricitabin und Tenofoviridisoproxil (im Folgenden: ATV/r+FTC/TDF) und
- Raltegravir in Kombination mit Emtricitabin und Tenofoviridisoproxil (im Folgenden: RAL+FTC/TDF).

Diese Regime repräsentieren somit die individualisierte zweckmäßige Vergleichstherapie und decken entsprechend den mit Abstand größten Teil der hier relevanten Patientenpopulation ab.

EFV/FTC/TDF: Efavirenz (EFV), Tenofoviridisoproxil (TDF) und Emtricitabin (FTC) werden in Form des STR (EFV/FTC/TDF) einmal täglich auf leeren Magen vor dem Zubettgehen oral eingenommen; eine Filmtablette enthält 600 mg des NNRTI Efavirenz, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil.

Efavirenz (EFV), Tenofoviridisoproxil (TDF) und Emtricitabin (FTC) sind in gleicher Dosierung in unterschiedlichen Formulierungen sowohl als Einzelsubstanzen als auch als STR verfügbar. Während die Einzelsubstanzen sowohl für vorbehandelte als auch für nicht-vorbehandelte Patienten zugelassen sind, ist die Zulassung der STR-Formulierung in Europa auf vorbehandelte, virussupprimierte Patienten beschränkt. Formal entspricht die Zulassung der in den Studien gewählten Formulierung somit nicht der dort untersuchten Patientenpopulation nicht-vorbehandelter Patienten – dessen ungeachtet sind beide Studien aus folgenden Gründen im vorliegenden Bewertungsverfahren zuzulassen und zu berücksichtigen:

1. Das STR ist bioäquivalent zu den Einzelsubstanzen.
2. Die EMA erkennt das STR im Zulassungsverfahren zu EVG/COBI/FTC/TDF als Vergleichstherapie in den Zulassungsstudien an.
3. Es liegen für das STR mittlerweile klinische Studien bei nicht-vorbehandelten Patienten vor.

4. Die Kombination EFV, FTC und TDF wird in der klinischen Praxis fast ausschließlich in Form des STR eingesetzt. Des Weiteren ist Atripla[®] in den USA und der Schweiz zur Anwendung sowohl für vorbehandelte, als auch für nicht-vorbehandelte Patienten zugelassen und wird breit eingesetzt.

Die medizinisch-klinische Rationale wurde dem Gemeinsamen Bundesausschuss (G-BA) mit dem Antrag auf Beratungsanforderung vom 11. Dezember 2012 vorgelegt und im Beratungsgespräch besprochen. Der G-BA hat beschieden, dass das Verfahren der frühen Nutzenbewertung der Bewertung von Substanzen und nicht Formulierungen dient, und beide klinische Studien somit zur Bewertung heranzuziehen sind [59]: „Die Geschäftsstelle erklärt, dass der Vergleich mit der zweckmäßigen Vergleichstherapie in der Nutzenbewertung nach §35a SGB V aufgrund der jeweiligen Wirkstoffe und deren Dosierungen erfolgt, nicht aber aufgrund der Unterschiede in der Art der Anwendung, wie z. B. des *Single-Tablet-Regimen*.“

Darüber hinaus findet sich in der Nutzenbewertung des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) zu Pixantron (IQWiG-Bericht Nr. 154, A12-17) [60] folgende Aussage, die den Einschluss beider Studien zusätzlich untermauert: „Laut der Festlegung des G-BA zur zweckmäßigen Vergleichstherapie ist der deutsche Zulassungsstatus inklusive der zugelassenen Dosierungen bei den gewählten Therapien in der Vergleichsgruppe jedoch zu berücksichtigen. Die Einhaltung des Zulassungsstatus stellt eine notwendige Voraussetzung für die zweckmäßige Vergleichstherapie dar, was sich aus der Verfahrensordnung des Gemeinsamen Bundesausschusses ergibt. Allein aus dieser Voraussetzung (Nutzenbewertung grundsätzlich innerhalb des in Deutschland gültigen Zulassungsstatus) folgt jedoch nicht zwangsläufig, dass Studien, in denen Patienten außerhalb des gültigen Zulassungsstatus behandelt wurden, nicht für die Bewertung relevant sind. Für solche Studien ist jeweils zu prüfen, ob die Studienergebnisse auf eine Behandlungssituation innerhalb des deutschen Zulassungsstatus anwendbar sind, das heißt, ob die Behandlungseffekte in einer nicht zulassungsgemäßen Behandlung mit einer Behandlung innerhalb der Zulassung hinreichend vergleichbar sind. Es wäre in einem solchen Fall also durchaus denkbar, dass sich Aussagen zum Zusatznutzen eines neuen Arzneimittels im Vergleich zu einer (zulassungsgemäßen) zweckmäßigen Vergleichstherapie auf Daten außerhalb der deutschen Zulassung stützen. Damit dies möglich ist, sind jedoch nachvollziehbare, datengestützte Überlegungen erforderlich, warum die in der Studie gewählte Therapie eine für die Patienten adäquate Option darstellt und der zweckmäßigen Vergleichstherapie zumindest nicht unterlegen ist.“

ATV/r+FTC/TDF: Sowohl Atazanavir als auch Ritonavir sind einmal täglich zu einer Mahlzeit oral einzunehmen. Eine Kapsel Atazanavir enthält 300 mg Atazanavir als Sulfat; eine Tablette Ritonavir enthält 100 mg des Boosters Ritonavir. Die Fixkombination aus FTC/TDF (200 mg Emtricitabin /245 mg Tenofoviridisoproxil) sollte ebenfalls möglichst zu einer Mahlzeit eingenommen werden.

RAL+FTC/TDF: Raltegravir wird zweimal täglich in einer Dosierung von 400 mg unabhängig von einer Mahlzeit eingenommen. Die Fixkombination FTC/TDF (200 mg

Emtricitabin /245 mg Tenofoviridisoproxil) sollte möglichst zu einer Mahlzeit eingenommen werden.

Endpunkte

Die Nutzenbewertung erfolgt anhand patientenrelevanter Endpunkte auf der Basis von randomisierten kontrollierten Studien (RCT). Folgende Endpunkte werden als patientenrelevant eingestuft und in die vorliegende Nutzenbewertung eingeschlossen:

- Virologisches Ansprechen (Absenkung der Viruslast unter die Nachweisgrenze)
- Zeit bis zum Verlust des virologischen Ansprechens
- Veränderung der CD4-Zellzahl (als Einheit der CD4-Zellzahl wird Zellen/ μ L oder Zellen/ mm^3 als Äquivalent verwendet)
- Gesamtrate unerwünschter Ereignisse
- Schwerwiegende unerwünschte Ereignisse
- Therapieabbrüche aufgrund unerwünschter Ereignisse
- Unerwünschte Ereignisse Grad 3-4
- Wichtige unerwünschte Ereignisse
 - Neurologische & psychiatrische Ereignisse
 - Hautausschläge
- Unerwünschte Ereignisse von Interesse
 - Renale Ereignisse
 - Knochenfrakturen
- Häufigste unerwünschte Ereignisse
 - Erkrankungen des Gastrointestinaltrakts (SOC)
 - Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)
 - Erkrankungen des Nervensystems (SOC)
 - Psychiatrische Erkrankungen (SOC)
 - Erkrankungen der Haut und des Unterhautzellgewebes (SOC)
 - Augenerkrankungen (SOC) (nur in Zusammenhang mit ATV relevant)

Studientypen

Da randomisierte kontrollierte Studien mit der geringsten Ergebnisunsicherheit behaftet sind, liefern sie, sofern methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt, die für die Nutzenbewertung einer medizinischen Intervention zuverlässigsten Ergebnisse.

Somit fließen in die Bewertung des Nutzens ausschließlich RCT ein.

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.

Gemäß den im vorherigen Abschnitt definierten Komponenten der Fragestellung wird nach Studien gesucht, die die folgenden Kriterien erfüllen:

Tabelle 4-2: Übersicht der Kriterien für den Studieneinschluss für den Vergleich von EVG/COBI/FTC/TDF gegen EFV/FTC/TDF, ATV/r+FTC/TDF und RAL+FTC/TDF

	Einschlusskriterien	E	Ausschlusskriterien	A
Population	Erwachsene vorbehandelte Patienten mit einer HIV-1 Infektion, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind.	1	<ul style="list-style-type: none"> • Patienten mit anderen Infektionen • Kinder und Jugendliche 	1
Intervention	Einmal täglich 150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (EVG/COBI/FTC/TDF)	2	Andere Intervention oder abweichende Dosierungen	2
Vergleichs-therapie	Einmal täglich 600 mg Efavirenz, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (EFV/FTC/TDF) Einmal täglich 300 mg Atazanavir, 100 mg Ritonavir, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (ATV/r+FTC/TDF) Zweimal täglich 400 mg Raltegravir in Kombination mit einmal täglich 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (RAL+FTC/TDF)	3	Andere Vergleichstherapie oder abweichende Dosierungen	3
Endpunkte	Mindestens einer der folgenden patientenrelevanten Endpunkte wird berichtet: <ul style="list-style-type: none"> • Virologisches Ansprechen (Absenkung der Viruslast unter die Nachweisgrenze) • Zeit bis zum Verlust des virologischen Ansprechens • Veränderung der CD4-Zellzahl • Gesamtrate unerwünschter Ereignisse • Schwerwiegende unerwünschte Ereignisse • Therapieabbrüche aufgrund unerwünschter Ereignisse • Unerwünschte Ereignisse Grad 3-4 • Wichtige unerwünschte Ereignisse <ul style="list-style-type: none"> ○ Hautausschläge ○ Neurologische & psychiatrische Ereignisse • Unerwünschte Ereignisse von Interesse <ul style="list-style-type: none"> ○ Renale Ereignisse ○ Knochenfrakturen • Häufigste unerwünschte Ereignisse <ul style="list-style-type: none"> ○ Erkrankungen des Gastrointestinaltrakts (SOC) ○ Allgemeine Erkrankungen und Beschwerden am 	4	Keiner der unter den Einschlusskriterien definierten Endpunkte wird berichtet	4

	Einschlusskriterien	E	Ausschlusskriterien	A
	Verabreichungsort (SOC) <ul style="list-style-type: none"> ○ Erkrankungen des Nervensystems (SOC) ○ Psychiatrische Erkrankungen (SOC) ○ Erkrankungen der Haut und des Unterhautzellgewebes (SOC) ○ Augenerkrankungen (SOC) (nur in Zusammenhang mit ATV relevant) 			
Studientyp	RCT	5	Nicht randomisierte Studien	5
Studiendauer	≥48 Wochen	6	<48 Wochen	6
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht	7	Review-Artikel, Case Reports; Keine Vollpublikation oder Bericht verfügbar oder es werden keine zusätzlichen Informationen dargestellt	7

Begründung der Wahl der Einschlusskriterien

Population

Die Patientenpopulation ergibt sich aus dem Wortlaut der Zulassung für EVG/COBI/FTC/TDF [1]. Die gemäß Indikation zugelassene Patientenpopulation umfasst zum einen nicht-vorbehandelte, erwachsene Patienten (≥18 Jahre), und zum anderen erwachsene Patienten (≥18 Jahre), bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind. Die Untersuchung der Patientenpopulation, definiert als erwachsene Patienten (≥18 Jahre) mit einer HIV-1 Infektion, die antiretroviral nicht-vorbehandelt sind, ist in Modul 4A zu finden.

Die Patientenpopulation der nicht-vorbehandelten Patienten (≥18 Jahre) umfasst auch die nicht-vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind. Der **nicht-vorbehandelte** Anteil der Gesamt-Population der Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, ist durch die Gesamt-Population der nicht-vorbehandelten Patienten abgedeckt. Somit werden im vorliegenden Modul 4B **vorbehandelte** Patienten (≥18 Jahre), bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, untersucht.

Intervention

Das Kriterium der Intervention ergibt sich aus dem in der vorliegenden Nutzenbewertung zu bewertenden Arzneimittel EVG/COBI/FTC/TDF, welches laut Fachinformation einmal täglich oral eingenommen wird und die folgenden Substanzen enthält: 150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil.

Vergleichstherapie

Für das Anwendungsgebiet „Behandlung der Infektion mit HIV-1 bei vorbehandelten Erwachsenen im Alter von 18 Jahren und darüber, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von Stribild® assoziiert sind“ fand das Beratungsgespräch zur Festlegung der zweckmäßigen Vergleichstherapie am 29.05.2013 statt.

Entsprechend der Festlegung des G-BA stellt die zweckmäßige Vergleichstherapie eine individuelle antiretrovirale Therapie in Abhängigkeit der Vortherapie(en) und unter Berücksichtigung des Grundes für den Therapiewechsel, insbesondere Therapieversagen aufgrund eines virologischen Versagens und etwaig einhergehender Resistenzbildung oder aufgrund von Nebenwirkungen dar und ist somit patientenindividuell unterschiedlich [58].

Da die festgelegte zweckmäßige Vergleichstherapie somit patientenindividuell unterschiedlich und damit nicht substanzspezifisch ist, umfasst diese per definitionem eine Vielzahl verschiedener antiretroviraler Therapien. Aus diesem Grund werden nicht nur Daten und Ergebnisse zu einer einzigen zweckmäßigen Vergleichstherapie präsentiert, sondern zu drei ausgewählten Vergleichstherapien (siehe Abschnitt 4.2.1), die die individualisierte zweckmäßige Vergleichstherapie repräsentieren und somit den mit Abstand größten Teil der hier relevanten Patientenpopulation abdecken.

Endpunkte

Das Kriterium der Endpunkte ergibt sich aus den in der Indikation der HIV-1 Infektion etablierten und validierten patientenrelevanten Endpunkten wie in Abschnitt 4.2.1 definiert.

Studientyp

RCT sind, sofern der jeweiligen Fragestellung entsprechend und methodisch adäquat durchgeführt, mit der geringsten Ergebnisunsicherheit behaftet. Es wird demnach nach RCT gesucht.

Studiendauer

Das Kriterium der Mindeststudiendauer ergibt sich aus der Leitlinie der EMA [61], die eine Bewertung des primären Endpunkts (Abfall der Viruslast unter die Nachweisgrenze von 50 Kopien/mL) nach 48 Wochen als adäquat einstuft.

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.

Da die erste Zulassung des zu bewertenden Arzneimittels am Tag der Einreichung des vorliegenden Nutzendossier beim G-BA nicht älter als 12 Monate ist (Datum der ersten Zulassung in den USA: 27. August 2012), ist die Durchführung einer bibliografischen Literaturrecherche nicht erforderlich. Es wird somit auf die Darstellung der Methodik verzichtet.

Da zur Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF gegenüber den zweckmäßigen Vergleichstherapien EFV/FTC/TDF, ATV/r+FTC/TDF und RAL+FTC/TDF in der Population der vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, in der Liste der Studien des pharmazeutischen Unternehmers keine geeignete Studie für einen direkten Vergleich identifiziert werden kann, wird eine bibliografische Literaturrecherche für die zweckmäßige Vergleichstherapie durchgeführt. Dabei wird nach RCT gesucht, die EFV/FTC/TDF, ATV/r+FTC/TDF bzw. RAL+FTC/TDF in der Indikation der HIV-1-Infektion bei erwachsenen vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, im Vergleich zu EVG/COBI/FTC/TDF untersuchen.

Die Literaturrecherche wird in den Datenbanken MEDLINE, EMBASE und Cochrane (Cochrane Central Register of Controlled Trials) durchgeführt, wobei für jede einzelne Datenbank eine sequentielle, individuell adaptierte Suchstrategie entwickelt und separat angewandt wird. Bei allen Recherchen werden nur Studien am Menschen berücksichtigt. Studien an gesunden Teilnehmern werden nicht berücksichtigt.

Mithilfe von aktuell validierten Filtern wird eine Einschränkung auf RCT vorgenommen. Für die Suche wird die Plattform OVID verwendet.

Die Ergebnisse der systematischen bibliografischen Literaturrecherche sind in Abschnitt 4.3.1.1.2 beschrieben.

Details zur Suchstrategie sind in Anhang 4-A zu finden.

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.

Die Studienregistersuche wird in den Portalen clinicaltrials.gov, clinicalstudyresults.org sowie in dem International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO) durchgeführt. Da die Seite clinicalstudyresults.org nicht mehr verfügbar ist, werden keine Ergebnisse zu dieser Suche dargestellt.

Es wird nach abgeschlossenen, laufenden und abgebrochenen Studien mit dem zu bewertenden Arzneimittel EVG/COBI/FTC/TDF gesucht, welches im Rahmen von Phase II, III oder IV Studien mit den jeweiligen zweckmäßigen Vergleichstherapien verglichen wurde.

Details zur Suchstrategie und dem Datum der Suche sind in Anhang 4-B zu finden.

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.

Das Screening der Ergebnisse aus der Studienregistersuche wird von zwei Personen unabhängig voneinander durchgeführt. Dabei werden die Studien anhand der in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien auf Ihre Relevanz geprüft. Im Falle von Diskrepanzen hinsichtlich der Bewertung der Relevanz fand eine Diskussion zur Klärung des Ein- oder Ausschlusses statt.

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

Bewertung des Verzerrungspotenzials

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

Verzerrungsaspekte der Ergebnisse auf Studienebene:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des behandelnden Arztes
- Ergebnisgesteuerte Berichterstattung
- Sonstige Aspekte (zum Beispiel intransparenter Patientenfluss)

Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- Ergebnisgesteuerte Berichterstattung
- Sonstige Aspekte (zum Beispiel Behandlung fehlender Werte und Drop-outs)

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

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

Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Eine relevante Verzerrung liegt vor, falls sich die Aussagen und Schlussfolgerungen bezüglich der Ergebnisse bei Behebung des verzerrenden Aspektes verändern würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führt nicht zum Ausschluss aus der Bewertung. Die Klassifizierung dient vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

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

Eine zusammenfassende Bewertung der Verzerrungsaspekte nach Anhang 4-G 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.

Im vorliegenden Dossier werden die in die Bewertung einfließenden Studien anhand von Design-Charakteristika (Studiendesign, Studiendauer, Zahl der randomisierten Patienten, Ort und Zeitraum der Studiendurchführung, Interventionen, primäres Zielkriterium und relevante sekundäre Zielkriterien) beschrieben. Die Beschreibung der Design-Charakteristika deckt die im CONSORT-Statement angegebenen Items 2b bis 14 ab. Darüber hinaus wird der

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.

Patientenfluss in den Einzelstudien mithilfe der Informationen aus dem CONSORT-Flow-Chart dargestellt. Nicht randomisierte Interventionsstudien werden mittels TREND-Statement dargestellt.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

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

Für jede der in die Nutzenbewertung eingeschlossenen Studien wird die jeweils untersuchte Studienpopulation genau charakterisiert. Hierfür werden neben demografischen Charakteristika auch die den Krankheitsstatus beschreibenden Charakteristika der Studienpopulationen betrachtet.

Die demografischen Charakteristika umfassen

- Alter
- Geschlecht
- Baseline Body Mass Index (BMI)
- Erkrankungsdauer
- Rasse und
- Ethnische Abstammung.

Als krankheitsbeschreibende Faktoren werden

- Ausgangsviruslast (HIV-1-RNA Kopien) zu Screening und Baseline (kategorial)
- CD4-Zellzahl zu Baseline (stetig und kategorial)
- HIV-Krankheitsstatus

- geschätzte glomeruläre Filtrationsrate (estimated glomerular filtration rate; eGFR) und
- Adhärenz

beschrieben. Darüber hinaus werden die Disposition der Patienten und deren Gründe für einen vorzeitigen Abbruch der Studie beschrieben. Der Patientenfluss der Studien wird in einem Flow-Chart dargestellt.

Die Ergebnisse der Einzelstudien werden mittels patientenrelevanter Endpunkte dargestellt und gegenübergestellt. Als patientenrelevant werden folgende Endpunkte betrachtet und in die Nutzenbewertung eingeschlossen:

- Virologisches Ansprechen (Absenkung der Viruslast unter die Nachweisgrenze von 50 HIV-1-RNA Kopien/mL)
- Zeit bis zum Verlust des virologischen Ansprechens (Time to loss of virologic response [TLOVR])
- Veränderung der CD4-Zellzahl
- Gesamtrate unerwünschter Ereignisse
- Schwerwiegende unerwünschte Ereignisse
- Therapieabbrüche aufgrund unerwünschter Ereignisse
- Unerwünschte Ereignisse Grad 3-4
- Wichtige unerwünschte Ereignisse
 - Neurologische und psychiatrische Ereignisse
 - Hautausschläge
- Unerwünschte Ereignisse von Interesse
 - Renale Ereignisse
 - Knochenfrakturen
- Häufigste unerwünschte Ereignisse
 - Erkrankungen des Gastrointestinaltrakts (SOC)
 - Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)
 - Erkrankungen des Nervensystems (SOC)
 - Psychiatrische Erkrankungen (SOC)
 - Erkrankungen der Haut und des Unterhautzellgewebes (SOC)
 - Augenerkrankungen (SOC) (nur in Zusammenhang mit ATV relevant)

Unter den einzelnen SOC werden die jeweils am häufigsten beobachteten Ereignisse als preferred terms (PT) betrachtet.

Die Patientenrelevanz und Validität der in die Nutzenbewertung eingeschlossenen Endpunkte wird im Folgenden begründet.

Aufgrund der großen Therapieerfolge, die mit der Einführung der hochwirksamen ART einhergehen, wird die Erhebung von Endpunkten wie Morbidität und Mortalität in klinischen Studien zunehmend impraktikabel. Das Fortschreiten der Erkrankung kann durch die ART stark verlangsamt werden, wodurch eine unverhältnismäßig lange Beobachtungszeit in klinischen Studien nötig wäre. Folglich ist es heutzutage üblich, Surrogatendpunkte einzusetzen, die aufgrund ihrer früheren Messbarkeit, eine wesentlich kürzere Studiendauer erlauben, um Aussagen zu Endpunkten wie Morbidität und Mortalität zu ermöglichen.

Virologisches Ansprechen

Wie bereits durch den G-BA und das IQWiG festgestellt, ist die Bedeutung der Viruslast als Endpunkt im vorliegenden Anwendungsgebiet grundsätzlich unbestritten, und der primäre Endpunkt virologisches Ansprechen wird in der Gesamtschau der nachfolgend aufgeführten Argumente als ausreichend valider Surrogatparameter für den kombinierten Endpunkt „Aids-definierende Erkrankungen/ Tod“ eingeschätzt [62]: Als Gründe für die Einschätzung des Endpunktes Viruslast als valides Surrogat stützt sich das IQWiG auf folgende Punkte:

- Arzneimittelstudien, die auf die Reduzierung der Viruslast ausgerichtet waren, führten im Behandlungsalltag zu einer dramatischen Verbesserung der Prognose für HIV-Patienten hinsichtlich Überleben und Ausbruch der Erkrankung.
- Das biologische Modell ist plausibel.
- Die Korrelation zwischen der individuellen Veränderung der Viruslast und dem Risiko für Aids-definierende Erkrankungen/ Tod ist sehr deutlich [63].
- Auch Daten auf Studienebene (Meta-Regressionen) widersprechen dieser Assoziation nicht, auch wenn eine klare Korrelation zwischen den Effekten auf das Surrogat und den Effekten auf den patientenrelevanten Endpunkt „Aids-definierende Erkrankungen/ Tod“ kaum gegeben ist. Bei diesen Studien wird der Zusammenhang wahrscheinlich durch den Einschluss vieler Nichtunterlegenheitsstudien (nahe am Nullpunkt) verwässert. Aus diesem Grund kann das Ausmaß des Zusatznutzens für den kombinierten Endpunkt „Aids-definierende Erkrankungen/ Tod“ auf Basis des Surrogats nicht quantifiziert werden.

Dieser Einschätzung wird hier gefolgt und dementsprechend stellt der Endpunkt „Virologisches Ansprechen“ ein ausreichend valides Surrogat für den patientenrelevanten Endpunkt „Aids-definierende Erkrankungen/ Tod“ dar.

Der Cut-Off des virologischen Ansprechens von 50 Kopien/mL ist ein in dieser Indikation etablierter und validierter Cut-Off Wert, der den Empfehlungen der EMA entspricht.

Zeit bis Verlust des virologischen Ansprechens

Primäres Ziel einer ART ist nicht nur der Abfall der Viruslast unter die Nachweisgrenze, sondern auch die dauerhafte virale Suppression. Die Deutsch-Österreichischen Leitlinien

definieren das Ziel der ART wie folgt: „Ziel der antiretroviralen Therapie (ART) ist es, durch die Hemmung der HIV-Replikation infektionsbedingte Symptome zu unterdrücken, die Krankheitsprogression zu vermindern, eine Rekonstitution der zellulären Immunität zu erreichen und die chronische Immunaktivierung mit ihren resultierenden Entzündungsprozessen zu reduzieren“ [2]. Die EMA empfiehlt entsprechend die Erfassung des Erreichens und Aufrechterhaltens des virologischen Ansprechens im Rahmen klinischer Studien [61]:

„The proportion of subjects that achieves and maintains undetectable plasma HIV-RNA (i.e. < 50 copies/ml) is the preferred primary efficacy endpoint for studies in treatment naïve as well as treatment experienced populations. [...]“

Eine höhere Plasmavirämie geht mit einem rascheren Absinken der CD4-Zellzahl und einem erhöhten Progressionsrisiko einher [64]. Durch das Erreichen und Aufrechterhalten des virologischen Ansprechens, der konsekutiven viralen Suppression und der assoziierten immunologischen Rekonstitution wird das Progressionsrisiko deutlich reduziert, und das Auftreten Aids-definierender Erkrankungen oder Tod können häufig vermieden werden.

Die Zeit, über die das virologische Ansprechen und damit die virale Suppression ohne Therapiewechsel aufrechterhalten werden, beschreibt die Dauer einer erfolgreichen ART (Persistenz). Dieser Zeitraum steht für direkt erlebbare, patientenrelevante Vorteile im Sinne einer symptomfreien Zeit und ist somit ein patientenrelevanter Morbiditäts-Endpunkt. Zudem ist dieser Parameter für die lebenslang notwendige ART insofern von Bedeutung, als dass eine Therapieveränderung aufgrund von virologischem Versagen im Grunde immer den dauerhaften Verlust von Therapieoption(en) bedeutet [65]. Resistenzen schränken nicht nur die therapeutischen Optionen ein, sondern erhöhen zudem Komplexität und Kosten der nachfolgend eingesetzten Regime für den individuellen Patienten; zudem stellt die Transmission resistenter Viren ein Problem für die öffentliche Gesundheit dar [66]. Darüber hinaus steigt das Risiko eines virologischen Rebounds mit der Komplexität des Regimes [45].

Es ist von besonderer klinischer Relevanz, dass nur durch eine **dauerhafte** Reduktion der Viruslast auf <50 Kopien/mL das Risiko für Resistenzentwicklung und Therapieversagen mit konsekutivem immunologischen Therapieversagen und Krankheitsprogression **nachhaltig** verhindert werden kann.

Gestützt wird dies durch die Ergebnisse SMART-Studie:

Eine wirksame ART ist mit substantiellen Vorteilen im Hinblick auf Mortalität und Morbidität assoziiert - allerdings auch mit potentiellen Nebenwirkungen. Basierend auf der Hypothese, Nebenwirkungen durch eine reduzierte Substanzexposition minimieren zu können, wurden in SMART zwei unterschiedliche Therapiestrategien miteinander verglichen: CD4-zellgesteuerte, episodische Einnahme versus kontinuierliche Einnahme der ART. Im Ergebnis zeigte sich, dass eine episodische ART das Risiko für Nebenwirkungen nicht reduziert, und zu einem signifikanten Anstieg des Risikos für opportunistische Infektionen

und Tod führt – dies primär als Konsequenz des Abfalls der CD4-Zellzahl und dem Anstieg der Viruslast [67].

Ein langfristiges, dauerhaftes virologisches Ansprechen führt zu einer fortschreitenden immunologischen Rekonstitution, wie die über Jahre ansteigenden CD4-Zellzahlen belegen [68].

Des Weiteren belegen verschiedene Studien belegen eine Assoziation zwischen Dauer der viralen Suppression und Rebound-Rate: Die Rate virologischer Rebounds nahm über die Zeit unter ART mit viraler Suppression deutlich ab [45, 69, 70]. Zudem ist das Risiko eines virologischen Rebounds bei nicht-vorbehandelten Patienten niedriger als bei behandelten Patienten, die ihre Therapie wechseln [45].

Dauer des virologischen Ansprechens und Persistenz sind somit kritische Parameter für die Reduktion von Morbidität und Mortalität, dies bedingt durch das assoziierte immunologische Ansprechen [42], als auch Risikofaktoren für einen virologischen Rebound mit konsekutivem Therapieversagen und Krankheitsprogression.

Vor diesem Hintergrund stellt der Endpunkt, welcher die Dauer des virologischen Ansprechens abbildet, ebenfalls ein ausreichend valides Surrogat für den patientenrelevanten Endpunkt „Aids-definierende Erkrankungen/Tod“ dar.

Veränderung der CD4-Zellzahl

Die absolute CD4-Zellzahl ist das Surrogat für den immunologischen Status und der primäre Indikator für den Therapiebeginn; die Veränderung der CD4-Zellzahl unter Therapie zeigt das immunologische Ansprechen. Neben dem virologischen Ansprechen, erfasst durch die Viruslast, sollte auch immer die CD4-Zellzahl untersucht und bewertet werden. Dies wird auch vom G-BA so empfohlen [71]:

„Der Einsatz von Surrogat-Parametern, die eine zeitnahe Bewertung des Krankheitsverlaufes ermöglichen, ist daher unumgänglich. Als Surrogatparameter zur Therapie HIV-Infizierter bewährt und etabliert sind die HI-Viruslast als Maß der Virusreplikation und die CD4-Zellzahl als Maß der ("quantitativen") Funktion des Immunsystems [...].“

Demnach werden die HI-Viruslast („...als Maß der Virusreplikation...“), wie auch die CD4-Zellzahl („...als Maß der ["quantitativen"] Funktion des Immunsystems...“) vom G-BA als valide Surrogatparameter eingestuft.

Auch die EMA stuft den Endpunkt CD4-Zellzahl als valides Surrogat ein, und empfiehlt eine gleichzeitige Erfassung von HI-Viruslast und CD4-Zellzahl [61]:

„Since the introduction of Highly Active Anti-retroviral Therapy (HAART = ART in this document), viral load and CD4+ T-cell counts have been generally accepted as surrogate markers for efficacy in studies with anti-retroviral agents. [...].“

„Effects on the CD4+ T-cell count should always be documented. [...]“

Gemäß Food and Drug Administration (FDA) sollte die Erfassung der CD4-Zellzahl gemeinsam mit der HI-Viruslast, und ggf. den CDC Klasse C Ereignissen, erfolgen:

„Analyses evaluating changes over time in CD4 cell counts should accompany the analyses of HIV RNA. Clinical endpoint data (CDC class C events) should also be collected, analyzed, and submitted with the NDA. However, the frequency of such events is likely to be low particularly over a 24-week time period.“

Marschner et al. [72] untersuchten den Zusammenhang zwischen HI-Viruslast bzw. CD4-Zellzahl und Krankheitsprogression. Dabei zeigte sich, dass ein Anstieg der CD4-Zellzahl versus Baseline um 100 Zellen/ μ L bis Woche 24 mit einem 60% geringeren Risiko einer Krankheitsprogression verbunden ist (95%-KI: 51%-64%, $p < 0,001$).

In einer Meta-Analyse von 15 Studien zeigen Hill et al. [73] einen signifikanten Zusammenhang zwischen der durchschnittlichen Veränderung von CD4-Zellzahl und patientenrelevanten Endpunkten wie Aids-definierende Erkrankungen/Tod.

Die resultierende Korrelation zwischen Behandlungseffekt auf CD4-Zellzahl und klinischem Nutzen zeigt Abbildung 4: Behandlungen, die einen größeren Anstieg der CD4-Zellzahl bedingen, zeigen eine höhere Reduktion der Rate an klinischer Progression. Mit zwei Ausnahmen zeigen alle Studien einen Anstieg der CD4-Zellzahl zusammen mit einem klinischen Vorteil; das Ausmaß des klinischen Vorteils war assoziiert mit dem Anstieg der CD4-Zellzahl.

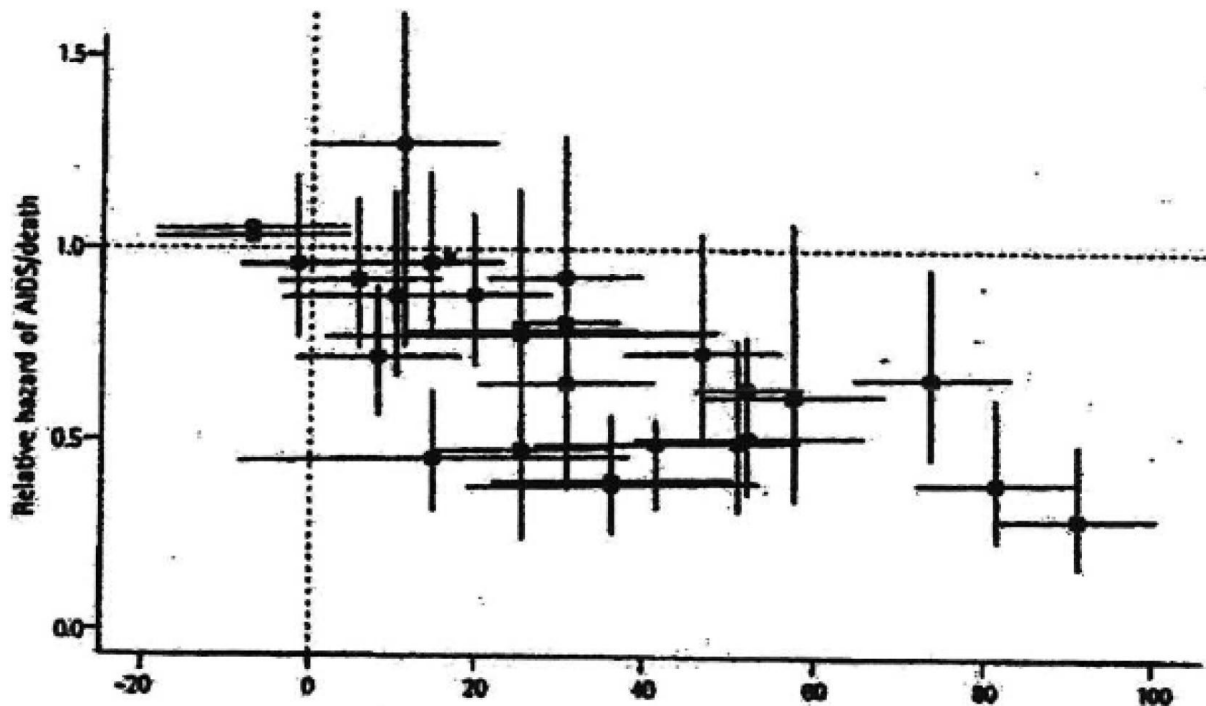


Abbildung 4: Relation zwischen Veränderung der CD4-Zellzahl und der Hazard für klinische Ereignisse (Aids/Tod) [73]

Die CD4-Zellzahl ist ein Surrogat für den immunologischen Status, das zuverlässig die Abschätzung des individuellen Risikos, an Aids zu erkranken, erlaubt und damit den primären Marker für den Beginn einer ART darstellt: Diese sollte spätestens bei einer CD4-Zellzahl von 350 CD4-Zellen/ μL begonnen werden [2]; ab einer CD4-Zellzahl von <200 CD4-Zellen/ μL muss mit zunehmender Dauer der Immunsuppression vermehrt mit dem Auftreten von Aids-definierenden Erkrankungen gerechnet werden [74]. Damit ist die Dynamik des CD4-Zellanstiegs gerade zu Beginn einer ART klinisch hochrelevant, da die Patienten bei schnellerem Anstieg der CD4-Zellzahl auch rascher oberhalb des potentiellen Risiko-Bereichs für die Entwicklung Aids-definierender Erkrankungen liegen. Die raschere Rekonstitution des Immunsystems ist somit direkt patientenrelevant.

Unerwünschte Ereignisse

Neben Wirksamkeitsparametern werden auch unerwünschte Ereignisse als Endpunkte in die Nutzenbewertung eingeschlossen. Relevante Ereignisse waren die Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4 sowie Therapieabbrüche aufgrund von unerwünschten Ereignissen.

Darüber hinaus werden unerwünschte Ereignisse, die für die Grunderkrankung oder die Arzneimittel relevant sind, im Rahmen der Nutzenbewertung betrachtet. Diese ‚wichtigen unerwünschten Ereignisse‘ und ‚unerwünschten Ereignisse von Interesse‘ umfassen:

- Wichtige unerwünschte Ereignisse:

- Neurologische und psychiatrische Ereignisse: Diese werden in Zusammenhang mit einer Behandlung mit EFV häufig beobachtet und als übliche Nebenwirkung dieser Behandlung in der Fachinformation beschrieben. Sie wurden deswegen in der vorliegenden Nutzenbewertung als relevant eingestuft. Ebenso treten bestimmte neurologische und psychiatrische Ereignisse häufig in Zusammenhang mit Raltegravir auf und werden aus diesem Grund als relevant eingeschätzt.

Obwohl diese Ereignisse keine häufig in Zusammenhang mit Atazanavir auftretenden Ereignisse sind, werden sie zum Zwecke der Vergleichbarkeit mit den anderen zweckmäßigen Vergleichstherapien dargestellt.

- Hautausschläge: Hautausschläge werden häufig in Zusammenhang mit der Substanzklasse der NNRTI und darunter auch EFV beobachtet und wurden deshalb in den klinischen Studien mit EVG/COBI/FTC/TDF monitoriert. Folglich werden diese Ereignisse hier als relevant eingestuft. Hautausschläge werden auch häufig in Zusammenhang mit Atazanavir (in EVG/COBI/FTC/TDF-Studien monitoriert) und Raltegravir beobachtet und werden auch hier deshalb als relevant eingestuft.
- Unerwünschte Ereignisse von Interesse:
 - Renale Ereignisse: Aufgrund des Zusammenhangs zwischen renalen Ereignissen und der Behandlung mit Tenofoviridisoproxil (TDF), welches in den in die Nutzenbewertung eingeschlossenen Studien zu nicht-vorbehandelten Patienten sowohl im EVG/COBI/FTC/TDF- als auch im jeweiligen Vergleichsarm gegeben wurde, wurden renale Ereignisse in den EVG/COBI/FTC/TDF-Studien monitoriert und hier als relevant eingeschätzt.
 - Knochenfrakturen: Aufgrund des Zusammenhangs zwischen Knochenfrakturen und der Behandlung mit TDF, welches in den in die Nutzenbewertung eingeschlossenen Studien zu nicht-vorbehandelten Patienten sowohl im EVG/COBI/FTC/TDF- als auch im jeweiligen Vergleichsarm gegeben wurde, wurden Knochenfrakturen in den EVG/COBI/FTC/TDF-Studien monitoriert. Sie werden hier demnach als relevant eingestuft.

Des Weiteren werden die im Rahmen der klinischen Studien am häufigsten vorkommenden unerwünschten Ereignisse (Grenze: $\geq 10\%$ in einem der Behandlungsarme) aus den folgenden SOC betrachtet:

- Erkrankungen des Gastrointestinaltrakts
- Allgemeine Erkrankungen und Beschwerden am Verabreichungsort
- Erkrankungen des Nervensystems
- Psychiatrische Erkrankungen
- Erkrankungen der Haut und des Unterhautzellgewebes und

- Augenerkrankungen (nur in Zusammenhang mit ATV relevant).

Die darunter betrachteten häufigsten unerwünschten Ereignisse werden als PT erfasst.

Sollten unter den Vergleichen zwischen EVG/COBI/FTC/TDF und ATV/r+FTC/TDF oder EVG/COBI/FTC/TDF und RAL+FTC/TDF unter einem der definierten SOC keine PT mit der Häufigkeit von $\geq 10\%$ auftreten, werden die in Modul 4A (EVG/COBI/FTC/TDF vs. EFV/FTC/TDF) unter dieser SOC aufgetretenen PT aus Vergleichbarkeitsgründen in Abschnitt 4.3.1.3.1 dennoch dargestellt, sofern es möglich ist.

Erkrankungen des Gastrointestinaltrakts (SOC), wie zum Beispiel Übelkeit und Diarrhoe, allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), wie zum Beispiel Fatigue, Erkrankungen des Nervensystems (SOC), wie zum Beispiel Schwindel und Kopfschmerzen, psychiatrische Erkrankungen (SOC), wie zum Beispiel abnorme Träume und Erkrankungen der Haut und des Unterhautzellgewebes (SOC), wie zum Beispiel Hautausschlag, und Augenerkrankungen (SOC), wie zum Beispiel skleraler Ikterus (nur in Zusammenhang mit ATV relevant), sind Ereignisse, die häufig in Zusammenhang mit EFV [9, 26], ATV [25] und RAL [18] auftreten und werden somit und aus folgenden Gründen hier als relevant eingeschätzt.

Bei der Verträglichkeit ist nicht nur der rein medizinische Aspekt von Bedeutung, sondern bei bestimmten Nebenwirkungen im gleichen Ausmaß auch der psychologische: Ein Hautausschlag beispielsweise ist ein äußerlich sichtbarer Effekt, der zu Fragen hinsichtlich der Ursachen und somit zur unbeabsichtigten Bekanntgabe einer HIV-Infektion führen kann. Dies ist besonders in der homosexuellen Community, die gut vernetzt und zum Thema HIV gut informiert ist, von hoher Relevanz. Da unerwünschte Ereignisse wie Hautausschlag im Kreis der mit HIV-1 infizierten Bevölkerung somit direkt stigmatisierend wirken, ist es von besonderem Interesse Ereignisse dieser Art weitgehend zu vermeiden. Die Stigmatisierung eines HIV-1 infizierten Patienten kann auch weitere Erkrankungen, so beispielsweise Depressionen, zur Folge haben. Daten zeigen, dass beeinträchtigende Nebenwirkungen und insbesondere auch Depressionen die Therapieadhärenz negativ beeinflussen, was wiederum das Risiko der Entwicklung von Resistenzen mit virologischem Versagen und konsekutiver Progression erhöht. Da jedoch gerade in der Behandlung der HIV-1 Infektion eine hohe Therapieadhärenz unabdingbar für einen Behandlungserfolg ist und einen direkten Einfluss auf Krankheitsprogression, Morbidität und Mortalität haben, spielt die Vermeidung von unerwünschten Ereignissen, welche die Therapieadhärenz negativ beeinflussen könnten, eine zentrale Rolle [4, 49, 50].

Abgesehen von im jeweiligen Endpunkt-Kapitel näher spezifizierten Ausnahmen wurden alle unerwünschten Ereignisse im Protokoll bzw. statistischen Analyseplan präspezifiziert. Die Erhebung der unerwünschten Ereignisse erfolgte standardisiert und nach den Regeln der ICH GCP. Die Gruppe der PT, die in den wichtigen unerwünschten Ereignissen und den unerwünschten Ereignissen von Interesse eingeschlossen wurden, wurde ebenso präspezifiziert. Multiple Ereignisse pro Patient in den jeweiligen Kategorien unerwünschter Ereignisse wurden nur einmal gezählt um Doppelungen zu vermeiden. Die häufigsten

unerwünschten Ereignisse wurden mittels MedDRA Kodierung 14.0 erfasst. In der STARTMRK-Studie wurde die MedDRA Version 11.0, bzw. später MedDRA Version 12.0 verwendet.

Die klinischen Endpunkte der unerwünschten Ereignisse reflektieren die therapieassoziierte und krankheitsbedingte Morbidität der Patienten und sind somit patientenrelevant, denn sie können mit der Notwendigkeit einer Therapieumstellung auf ein möglicherweise weniger Adhärenz-förderliches Regime oder mit einem Therapieabbruch durch den Patienten einhergehen. Beides kann zu einem erhöhten Risiko für ein virologisches Versagen führen [2, 4].

4.2.5.3 Meta-Analysen

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

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten⁴ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär anhand des Odds Ratio durchgeführt werden. In begründeten 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

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.

Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

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

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

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

Für dichotome Endpunkte wird der Therapieeffekt in der Regel anhand des Odds Ratio (OR) zusammengefasst. Gegebenenfalls, z. B. im Falle nicht hinreichend homogener Studienresultate, werden auch andere Effektmaße berechnet (wie z. B. das relative Risiko [RR]) und für Sensitivitätsanalysen verwendet (siehe Abschnitt 4.2.5.4). Für kontinuierliche Variablen werden Mittelwertdifferenzen berechnet. Für Endpunkte, für die laut Studienprotokoll die Nicht-Unterlegenheit zu einer präspezifizierten Grenze gezeigt werden sollte, werden Risikodifferenzen dargestellt.

Insbesondere für quantitative Parameter werden, falls notwendig, die Therapieeffekte zusätzlich hinsichtlich ihrer klinischen Relevanz bewertet. Zur Abschätzung der klinischen Relevanz wird neben der Betrachtung der Größe des Punktschätzers die Lage des Konfidenzintervalls in Bezug auf eine Irrelevanzgrenze herangezogen. Dieses Verfahren ist äquivalent zum Testen verschobener Nullhypothesen, wenn man verlangt, dass das gesamte Konfidenzintervall auf der richtigen Seite der Irrelevanzgrenze liegt [75]. Falls keine validierte Irrelevanzgrenze vorhanden ist, wird die standardisierte Mittelwertdifferenz (Hedges' g) und die Lage des Konfidenzintervalls bezogen auf die Irrelevanzgrenze 0,2 betrachtet [76].

Die zusammenfassende Darstellung der Resultate der Einzelstudien sowie eine visuelle Einschätzung möglicher Heterogenität der Studienergebnisse erfolgt mit Hilfe eines Forest-Plots, in dem die Therapieeffekte der Einzelstudien (sowie – im Fall hinreichender Homogenität – ein zusammenfassender Effektschätzer) mit dem entsprechenden Konfidenzintervall abgetragen werden. Falls nur eine 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 [77]. Zur Feststellung statistischer Heterogenität wird ein Signifikanzlevel von 0,2 verwendet. Für die I^2 -Statistik gelten die im Cochrane-Handbuch [78] 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. Im Fall erheblicher Heterogenität werden mögliche Ursachen der Heterogenität diskutiert und mittels Sensitivitäts- und Subgruppenanalysen untersucht.

In folgenden Fällen werden keine Meta-Analysen durchgeführt:

- falls nur eine für die Fragestellung relevante RCT zur Verfügung steht.
- 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, Endpunktoperationalisierungen, Beobachtungsdauer, etc.

In diesen Fällen werden die Resultate zu patientenrelevanten Endpunkte, wie in der/den Einzelstudie(n) analysiert, dargestellt. Im Falle gleichgerichteter Effekte und statistisch signifikanter Ergebnisse in mindestens zwei der in die Meta-Analyse eingeschlossenen Studien wurde der Gesamtschätzer für die Nutzenbewertung herangezogen, wenn die Studien ein Gesamtgewicht von mindestens 80% am Gesamtschätzer hatten.

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.

Auf Studienebene dienen Sensitivitätsanalysen zur Überprüfung der Robustheit gegenüber methodischer Variation. Dabei werden zum Beispiel die Wahl der Auswertungspopulation (Per-Protokoll-Auswertungen versus ITT-Auswertungen), der Imputationsalgorithmus (Snapshot-Auswertung versus TLOVR-Auswertung) oder der Einschluss von Patienten, die aus besonderen Gründen nicht in der primären Analyse berücksichtigt wurden, variiert. Dabei ist man auf die Verfügbarkeit solcher Analysen aus den Einzelstudien angewiesen.

Des Weiteren können Sensitivitätsanalysen zur Überprüfung der Robustheit der Ergebnisse im Rahmen von Meta-Analysen, für folgende Faktoren durchgeführt werden:

- Qualität/Verzerrungspotenzial der Einzelstudien (Welchen Effekt hatte der Ausschluss von Studien bei denen das Verzerrungspotenzial als „hoch“ eingestuft wird?)
- Modellwahl bei der Meta-Analyse (FEM statt REM)
- Wahl des Effektmaßes (standardisierte Mittelwertdifferenz statt Mittelwertdifferenz)

Sensitivitätsanalysen, die die Robustheit der Ergebnisse bei Verwendung von Modellen mit festen Effekten (FEM) für die Berechnung des Gesamtschätzers zeigen sollen, werden für alle Endpunkte durchgeführt. Die Wahl des Modells für die Meta-Analyse beeinflusst einerseits die Präzision des Gesamtschätzers und andererseits die Gewichtung der Einzelstudien in der Meta-Analyse. Die Resultate der Sensitivitätsanalysen werden immer dann präsentiert, wenn sie von denen der Hauptanalyse abweichen (z. B. signifikant versus nicht signifikant) oder wenn sie Erklärungen im Falle von Heterogenität liefern können.

In der vorliegenden Indikation der HIV-1 Infektion werden, sofern möglich, im Speziellen Sensitivitätsanalysen hinsichtlich der beiden Imputationsalgorithmen der FDA, TLOVR und Snapshot, durchgeführt [79].

Die standardisierte Mittelwertdifferenz wird zur Bewertung der klinischen Relevanz für quantitative Endpunkte herangezogen und dann dargestellt, wenn die Analyse der Mittelwertdifferenz einen statistisch signifikanten Unterschied zwischen den beiden Behandlungen zeigt.

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 Effektmodi-

fikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollten, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. -stadium

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

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

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

Es werden, sofern möglich, Subgruppenanalysen durchgeführt, um zu überprüfen, ob der Therapieeffekt über die Subgruppen hinweg konstant ist oder ob es Hinweise für eine Effektmodifikation gibt. Datenquellen für Subgruppenresultate sind Subgruppenanalysen aus Studienberichten, Publikationen sowie post-hoc Analysen. Folgende Subgruppen werden getrennt betrachtet:

- Alter (<40 Jahre versus \geq 40 Jahre),
- Geschlecht (Männer versus Frauen),
- Rasse (Kaukasisch versus nicht-kaukasisch),
- Therapieadhärenz (<95% versus \geq 95%),

sowie als Maß der Krankheitsschwere

- Ausgangsviruslast (≤ 100.000 versus > 100.000 HIV-1-RNA-Kopien/mL) und
- CD4-Zellzahl zu Baseline (≤ 350 versus > 350 Zellen/ μ L).

Rasse und Ethnie sind zwei Begriffe, die im europäischen Sprachraum fast immer synonym verwendet werden. Die aktuell vorgenommene Unterscheidung und Definition stammt aus den USA (<http://www.census.gov>) und findet insofern trotz ihrer geringen Relevanz für Europa entsprechende Anwendung in vielen US-initiierten klinischen Studien: Während Ethnie durch soziologische Faktoren wie Nationalität und Herkunft beschrieben wird, liegen dem Begriff der Rasse hingegen biologische Merkmale und physische Charakteristika wie Haut- oder Augenfarbe oder Knochenstruktur zugrunde.

Rassenabhängige (nicht jedoch Ethnie-abhängige) physiologisch bedingte Unterschiede in Pharmakokinetik und Pharmakodynamik sind auch bei HIV-relevanten Fragestellungen bekannt:

Eine HIV-assoziierte Nephropathie (HIVAN) tritt aufgrund einer genetischen Prädisposition (Mutation im Apo-L1-Gen) fast hauptsächlich bei schwarzafrikanischen Menschen auf [80]. Desgleichen zeigen sich rassenabhängige genetische Unterschiede im Cytochrom-P450-System, wodurch der Metabolismus antiretroviraler Substanzen beeinflusst werden kann: Personen schwarzafrikanischen Ursprungs weisen unterschiedliche genetische Polymorphismen im CYP450 auf, die möglicherweise zu einem verlangsamten Metabolismus von Efavirenz führen können. Des Weiteren haben Menschen schwarzafrikanischen Ursprungs unabhängig von ihrem HIV-Status ein höheres Risiko für die Entwicklung von Diabetes [81]. In einer (älteren) Studie [82] zeigten sich bei schwarzafrikanischen Menschen höhere Nebenwirkungsraten sowie eine höhere Mortalität und niedrigere Abbruchraten aufgrund von unerwünschten Ereignissen. Adjustierte man allerdings für bekannte prädiktive Faktoren wie CD4-Zellzahl und Aids in der Vorgeschichte, verschwanden die Unterschiede.

Da auf Basis der vorliegenden klinischen Evidenz im Gegensatz zur Rasse kein Einfluss der Ethnie – zumal ethnische Gruppen in sich eine ausgeprägte Heterogenität aufweisen können - auf Wirkung oder Nebenwirkung einer Medikation zu erwarten ist, wurden die Subgruppenanalysen auf die Rasse beschränkt - dies auch vor dem Hintergrund der insgesamt geringfügigen Relevanz der Fragestellung im deutschen Versorgungskontext.

Möglicherweise variierende Effekte zwischen verschiedenen Subgruppen werden mittels Interaktionstests bewertet. Zur Ableitung von Hinweisen oder Belegen auf Effektmodifikation wird primär der p-Wert herangezogen. Das I^2 -Maß wird jedoch zusätzlich dargestellt. Bedeutsame Heterogenität zwischen den Subgruppen wird ab einem $p < 0,20$ bzw. einem $I^2 > 50\%$ angenommen. Dabei besteht bei einem p-Wert von $\geq 0,05$ und $< 0,20$ ein Hinweis auf Effektmodifikation und ab $p < 0,05$ wird ein Beleg auf Effektmodifikation abgeleitet. Bei einem signifikanten Gesamtschätzer und einem statistisch signifikanten Effekt in einer Subgruppe und einem nicht signifikanten Effekt in der anderen Subgruppe führt ein Hinweis auf Effektmodifikation zu einer Herabstufung der Aussagesicherheit der Nutzensaussage für die Subgruppe mit dem nicht signifikanten Effekt.

Im Falle signifikanter Interaktionen (Signifikanzniveau 0,2) werden Meta-Analysen getrennt für die entsprechenden Subgruppen durchgeführt. Falls kein Hinweis auf eine signifikante Interaktion zwischen Behandlung und dem subgruppenbildenden Faktor besteht, werden keine nach Subgruppen getrennten Meta-Analysen durchgeführt. Meta-Regressionen werden bei geringer Studienzahl (<5) nicht durchgeführt.

Die Trennpunkte für die Subgruppen, bei denen der subgruppenbildende Faktor ein quantitatives Merkmal war, d. h. Alter, CD4-Zellzahl oder Viruslast, wurden a priori im Studienprotokoll festgelegt. Vor allem die Trennpunkte zu CD4-Zellzahl und Viruslast bilden die in dieser Indikation üblichen Subgruppenkategorien.

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 Analyseverfahren dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich⁷. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“⁸, „Multiple-Treatment-Meta-Analysen“⁹ oder auch „Netzwerk-Meta-Analysen“¹⁰, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹¹. Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

⁷ Glenny AM, Altman DG, Song F, 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.

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.

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

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

Es liegt keine direkte vergleichende Studie für den Vergleich von EVG/COBI/FTC/TDF mit RAL+FTC/TDF vor. Aus diesem Grund wird ein adjustierter indirekter Vergleich durchgeführt. Als Brückenkomparatoren werden Efavirenz bzw. Ritonavir-geboostertes Atazanavir, jeweils in Kombination mit Emtricitabin und Tenofoviridisoproxil verwendet.

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 [83].

Der einfache adjustierte indirekte Vergleich wird mittels eines frequentistischen Ansatzes durchgeführt. Er beruht auf einem einfachen Netzwerk und wird durchgeführt, wenn kein direkter Vergleich möglich ist.

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 relative Risiko (RR), das Odds Ratio (OR), sowie das Hazard Ratio (HR). Für stetige Variablen wird dieses Verfahren analog angewendet, mit der Ausnahme, dass die Effektmaße hier nicht logarithmiert werden.

Im vorliegenden Dossier zur Nutzenbewertung wird als Effektmaß primär das OR verwendet. Da jedoch v. a. bei der Durchführung von indirekten Vergleichen die Wahl des Effektmaßes eine besondere Bedeutung hat, werden hier Sensitivitätsanalysen unter Verwendung des RR als Effektmaß durchgeführt.

Grundannahme für die Gültigkeit dieser Zusammenhänge ist die Abwesenheit bedeutsamer Heterogenität zwischen den eingeschlossenen Studien (methodische Faktoren, Effektmodifikatoren). Für den Fall, dass vor Durchführung des adjustierten indirekten Vergleiches eine Meta-Analyse durchgeführt werden muss, ist die Untersuchung der Heterogenität der eingeschlossenen Studien mittels statistischer Tests sowie gebräuchlicher Heterogenitätsmaße vorgesehen. Des Weiteren werden 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 Vergleichs verzerren könnten.

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.

Es werden nur Studien des pharmazeutischen Unternehmers mit dem zu bewertenden Arzneimittel berücksichtigt, die entweder mit dem Single-Tablet-Regimen (EVG/COBI/FTC/TDF) oder aber allen vier Komponenten in Kombination (EVG+COBI+FTC+TDF) in einem Studienarm durchgeführt wurden [59]. Studien, die nur mit einzelnen, aber nicht allen vier Komponenten in Kombination durchgeführt wurden, werden hier nicht gelistet, unabhängig davon, ob sie mit der Marketing Authorisation eingereicht wurden. Zudem werden in der folgenden Tabelle nur die Studien im entsprechenden Anwendungsgebiet aufgenommen (Infektion mit dem Humanen Immundefizienzvirus 1 (HIV-1) bei vorbehandelten Erwachsenen im Alter von 18 Jahren und darüber, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind).

Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
GS-US-236-0115	nein	laufend	48 Wochen (ggf. Erweiterung um	Switch zu Elvitegravir 150 mg/ Cobicistat

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
			48 Wochen für Patienten in deren Land EVG/COBI/FTC /TDF nach 48 Wochen Behandlung noch nicht kommerziell verfügbar ist)	150 mg/ Emtricitabin 200 mg/ Tenofovir- disoproxil 245 mg Weitere Einnahme von Ritonavir-geboostertem Proteaseinhibitor + Emtricitabin 200 mg/ Tenofovirdisoproxil 245 mg
GS-US-236-0121	nein	laufend	48 Wochen (ggf. Erweiterung um 48 Wochen für Patienten in deren Land EVG/COBI/FTC /TDF nach 48 Wochen Behandlung noch nicht kommerziell verfügbar ist)	Switch zu Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofovir- disoproxil 245 mg Weitere Einnahme von NNRTI + Emtricitabin 200 mg/ Tenofovir- disoproxil 245 mg (Efavirenz, Nevirapin, und Rilpivirin sind die einzigen erlaubten NNRTI in Kombination mit FTC/TDF)
GS-US-292-0109	nein	laufend	48 Wochen	Switch zu Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofovir- alafenamid 10 mg EVG/COBI/FTC/TAF Weitere Einnahme von Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofovirdisoproxil 245 mg oder Efavirenz 600 mg/ Emtricitabin 200 mg/ Tenofovirdisoproxil 245 mg oder Ritonavir 100 mg + Atazanavir 300 mg + Emtricitabin 200 mg/ Tenofovirdisoproxil 245 mg oder Cobicistat 150 mg +

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
				Atazanavir 300 mg + Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg

Geben Sie an, welchen Stand die Information in Tabelle 4-3 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 Stand der Tabelle 4-3 ist der 30. April 2013.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-3 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

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

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
GS-US-236-0115	Zu dieser Studie wurden noch keine Daten berichtet.
GS-US-236-0121	Zu dieser Studie wurden noch keine Daten berichtet.
GS-US-292-0109	Zu dieser Studie wurden noch keine Daten berichtet. Es handelt sich um eine Studie zu dem aus TDF weiterentwickelten Wirkstoff Tenofovirafenamid (TAF).

Zum Zeitpunkt der Dossiereinreichung stehen noch keine Daten aus den Studien GS-US-236-0121 und GS-US-236-0115 zur Verfügung. Im Folgenden wird somit nur das Design der beiden Switch-Studien für EVG/COBI/FTC/TDF beschrieben.

GS-US-236-0115

Studie GS-US-236-0115 ist eine randomisierte aktiv-kontrollierte Studie, die in einem offenen Parallelgruppendesign durchgeführt wurde. Die untersuchte Population besteht aus HIV-1-infizierten erwachsenen Patienten, die unter einer Therapie mit geboostertem Protease-Inhibitor (PI/r) in Kombination mit FTC/TDF seit mehr als sechs Monaten virologisch supprimiert sind. Die Patienten wurden in einem Verhältnis von 2:1 entweder zu einem Wechsel des Behandlungsregimes zu EVG/COBI/FTC/TDF (n=280) oder zur weiteren Einnahme des bestehenden Behandlungsregimes PI/r+FTC/TDF (n=140) randomisiert.

Nach einer maximal 30-tägigen Screening-Phase wurden die Patienten für 48 Wochen behandelt. Patienten, in deren Ländern EVG/COBI/FTC/TDF nach dieser Behandlungsphase

kommerziell verfügbar war, traten in eine 30-tägige Follow-Up-Phase (vor Ort oder per Telefon) ein. Patienten, in deren Ländern EVG/COBI/FTC/TDF nach den ersten 48 Wochen Behandlung noch nicht kommerziell verfügbar war, hatten die Möglichkeit in eine 30-tägige Follow-Up-Phase (vor Ort oder per Telefon) einzutreten, oder EVG/COBI/FTC/TDF für weitere 48 Wochen oder bis zum Stopp des klinischen Programms zu EVG/COBI/FTC/TDF durch Gilead, zu erhalten. Dies gilt auch für Patienten, die zur weiteren Einnahme des bestehenden Behandlungsregimes randomisiert wurden. Es konnten nur Patienten in die Extensionsphase eintreten, die kein virologisches Versagen aufwiesen.

Die Patienten mussten seit mindestens sechs Monaten vor Screening unter einer stabilen Initial- oder Zweitlinientherapie bestehend aus PI/r+FTC/TDF virologisch supprimiert sein (Viruslastgrenze gemäß lokal verwendetem Assay, mindestens zweimal gemessen). Im Falle des zweiten Regimes durfte die HIV-1 RNA Viruslast zum Zeitpunkt des Behandlungswechsels vom initialen auf das zweite Regime nicht über der Nachweisgrenze gewesen sein (gemäß lokal verwendetem Assay) noch durfte der Patient nach einem bestätigten virologischen Ansprechen je zwei aufeinanderfolgende HIV-1-RNA Messungen über der Nachweisgrenze gehabt haben (Viruslastgrenze gemäß lokal verwendetem Assay). Im Original-Protokoll war eine eGFR ≥ 90 mL/min vorgeschrieben, dies wurde jedoch in einem Amendment zu ≥ 70 mL/min geändert. Von der Studienteilnahme ausgeschlossen wurden unter anderem Patienten, die jemals einen INI eingenommen haben, die innerhalb von 30 Tagen vor dem Screening eine neue Aids-definierende Erkrankung hatten, schwanger waren oder stillten oder die eine Hepatitis C-Behandlung bekamen oder bekommen sollten.

Die Patienten wurden per IVRS/IWRS auf die Behandlungsarme randomisiert und erhielten entweder 150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (EVG/COBI/FTC/TDF) oder nahmen weiter einen Ritonavir-geboosterten PI in Kombination mit 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (PI/r+FTC/TDF) ein.

Primärer Endpunkt ist der Anteil an Patienten mit HIV-1 RNA < 50 Kopien/mL zu Woche 48, definiert durch den FDA Snapshot-Algorithmus (Nichtunterlegenheitshypothese zur 12%-Grenze). Sekundäre Endpunkte sind die Sicherheit und Verträglichkeit der beiden Behandlungen zu Woche 48 und die Veränderung der CD4-Zellzahl zu Woche 48. Es werden außerdem patientenberichtete Outcomes erhoben, wie zum Beispiel der HIV Treatment Satisfaction Questionnaire Status (HIVTSQs) und der Short Form-36 (SF-36).

Die Studie startete im November 2011 und wird voraussichtlich bis Dezember 2013 (Woche 48) fertiggestellt. Erste Ergebnisse werden Anfang des Jahres 2014 erwartet.

GS-US-236-0121

Studie GS-US-236-0121 ist eine randomisierte aktiv-kontrollierte Studie, die in einem offenen Parallelgruppendesign durchgeführt wurde. Die untersuchte Population besteht aus HIV-1-infizierten erwachsenen Patienten, die unter einer Therapie mit NNRTI in Kombination mit FTC/TDF seit mehr als sechs Monaten virologisch supprimiert sind. Die Patienten wurden in einem Verhältnis von 2:1 entweder zu einem Wechsel des

Behandlungsregimes zu EVG/COBI/FTC/TDF (n=280) oder zur weiteren Einnahme des bestehenden Behandlungsregimes, NNRTI+FTC/TDF (n=140), randomisiert.

Nach einer maximal 30-tägigen Screening-Phase wurden die Patienten für 48 Wochen behandelt. Patienten, in deren Ländern EVG/COBI/FTC/TDF nach dieser Behandlungsphase kommerziell verfügbar war, traten in eine 30-tägige Follow-Up Phase (Vor Ort oder per Telefon) ein. Patienten, in deren Ländern EVG/COBI/FTC/TDF nach den ersten 48 Wochen Behandlung noch nicht kommerziell verfügbar war, hatten die Möglichkeit in eine 30-tägige Follow-Up-Phase (vor Ort oder per Telefon) einzutreten oder EVG/COBI/FTC/TDF für weitere 48 Wochen oder bis zum Stopp des klinischen Programms zu EVG/COBI/FTC/TDF durch Gilead, zu erhalten. Dies gilt auch für Patienten, die zur weiteren Einnahme des bestehenden Behandlungsregimes randomisiert wurden. Es konnten nur Patienten in die Extensionsphase eintreten, die kein virologisches Versagen aufwiesen.

Die Patienten mussten seit mindestens sechs Monaten vor Screening unter einer stabilen Initialtherapie bestehend aus NNRTI+FTC/TDF virologisch supprimiert (Viruslastgrenze gemäß lokal verwendetem Assay, mindestens zweimal gemessen) sein. Dies schließt auch Patienten ein, die von einer Behandlung mit Einzelkomponenten zu einer Behandlung mit einer fixen Kombination derselben Wirkstoffe wechselten. Im Original-Protokoll war eine eGFR ≥ 90 mL/min vorgeschrieben, dies wurde jedoch in einem Amendment zu ≥ 70 mL/min geändert. Ebenfalls in einem Amendment geändert wurde das Kriterium der Erstbehandlung. Neues Einschlusskriterium war, dass die Patienten auf ihrer Initialbehandlung oder ihrer zweiten Behandlung virologisch supprimiert sein mussten. Von der Studienteilnahme ausgeschlossen wurden unter anderem Patienten, die jemals einen INI eingenommen haben, die innerhalb von 30 Tagen vor dem Screening eine neue Aids-definierende Erkrankung hatten, schwanger oder stillend waren oder die eine Hepatitis C-Behandlung bekamen oder bekommen sollten.

Die Patienten wurden per IVRS/IWRS auf die Behandlungsarme randomisiert und erhielten entweder 150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil oder nahmen weiter NNRTI in Kombination mit 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil ein. Bei der Randomisierung wurde nach Einnahme von Efavirenz (ja oder nein) stratifiziert.

Primärer Endpunkt ist der Anteil an Patienten mit HIV-1 RNA < 50 Kopien/mL zu Woche 48, definiert durch den FDA Snapshot-Algorithmus (Nichtunterlegenheitshypothese zur 12%-Grenze). Sekundäre Endpunkte sind die Sicherheit und Verträglichkeit der beiden Behandlungen zu Woche 48 und die Veränderung der CD4-Zellzahl zu Woche 48. Es werden außerdem patientenberichtete Outcomes erhoben, wie zum Beispiel der HIVTSQs und der SF-36.

Die Studie startete im November 2011 und wird voraussichtlich bis Dezember 2013 (Woche 48) fertiggestellt. Erste Ergebnisse werden Anfang des Jahres 2014 erwartet.

Es konnte in der Liste der Studien des pharmazeutischen Unternehmers keine Studie identifiziert werden, die sich zur Beantwortung der Fragestellung von EVG/COBI/FTC/TDF bei vorbehandelten Patienten mit HIV-1, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, im Vergleich zur jeweiligen zweckmäßigen Vergleichstherapie geeignet hätte.

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

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

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

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

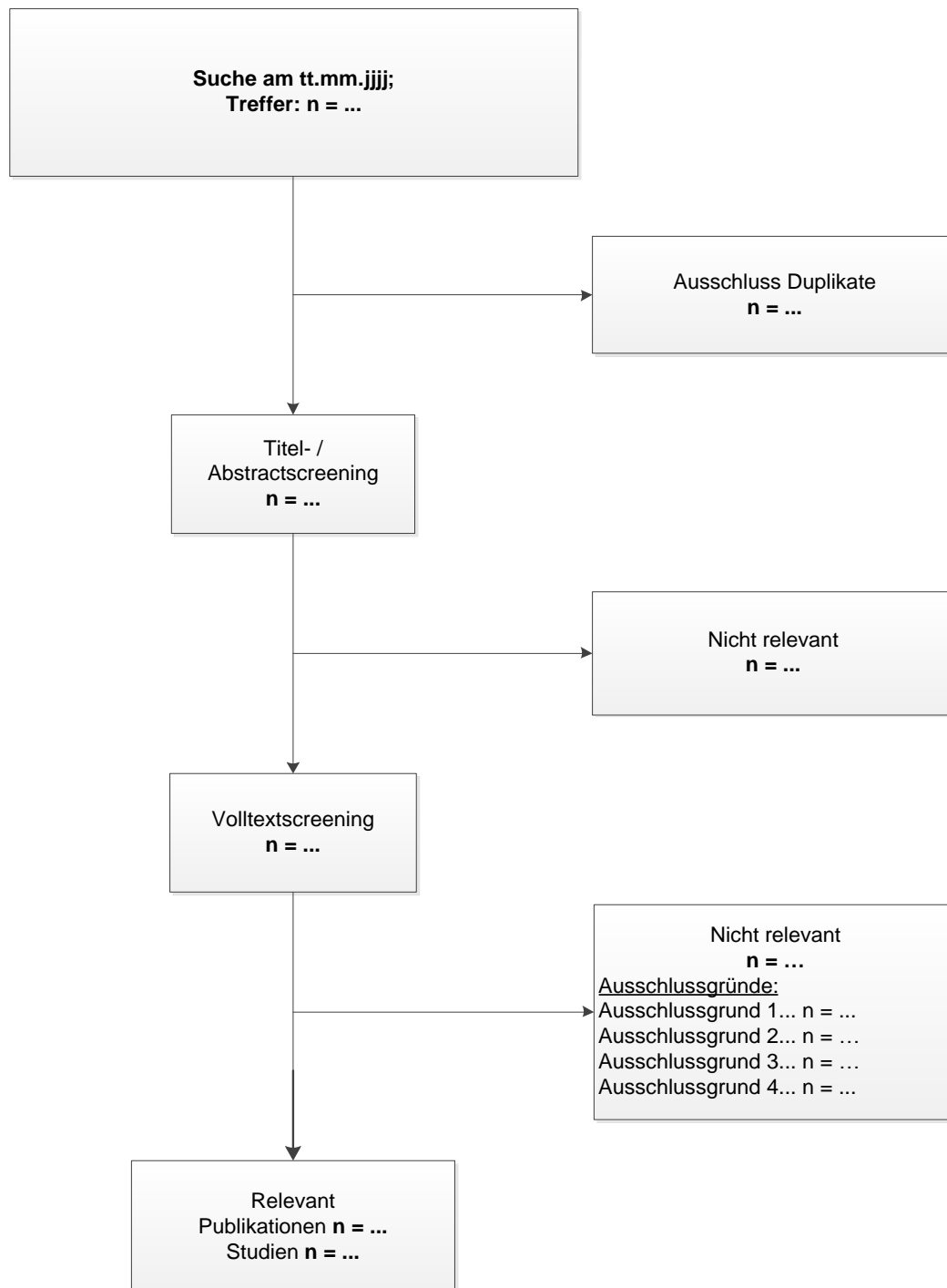


Abbildung 5: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Wie bereits in Abschnitt 4.2.3.2 beschrieben, wurde keine bibliographische Literaturrecherche für das zu bewertende Arzneimittel durchgeführt.

Da in der Liste der Studien des pharmazeutischen Unternehmers keine geeignete Studie für einen direkten oder indirekten Vergleich identifiziert werden kann, wird eine bibliografische Literaturrecherche durchgeführt:

Wie in Abschnitt 4.2.1 dargestellt, wurde die Zulassung für die Patientenpopulation der vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, von der Europäischen Zulassungsbehörde auf Basis der Daten zu den nicht-vorbehandelten Patienten erteilt. Das klinische Entwicklungsprogramm zur Umstellung vorbehandelter Patienten läuft derzeit (siehe Tabelle 4-3) – es liegen jedoch noch keine Daten aus randomisierten klinischen Studien vor (siehe Tabelle 4-4).

Vor diesem Hintergrund und basierend auf den in Abschnitt 4.2.1 ausgeführten Erwägungen wurde die Literaturrecherche entsprechend für die drei repräsentativen, substanzspezifischen zweckmäßigen Vergleichstherapien bestehend aus Efavirenz (EFV), Atazanavir/r (ATV/r) und Raltegravir (RAL), jeweils in Kombination mit FTC/TDF, durchgeführt.

Es wird dabei nach RCT gesucht, die EFV/FTC/TDF, ATV/r+FTC/TDF bzw. RAL+FTC/TDF in der Indikation der HIV-1-Infektion bei erwachsenen Patienten im direkten Vergleich zu EVG/COBI/FTC/TDF untersuchen.

Die umfassende bibliographische Literaturrecherche vom 08.04.2013 nach RCT mit der zweckmäßigen Vergleichstherapie Efavirenz erzielte insgesamt 757 Treffer (ohne Duplikate), die zunächst auf Grundlage von Titel und Abstract nach den in Abschnitt 4.2.2 definierten Kriterien gescreent wurden. Dabei wurden alle 757 Publikationen als nicht relevant eingestuft und ausgeschlossen.

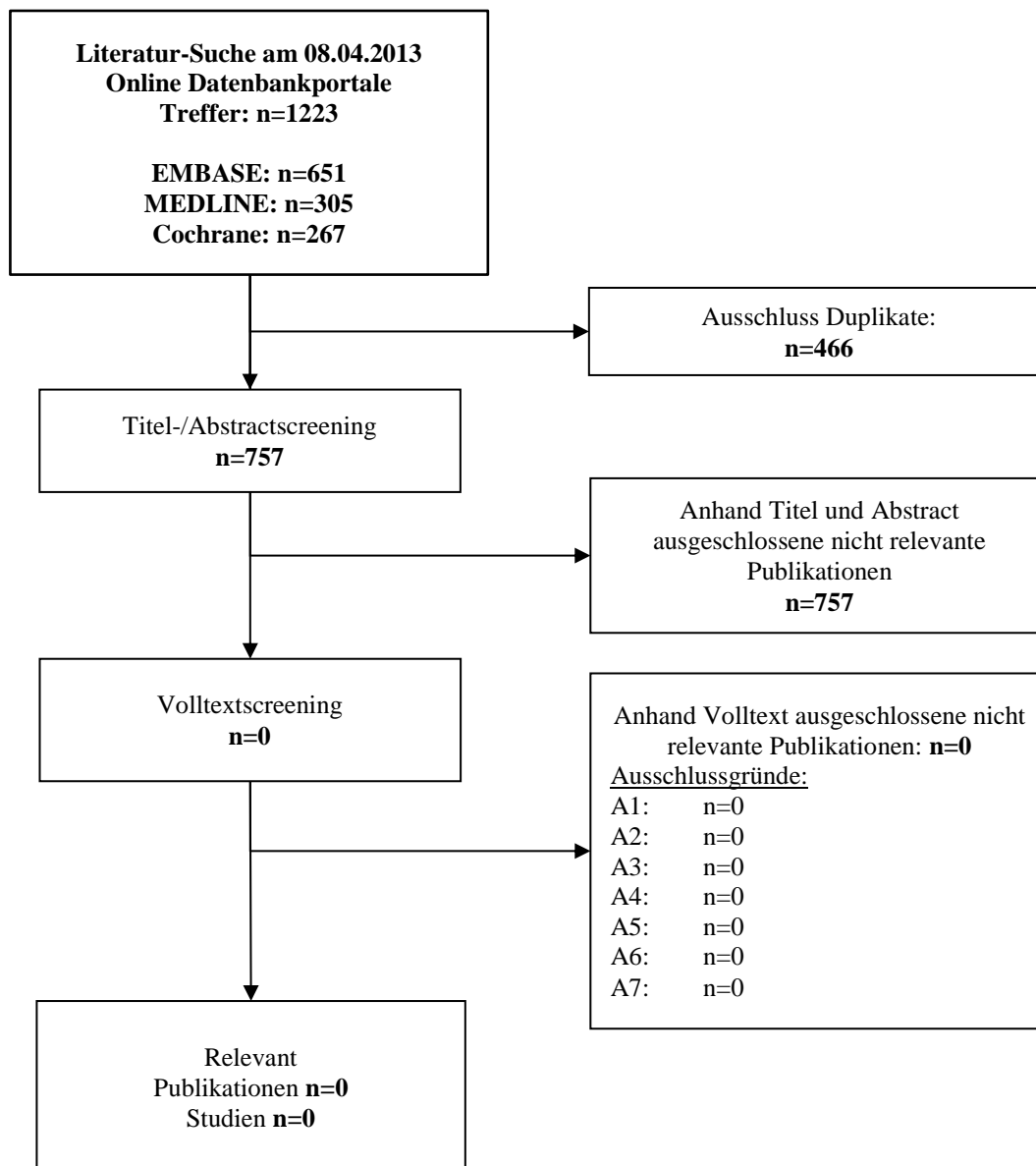


Abbildung 6: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie EFV/FTC/TDF im Vergleich mit dem zu bewertenden Arzneimittel

Die umfassende bibliographische Literaturrecherche vom 08.04.2013 nach RCT mit der zweckmäßigen Vergleichstherapie Atazanavir erzielte insgesamt 267 Treffer (ohne Duplikate), die zunächst auf Grundlage von Titel und Abstract nach den in Abschnitt 4.2.2 definierten Kriterien gescreent wurden. Dabei wurden alle 267 Publikationen als nicht relevant eingestuft und ausgeschlossen.

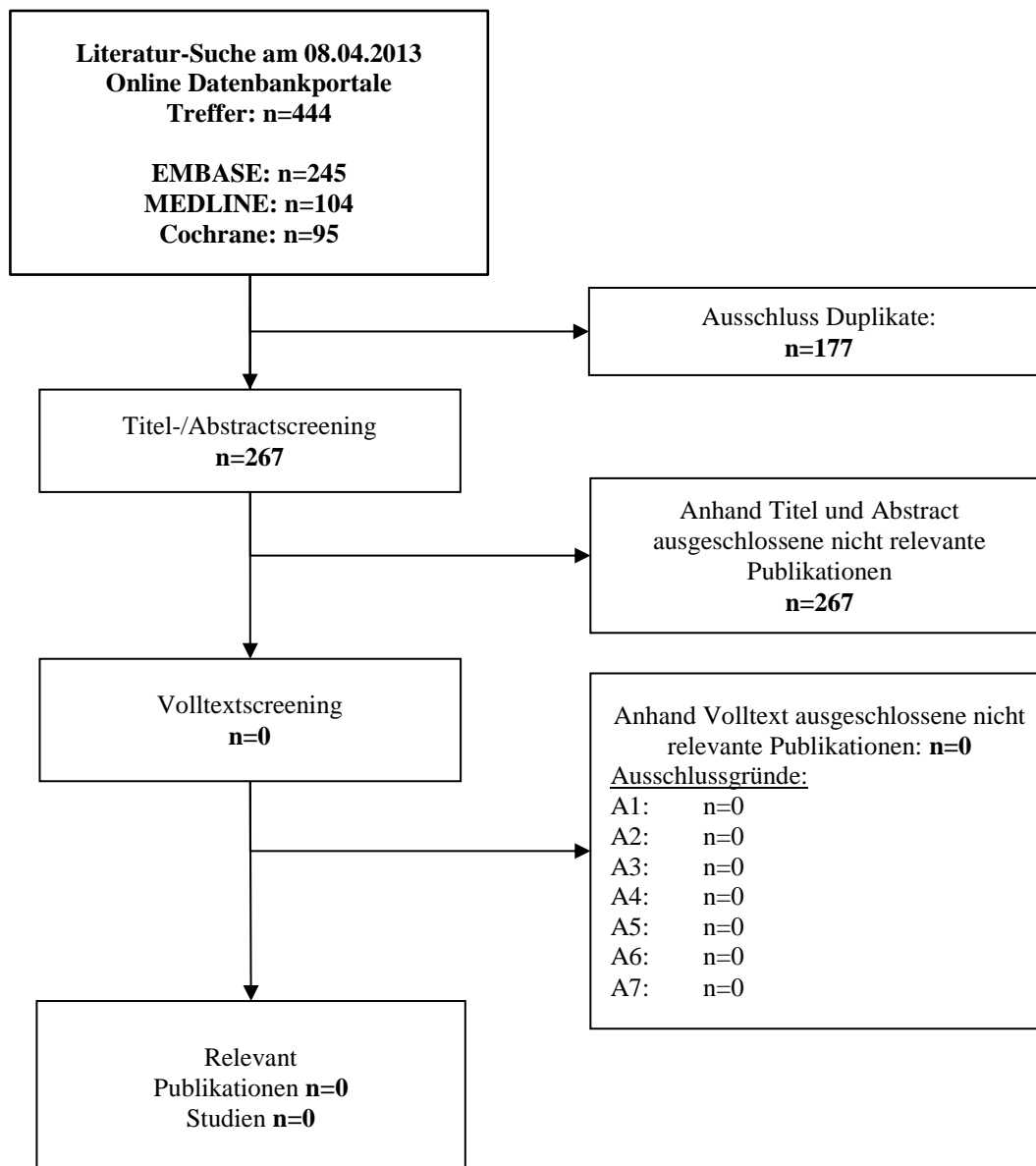


Abbildung 7: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie ATV/r+FTC/TDF im Vergleich mit dem zu bewertenden Arzneimittel

Die umfassende bibliographische Literaturrecherche vom 08.04.2013 nach RCT mit der zweckmäßigen Vergleichstherapie Raltegravir erzielte insgesamt 124 Treffer (ohne Duplikate), die zunächst auf Grundlage von Titel und Abstract nach den in Abschnitt 4.2.2 definierten Kriterien gescreent wurden. Dabei wurden alle 124 Publikationen als nicht relevant eingestuft und ausgeschlossen.

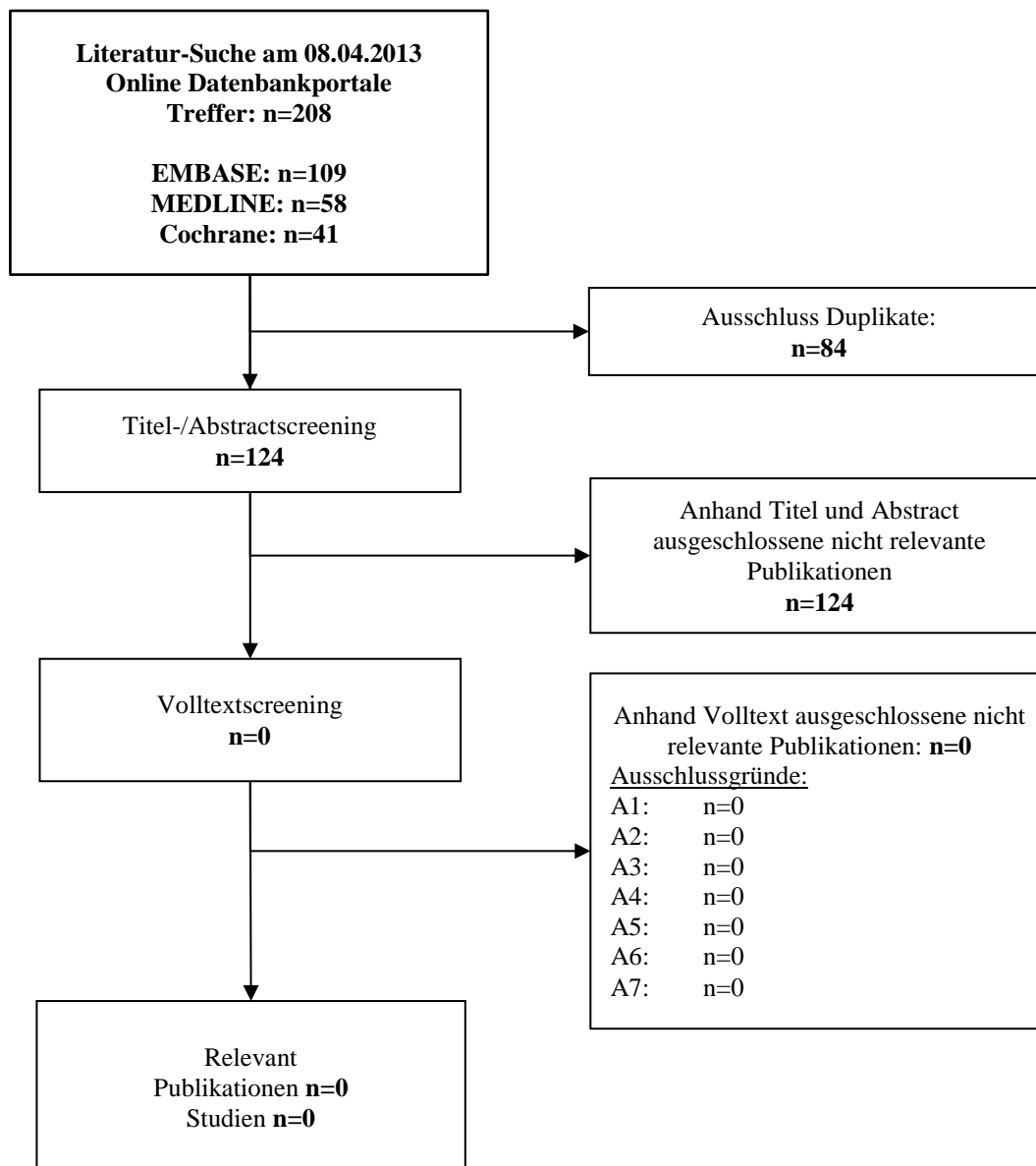


Abbildung 8: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie RAL+FTC/TDF im Vergleich mit dem zu bewertenden Arzneimittel

Es konnten keine Studien identifiziert werden, die sich für einen direkten Vergleich geeignet hätten.

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 4-3) 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 4-5: 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)
<Studie 1>	clinicaltrials.gov [Zitate], clinicalstudyresults.org [Zitate]	ja	nein

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

Es konnte in der Studienregistersuche keine Studie identifiziert werden, die sich zur Beantwortung der Fragestellung von EVG/COBI/FTC/TDF bei vorbehandelten Patienten mit HIV-1, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, im Vergleich zur jeweiligen zweckmäßigen Vergleichstherapie geeignet hätte.

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

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

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. Orientieren Sie sich dabei an der beispiel-

haften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-6: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						

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.

Die vom G-BA festgelegte zweckmäßige Vergleichstherapie ist eine individuelle antiretrovirale Therapie in Abhängigkeit der Vortherapie(en) und unter Berücksichtigung des Grundes für den Therapiewechsel, insbesondere Therapieversagen aufgrund eines virologischen Versagens und etwaig einhergehender Resistenzbildung oder aufgrund von Nebenwirkungen – die zweckmäßige Vergleichstherapie ist demzufolge eine individualisierte und nicht substanzspezifizierte Therapie. Zur Konkretisierung der Fragestellung wurden drei substanzspezifische Regime ausgewählt, die für die individualisierte Therapie als repräsentativ erachtet wurden.

Es konnte durch die Informationsbeschaffung keine Studie identifiziert werden, die sich zur Beantwortung der Fragestellung, EVG/COBI/FTC/TDF im Vergleich zur festgelegten zweckmäßigen Vergleichstherapie - repräsentiert durch EFV/FTC/TDF, ATV/r+FTC/TDF bzw. RAL+FTC/TDF -, bei vorbehandelten erwachsenen Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, geeignet hätte.

Allerdings können die Daten aus den klinischen Studien zu nicht-vorbehandelten Patienten auf Basis der folgenden medizinischen Rationale auf die vorbehandelte Patientenpopulation, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, übertragen werden:

Bei den Therapieprinzipien der ART gilt es als generelles Ziel der Therapie, durch die Hemmung der HIV-Replikation infektionsbedingte Symptome zu unterdrücken, Krankheitsprogression zu vermindern, eine immunologische Rekonstitution zu erreichen und die chronische Immunaktivierung zu reduzieren. Nur eine hohe virologische Wirksamkeit der ART und eine dauerhafte Suppression der Viruslast auf unter 50 RNA-Kopien/mL können Resistenzentwicklung und konsekutives Therapieversagen nachhaltig verhindern [2].

Grundsätzlich unterscheidet man in der ART zwei Patientenpopulationen, für die die übergeordneten Therapieprinzipien (dauerhafte Hemmung der Replikation und virale Suppression auf <50 Kopien/mL) gleichermaßen gültig sind:

1. Patienten, die noch keine antiretrovirale Therapie erhalten (nicht-vorbehandelte Patienten):
2. Patienten, die antiretroviral behandelt werden (vorbehandelte Patienten)

In den Empfehlungen zur Umstellung antiretroviral vorbehandelter Patienten wird in allen nationalen und internationalen Leitlinien zwischen zwei Subpopulationen, bzw. Indikationen zur Umstellung unterschieden [2-6]:

- Supprimierte Patienten: Umstellung aufgrund von Nebenwirkungen oder zur Therapievereinfachung und Unterstützung, bzw. Verbesserung der Adhärenz
- Patienten von mit virologischem Versagen: Umstellung aufgrund des virologischen Versagens und konsekutiver Virämie

Diese Unterscheidung ergibt sich aus dem entscheidenden Kriterium, das die Grundvoraussetzung für die virologische Wirksamkeit einer ART darstellt: Dem Vorliegen Resistenz-assoziiierter Mutationen gegen die einzusetzenden Substanzen.

Die Unterscheidung zwischen Patienten mit und ohne Resistenz-assoziierten Mutationen ist von elementarer Bedeutung: Ein vorbehandelter Patient, dessen Virus keine Resistenz-assoziierten Mutationen gegen antiretrovirale Substanzen aufweist, entspricht hinsichtlich des Resistenz-Status einem nicht-vorbehandelten Patienten. Damit ist dieser Patient im Hinblick auf die entsprechenden antiretroviralen Substanzen als funktionell nicht-vorbehandelt zu definieren - womit beide Populationen (nicht-vorbehandelt und vorbehandelt ohne Resistenzen) als gleichgestellt zu betrachten sind. Vor dem Hintergrund der Gleichstellung der Population können klinische Daten zu nicht-vorbehandelten Patienten somit auf vorbehandelte Patienten ohne Resistenzen extrapoliert und damit übertragen werden [7].

Die Rationale der Übertragbarkeit klinischer Daten von nicht-vorbehandelten Patienten auf vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, lässt sich zudem wie folgt belegen:

1. Die DHHS-Leitlinien führen in den Empfehlungen zur Umstellung supprimierter Patienten explizit aus: „Für Regime, die zur Therapievereinfachung eingesetzt werden, sollte grundsätzlich eine hohe Wirksamkeit bei antiretroviral nicht-vorbehandelten Patienten nachgewiesen worden sein, oder es sollte für den Patienten anhand seiner individuellen Vorbehandlung und seines Resistenzprofils eine hohe Wirksamkeit vorhergesagt werden können“ [4]. Gemäß der DHHS-Leitlinie ist somit davon auszugehen, dass die nachgewiesene Wirksamkeit eines antiretroviralen Regimes bei nicht-vorbehandelten Patienten auf suppressierte vorbehandelte Patienten extrapoliert und damit übertragen werden kann.
2. Die Übertragbarkeit der Daten ist hinsichtlich der Wirksamkeit und des Verträglichkeitsprofils auch rein mechanistisch und zwar sowohl pharmakologisch als auch biologisch nachvollziehbar:

Eine Initialtherapie muss ausreichend virologisch wirksam sein, um das Virus erfolgreich zu kontrollieren und die Ausgangsviruslast unabhängig von ihrer Höhe (etwa 50% der Patienten weisen eine Ausgangsviruslast von >100.000 Kopien/mL auf) unter die Nachweisgrenze zu senken. Antiretrovirale Therapieregime, die zur Umstellung supprimierter Patienten eingesetzt werden, hingegen müssen lediglich ausreichend virologisch wirksam sein, um die Viruslast unter der Nachweisgrenze zu halten. Entsprechend muss die virologische Potenz einer antiretroviralen Initialtherapie mindestens so hoch sein wie die eines Regimes, auf das ein suppressierter Patient umgestellt wird. Vor diesem Hintergrund ist davon auszugehen, dass klinische Daten zur virologischen Wirksamkeit, die bei nicht-vorbehandelten Patienten erhoben werden, direkt auf die Umstellung vorbehandelter Patienten extrapolierbar sind, sofern diese keine Resistenzen gegen die Substanzen aufweisen.

Unter Therapie auftretende Nebenwirkungen sind primär Substanz- und nicht Virus-assoziiert, woraus schlussgefolgert werden kann, dass die Höhe der Viruslast zu Beginn einer bestimmten Therapie keinen Einfluss auf die auftretenden Nebenwirkungen hat. Entsprechend wird auch im Bericht zum Sicherheitsprofil in den Fachinformationen [8-35] zu den verschiedenen antiretroviralen Substanzen nicht zwischen vorbehandelten und nicht-vorbehandelten Patienten unterschieden – ganz im Gegenteil beruht das in den Fachinformationen berichtete Sicherheitsprofil stets auf Basis des gesamten klinischen Entwicklungsprogramms.

3. Die EMA hat für Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil sowohl die Zulassung für nicht-vorbehandelte Patienten, als auch die Zulassung für die Patienten erteilt, die keine Resistenzen gegen einen der drei enthaltenen antiretroviralen Substanzen aufweisen. Da diese Zulassung ausschließlich auf Basis von klinischen Daten zu nicht-vorbehandelten Patienten eingereicht wurde, muss

geschlussfolgert werden, dass die EMA die beiden Population auf Basis der dargestellten wissenschaftlichen Rationale funktionell gleichgestellt und die Daten entsprechend extrapoliert hat. Dies wird auch durch ein von der EMA erstelltes Konzeptpapier zur Extrapolation von klinischen Daten klar unterstützt: Die angestrebte Extrapolation ist definiert als Deduktion von Studienergebnissen in bestimmten Subpopulation (Quellenpopulation) auf andere Subpopulationen (Zielpopulation) – analog gilt dies für verwandte Erkrankungen oder Substanzen. Zielsetzung der Extrapolation von Daten ist zum einen die Reduktion des Bedarfs der zusätzlichen Datengenerierung für die Zielpopulation, und zum anderen die Vermeidung unnötiger Studien und Allokation von Ressourcen in die klinischen Studien, die tatsächlich notwendig sind. Grundvoraussetzung für die Übertragbarkeit ist eine pharmakologische und biologische Rationale, die für Quellen- und Zielpopulation anhand von Erkrankung, Medikation (Wirkmechanismus, Pharmakokinetik und Pharmakodynamik) sowie Anwendbarkeit von Wirksamkeits- und Sicherheitsendpunkten bewertet werden sollte [7]. Für die beiden hier vorliegenden Populationen der nicht-vorbehandelten Quellenpopulation und der Zielpopulation von Patienten ohne Resistenz-assoziierte Mutationen, zeigt sich die geforderte pharmakologische und biologische Rationale deutlich bei der vollständigen Übereinstimmung der genannten Parameter. Vor diesem Hintergrund ist mit Erteilung der Zulassung also davon auszugehen, dass die Europäische Zulassungsbehörde die Übertragbarkeit der Daten von den nicht-vorbehandelten Patienten auf die vorbehandelten Patienten ohne Resistenz-assoziierte Mutationen als gegeben ansieht.

In logischer Konsequenz bedeutet die dargestellte grundsätzliche Übertragbarkeit der Daten, dass auch ein in den klinischen Studien zu nicht-vorbehandelten Patienten gefundener Zusatznutzen für auf die vorbehandelte Patientenpopulation, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, zu übertragen ist.

Die Übertragbarkeit der Daten wird zusätzlich gestützt durch die Ergebnisse der einarmigen Studie GS-US-236-0123, in welcher Patienten, die mindestens sechs Monate mit RAL+FTC/TDF vorbehandelt wurden, auf EVG/COBI/FTC/TDF umgestellt wurden. Nach 12 Wochen Behandlung mit EVG/COBI/FTC/TDF zeigten alle 48 (100%) Patienten virologisches Ansprechen. Hinsichtlich der CD4-Zellzahl zeigten die Patienten eine numerische, aber nicht signifikante Verbesserung (Baseline: 711 Zellen/ μ L, zu Woche 12: 744 Zellen/ μ L). In der Studie traten weder Todesfälle, noch schwerwiegende unerwünschte Ereignisse oder Therapieabbrüche aufgrund unerwünschter Ereignisse auf. 36 (75,0%) Patienten hatten unerwünschte Ereignisse; davon wurde bei elf Patienten (22,9%) das unerwünschte Ereignis als Grad 2 eingestuft. Unerwünschte Ereignisse höheren Grads traten nicht auf. 12 (25,0%) Patienten wiesen Erkrankungen des Gastrointestinaltrakts auf; hier war Diarrhoe das am häufigsten auftretende Ereignis aus dieser SOC (fünf [10,4%] Patienten). Psychiatrische Erkrankungen traten bei elf (22,9%) Patienten auf, mit Schlaflosigkeit als am häufigsten auftretendes Ereignis (fünf [10,4%] Patienten). Für eine detailliertere Übersicht der Ergebnisse aus dieser Studie siehe Abschnitt 4.3.2.3.

Aufgrund der soeben dargestellten grundsätzlichen Übertragbarkeit der Daten von nicht-vorbehandelten auf vorbehandelte Patienten, ist zu folgern, dass auch ein in den klinischen Studien zu nicht-vorbehandelten Patienten gefundener Zusatznutzen auf die vorbehandelte Patientenpopulation, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, zu übertragen ist. Dementsprechend werden für die Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF in der Population der vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, RCT mit dem zu bewertenden Arzneimittel herangezogen, die in der Population der nicht-vorbehandelten Patienten durchgeführt wurden.

Die Informationsbeschaffung für die nicht-vorbehandelten HIV-1-infizierten Patienten wird im Folgenden dargestellt. Ein- und Ausschlusskriterien der vorbehandelten Patienten (siehe Abschnitt 4.2.2) treffen hier äquivalent zu, mit dem Unterschied, dass nun nach antiretroviral nicht-vorbehandelten erwachsenen Patienten mit einer HIV-1 Infektion (Populationskriterium) gesucht wird.

Ergebnisse der Informationsbeschaffung– RCT mit dem zu bewertenden Arzneimittel (nicht-vorbehandelte Patienten)

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.

Es werden nur Studien des pharmazeutischen Unternehmers mit dem zu bewertenden Arzneimittel berücksichtigt, die entweder mit dem Single-Tablet-Regimen (EVG/COBI/FTC/TDF) oder mit den einzelnen, aber allen vier Komponenten in Kombination (EVG+COBI+FTC+TDF) in einem Studienarm durchgeführt wurden [59]. Studien, die nur

mit einzelnen, aber nicht allen vier Komponenten in Kombination durchgeführt wurden, werden hier nicht gelistet, unabhängig davon, ob sie mit der Marketing Authorisation eingereicht wurden.

Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel (nicht-vorbehandelte Patienten)

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
GS-US-236-0102	ja	laufend 48-Wochen-Interim-Bericht verfügbar 96-Wochen-Interim-Bericht verfügbar	192 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo Efavirenz 600 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo
GS-US-236-0103	ja	laufend 48-Wochen-Interim-Bericht verfügbar 96-Wochen-Interim-Bericht verfügbar	192 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo Ritonavir 100 mg + Atazanavir 300 mg + Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo
GS-US-236-0104	ja	abgeschlossen 96-Wochen-Interim-Bericht verfügbar	60 Wochen, im Anschluss daran Rollover Open-label Extension	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo Efavirenz 600 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo
GS-US-236-0128	nein	laufend	48 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo Ritonavir 100 mg + Atazanavir 300 mg + Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo
GS-US-292-0102	nein	laufend	mindestens 48 Wochen, im Anschluss daran Rollover Open-label Extension	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviralafenamid 10 mg plus Placebo Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
				245 mg plus Placebo
GS-US-292-0111	nein	laufend	mindestens 48 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviralfenamid 10 mg plus Placebo Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofovirdisoproxil 245 mg plus Placebo
GS-US-292-0104	nein	laufend	96 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviralfenamid 10 mg plus Placebo Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofovirdisoproxil 245 mg plus Placebo

Geben Sie an, welchen Stand die Information in Tabelle 4-3 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 Stand der Information in Tabelle 4-3 ist der 30. April 2013.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-3 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
GS-US-236-0128	Es liegen keine Studienergebnisse vor.
GS-US-292-0102	Es liegen lediglich 24-Wochen-Daten vor. Es handelt sich außerdem um eine Studie zu dem in Entwicklung befindlichen, aus TDF weiterentwickelten Wirkstoff Tenofovirafenamid (TAF).
GS-US-292-0111	Es liegen keine Studienergebnisse vor.
GS-US-292-0104	Es liegen keine Studienergebnisse vor.

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

Studien aus der bibliografischen Literaturrecherche

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

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

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

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

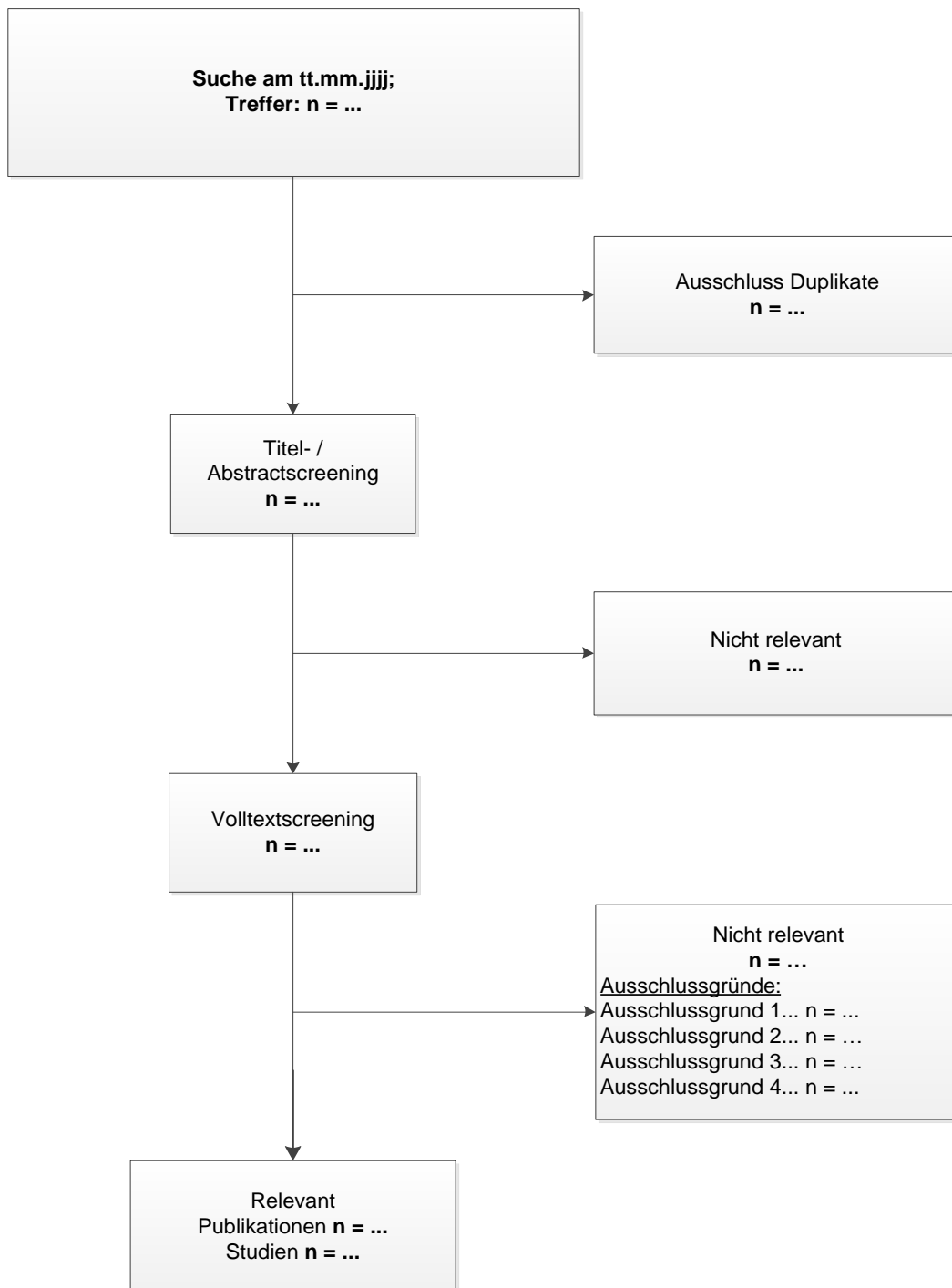


Abbildung 9: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Wie in Abschnitt 4.2.3.2 beschrieben, wurde keine bibliographische Literaturrecherche für das zu bewertende Arzneimittel durchgeführt.

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 4-3) 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 4-9: 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)
GS-US-236-0102	clinicaltrials.gov [NCT01095796] [84], ICTRP [NCT01095796] [85]	ja	n.d.
GS-US-236-0103	clinicaltrials.gov [NCT01106586], [86] ICTRP Search Portal [NCT01106586] [87]	ja	n.d.
GS-US-236-0104	clinicaltrials.gov [NCT00869557] [88], ICTRP [NCT00869557] [89]	ja	n.d.

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

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 Zwischenzeilen-

überschriften 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 4-10: 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						
Nicht zutreffend						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
GS-US-236-0102	ja	ja	ja	ja [90, 91]	ja [84, 85]	ja [92]
GS-US-236-0104	ja	ja	ja	ja [93]	ja [88, 89]	ja [94]
GS-US-236-0103	ja	ja	nein	ja [95, 96]	ja [86, 87]	ja [97]
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.						

Es konnten zwei randomisierte kontrollierte Studien für einen direkten Vergleich von EVG/COBI/FTC/TDF mit EFV/FTC/TDF, eine randomisierte kontrollierte Studie für einen direkten Vergleich von EVG/COBI/FTC/TDF mit ATV/r+FTC/TDF, jedoch keine Studie für einen direkten Vergleich von EVG/COBI/FTC/TDF mit RAL+FTC/TDF gefunden werden.

Gemäß ihrer Ein-, resp. Ausschlusskriterien entsprechen die Studienpopulationen der Population der Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind:

- In Studie GS-236-0102 wurde jegliche antiretrovirale Vorbehandlung ausgeschlossen und darüberhinaus musste eine bestätigte Sensitivität auf EFV, FTC und TDF vorliegen (genotypischer Resistenztest).
- In Studie GS-236-0103 wurde jegliche antiretrovirale Vorbehandlung ausgeschlossen und darüberhinaus musste eine bestätigte Sensitivität auf ATV, FTC und TDF vorliegen (genotypischer Resistenztest).

- In Studie GS-236-0104 wurde jegliche antiretrovirale Vorbehandlung ausgeschlossen und darüberhinaus musste eine bestätigte Sensitivität auf NRTI, NNRTI und PI vorliegen (genotypischer Resistenztest).

Dementsprechend basiert im Folgenden die Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF auf zwei aktiv-kontrollierten randomisierten Studien des pharmazeutischen Unternehmers, die das zu bewertende Arzneimittel EVG/COBI/FTC/TDF mit der zweckmäßigen Vergleichstherapie EFV/FTC/TDF direkt vergleichen (GS-US-236-0104 und GS-US-236-0102). Zusätzlich zu den Informationen aus dem Studienbericht werden post-hoc durchgeführte Subgruppenanalysen dargestellt. Die Ergebnisse aus diesem direkten Vergleich werden in Modul 4A dargestellt und auf die Population der vorbehandelten Patienten in diesem Modul 4B, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, übertragen. Die Daten werden hier nicht zusätzlich dargestellt. Für Details zu Charakteristika, Verzerrungspotenzial und Ergebnissen aus diesen Studien siehe Modul 4A.

Die Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF basiert im Folgenden auf einer aktiv-kontrollierten randomisierten Studie des pharmazeutischen Unternehmers, die das zu bewertende Arzneimittel EVG/COBI/FTC/TDF mit ATV/r+FTC/TDF direkt vergleicht (GS-US-236-0103). Es werden Daten zu Woche 48 und 96 präsentiert. Zusätzlich zu den Informationen aus dem Studienbericht werden post-hoc durchgeführte Subgruppenanalysen dargestellt.

Für Details zur Informationsbeschaffung und Ergebnissen zum Vergleich von EVG/COBI/FTC/TDF und RAL+FTC/TDF siehe Abschnitt 4.3.2.1.

Im Weiteren werden die Studien GS-US-236-0102, GS-US-236-0103 und GS-US-236-0104 aus Übersichtlichkeitsgründen mit S102, S103 und S104 abgekürzt.

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.

Tabelle 4-11: 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
S103	RCT, doppelblind, parallel	Antiretroviral nicht- vorbehandelte HIV-1- infizierte Erwachsene	EVG/COBI/FTC/ TDF: n=357 ATV/r+FTC/TDF: n=358	Screening: 5 Wochen Behandlung: 96 Wochen + Zeit bis Entblindung dann entweder: Open-label- Behandlung: bis Produkt kommerziell verfügbar oder Gilead das Forschungsprogramm dazu stoppt oder Follow-Up: 30 Tage	International: Australien, Belgien, Dänemark, Deutschland, Frankreich, Italien, Kanada, Mexiko, Niederlande, Österreich, Portugal, Puerto Rico, Schweden, Schweiz, Thailand, UK, USA Woche 48: 04/2010-09/2011 Woche 96: 04/2010-08/2012	<u>Primärer Endpunkt:</u> Virologisches Ansprechen zu Woche 48 <u>Sekundäre Endpunkte:</u> Zeit bis Verlust des virologischen Ansprechens Veränderung der CD4- Zellzahl Mortalität Unerwünschte Ereignisse

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

Studie	EVG/COBI/FTC/TDF	ATV/r+FTC/TDF	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
S103	150 mg Elvitegravir/ 150 mg Cobicistat/ 200 mg Emtricitabin/ 245 mg Tenofoviridisoproxil einmal täglich (EVG/COBI/FTC/TDF)	300 mg Atazanavir + 100 mg Ritonavir + 200 mg Emtricitabin/ 245 mg Tenofoviridisoproxil (FTC/TDF) einmal täglich	Screening: 5 Wochen Behandlung: 96 Wochen + Zeit bis zur Entblindung (Patienten kamen alle 12 Wochen zur Visite) Anschließend: Rollover Open-label Studie bis zur kommerziellen Verfügbarkeit des Medikaments oder bis Gilead das Forschungsprogramm dazu stoppt oder Follow-Up: 30 Tage

Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Alter, Geschlecht, BMI; Safety Analyse-Set)

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD)	Geschlecht w/m (%)	Baseline BMI (kg/m ²) Mittelwert (SD)
S103				
EVG/COBI/FTC/TDF	353	38 (10,5)	8,2/91,8	25,4 (5,3)
ATV/r+FTC/TDF	355	39 (9,8)	11,0/89,0	25,8 (4,8)

BMI = Body Mass Index, SD = Standardabweichung, w= weiblich, m = männlich

Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Rasse, Safety Analyse-Set)

Studie Gruppe	Rasse n (%)					
	Kaukasisch	Asiatisch	Schwarz oder afrikanischer Abstammung	Indianer/ Urein- wohner Alaskas	Hawaiianer/ Pazifik- insulaner	Andere
S103						
EVG/COBI/FTC/TDF	250 (70,8)	17 (4,8)	72 (20,4)	2 (0,6)	1 (0,3)	11 (3,1)
ATV/r+FTC/TDF	277 (78,0)	17 (4,8)	47 (13,2)	3 (0,8)	2 (0,6)	9 (2,5)

k.A. = keine Angaben, n: Anzahl Patienten in der Kategorie

Tabelle 4-15: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Ethnische Abstammung, Safety Analyse-Set)

Studie Gruppe	Ethnische Abstammung n (%)		
	Hispanisch/Latino	Nicht hispanisch/Latino	Nicht erlaubt zu fragen
S103			
EVG/COBI/FTC/TDF	64 (18,1)	284 (80,5)	5 (1,4)
ATV/r+FTC/TDF	47 (13,2)	298 (83,9)	10 (2,8)

n: Anzahl der Patienten in der Kategorie

Tabelle 4-16: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (HIV-1-RNA, Safety Analyse-Set)

Studie Gruppe	Baseline HIV-1-RNA (log ₁₀ Kopien/mL) Mittelwert (SD)	Baseline HIV-1-RNA Kopien/mL n (%)	
		≤100.000	>100.000
S103			
EVG/COBI/FTC/TDF	4,82 (0,6)	203 (57,5)	150 (42,5)
ATV/r+FTC/TDF	4,80 (0,6)	214 (60,3)	141 (39,7)

n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung, HIV = Humanes Immundefizienz-Virus

Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (CD4-Zellzahl, Safety Analyse-Set)

Studie Gruppe	CD4-Zellzahl zu Baseline (μL) Mittelwert (SD)	CD4-Zellzahl zu Baseline (μL) n (%)				
		≤50	51-≤200	201-≤350	351-≤500	>500
S103						
EVG/COBI/FTC/TDF	364 (180,6)	7 (2,0)	42 (11,9)	122 (34,6)	122 (34,6)	55 (15,6)
ATV/r+FTC/TDF	375 (158,9)	6 (1,7)	34 (9,6)	124 (34,9)	122 (34,4)	70 (19,7)

n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung

Tabelle 4-18: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (HIV-Status, Safety Analyse-Set)

Studie Gruppe	HIV Krankheitsstatus n (%)			eGFR nach Cockcroft-Gault (mL/min) Mittelwert (SD)
	Asymptomatisch	Symptomatische HIV Infektionen	Aids	
S103				
EVG/COBI/FTC/TDF	285 (80,7)	36 (10,2)	32 (9,1)	119,6 (33,6)
ATV/r+FTC/TDF	293 (82,5)	38 (10,7)	24 (6,8)	119,4 (31,3)

Aids = Erworbenes Immundefektsyndrom (Acquired immunodeficiency syndrome), eGFR = Geschätzte glomeruläre Filtrationsrate (estimated Glomerular Filtration Rate), HIV = Humanes Immundefizienz-Virus, n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung

Tabelle 4-19: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Adhärenz, Safety Analyse-Set)

Studie Gruppe	Anzahl Patienten, die mindestens eine Flasche wieder zurückgaben n (%)	Adhärenz-Rate ^a n (%)				Adhärenz- Rate Mittelwert (SD)
		<80%	≥80-<90%	≥90-<95%	≥95%	
Woche 48						
S103 ^{b,c}						
EVG/COBI/FTC/TDF	352 (99,7)	12 (3,4)	26 (7,4)	56 (15,9)	258 (73,3)	96,0 (6,2)
ATV/r+FTC/TDF	354 (99,7)	8 (2,3)	29 (8,2)	41 (11,6)	276 (78,0)	96,2 (6,2)
Woche 96						
S103 ^{b,c}						
EVG/COBI/FTC/TDF	352 (99,7)	11 (3,1)	25 (7,1)	61 (17,3)	255 (72,4)	95,7 (5,8)
ATV/r+FTC/TDF	354 (99,7)	13 (3,7)	30 (8,5)	60 (16,9)	251 (70,9)	95,4 (6,5)
^a Adhärenz-Rate berechnet sich basierend auf der Anzahl an Patienten die mindestens eine Flasche zurückgaben ^b Adhärenz-Rate wurde berechnet basierend nur auf Tabletten-Anzahl der aktiven Medikamente ^c Adhärenz-Rate war wie folgt definiert: 100 x (Gesamtanzahl eingenommener Tabletten / Gesamtanzahl verschriebener Tabletten) n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung						

Tabelle 4-20: Charakterisierung der Studienpopulationen (Patientendisposition) – RCT mit dem zu bewertenden Arzneimittel

Studie Disposition	EVG/COBI/FTC/TDF n (%)	ATV/r+FTC/TDF n (%)
S103		
Randomisiert	357	358
Behandlung erhalten	353 (100,0)	355 (100,0)
Woche 48 bis zum Daten Cut-Off in Behandlung abgeschlossen	320 (90,7)	315 (88,7)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	13 (3,7)	18 (5,1)
Tod	0 (0,0)	0 (0,0)
Schwangerschaft	1 (0,3)	0 (0,0)
Fehlende Wirksamkeit	4 (1,1)	1 (0,3)
Entscheidung des Prüfarztes	1 (0,3)	3 (0,8)
Einverständnis zurückgezogen	1 (0,3)	6 (1,7)
Lost-to-follow-up	7 (2,0)	7 (2,0)
Non-Compliance	5 (1,4)	5 (1,4)
Protokollverletzung	1 (0,3)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Woche 48 bis zum Daten Cut-Off in Studie	331 (93,8)	324 (91,3)

Studie	EVG/COBI/FTC/TDF	ATV/r+FTC/TDF
Disposition	n (%)	n (%)
abgeschlossen		
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	5 (1,4)	9 (2,5)
Tod	0 (0,0)	3 (0,8)
Schwangerschaft	1 (0,3)	0 (0,0)
Fehlende Wirksamkeit	0 (0,0)	0 (0,0)
Entscheidung des Prüfarztes	1 (0,3)	2 (0,6)
Einverständnis zurückgezogen	3 (0,8)	6 (1,7)
Lost-to-follow-up	8 (2,3)	7 (2,0)
Non-Compliance	3 (0,8)	4 (1,1)
Protokollverletzung	1 (0,3)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Woche 96 bis zum Daten Cut-Off in Behandlung abgeschlossen	304 (86,1)	300 (84,5)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	15 (4,2)	21 (5,9)
Tod	0 (0,0)	0 (0,0)
Schwangerschaft	3 (0,8)	1 (0,3)
Fehlende Wirksamkeit	4 (1,1)	1 (0,3)
Entscheidung des Prüfarztes	4 (1,1)	5 (1,4)
Einverständnis zurückgezogen	3 (0,8)	11 (3,1)
Lost-to-follow-up	10 (2,8)	10 (2,8)
Non-Compliance	9 (2,5)	6 (1,7)
Protokollverletzung	1 (0,3)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Woche 96 bis zum Daten Cut-Off in Studie abgeschlossen	313 (88,7)	310 (87,3)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	5 (1,4)	10 (2,8)
Tod	1 (0,3)	3 (0,8)
Schwangerschaft	1 (0,3)	1 (0,3)
Fehlende Wirksamkeit	0 (0,0)	0 (0,0)
Entscheidung des Prüfarztes	3 (0,8)	5 (1,4)
Einverständnis zurückgezogen	9 (2,5)	11 (3,1)
Lost-to-follow-up	12 (3,4)	10 (2,8)
Non-Compliance	8 (2,3)	5 (1,4)
Protokollverletzung	1 (0,3)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
n: Anzahl der Patienten in der Kategorie		

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

Design

Die in die Nutzenbewertung von EVG/COBI/FTC/TDF gegenüber der Vergleichstherapie ATV/r+FTC/TDF eingeschlossene Studie wurde in einem randomisierten aktiv-kontrollierten Parallelgruppendesign durchgeführt. Die Patienten wurden entweder mit EVG/COBI/FTC/TDF (150 mg Elvitegravir/150 mg Cobicistat/200 mg Emtricitabin/245 mg Tenofoviridisoproxil) oder mit dem Protease-Inhibitor Atazanavir (300 mg), geboostert mit 100 mg Ritonavir, in Kombination mit 200 mg Emtricitabin/245 mg Tenofoviridisoproxil FTC/TDF behandelt.

Nach einer 5-wöchigen Screening-Phase wurden die Patienten 96 Wochen lang in einem doppelblinden Design behandelt. Anschließend wurden die Patienten weiter bis zur Entblindung verblindet behandelt, wobei sie alle 12 Wochen zur Visite erscheinen mussten. Danach hatten die Patienten die Möglichkeit eines Rollovers in eine Open-Label-Studie, in der sie mit EVG/COBI/FTC/TDF behandelt wurden, bis dieses Arzneimittel kommerziell verfügbar war oder Gilead das Forschungsprogramm dazu stoppt. Nutzten die Patienten diese Option nicht, traten sie in die 30-tägige Follow-up Phase ein.

Die Studie war international und multizentrisch. Die Zentren erstreckten sich auf die folgenden Länder: Australien, Belgien, Dänemark, Deutschland, Frankreich, Italien, Kanada, Mexiko, Niederlande, Österreich, Portugal, Puerto Rico, Schweden, Schweiz, Thailand, UK und USA.

Die Studie wurde in einem doppel-blinden double-dummy Design durchgeführt, um die Verblindung zu gewährleisten, d. h. weder den Prüfärzten noch den Patienten war die Gruppenzugehörigkeit der Patienten bekannt. Dazu wies das interactive voice response system (IVRS)/ interactive web response system (IWRS) den Studienmedikations-Flaschen verblindete Flaschennummern zu.

Die Patienten wurden durch das IVRS/IWRS in einem Verhältnis von 1:1 auf die beiden Behandlungsarme randomisiert (Randomisierung durchgeführt durch Bracket). Dabei wurde nach Screening HIV-1-RNA Level (≤ 100.000 Kopien/mL oder > 100.000 Kopien/mL) stratifiziert. Es wurde eine Blocklänge von 4 verwendet.

Studienpopulation

In Studie S103 wurden erwachsene Patienten eingeschlossen, die eine Infektion mit HIV-1 aufwiesen und antiretroviral nicht-vorbehandelt waren. Die Patienten mussten außerdem unter anderem eine Viruslast ≥ 5.000 Kopien/mL aufweisen und außerdem eine adäquate Nierenfunktion, definiert durch eGFR ≥ 70 mL/min gemäß Cockcroft-Gault, aufweisen. Des Weiteren durften die Patienten keinerlei antiretrovirale Vorbehandlung bekommen haben; die Sensitivität gegenüber den Wirkstoffen FTC, TDF und ATV musste zusätzlich anhand eines genotypischen Resistenztests dokumentiert worden sein. Da INI-assoziierte Primärresistenzen nicht beschrieben sind (siehe Module 3A und 3B), wurden durch die im Studiendesign

vorgegebene fehlende Vorbehandlung per se nur Patienten eingeschlossen, die keine Resistenz-assoziierten Mutationen gegen EVG aufweisen; Resistenz-assoziierte Mutationen gegen FTC, TDF und ATV wurden zusätzlich direkt ausgeschlossen (genotypischer Resistenztest). Somit entspricht diese Studienpopulation der Population von HIV-1 infizierten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind. Wichtige Ausschlusskriterien waren unter anderem eine Hepatitis C-Behandlung, Schwangerschaft, schwere Infektionen oder die Einnahme von Medikamenten, die nicht mit den in der Studie verabreichten Behandlungen eingenommen werden sollte.

Die Patienten waren im Mittel 38 (EVG/COBI/FTC/TDF-Arm) bzw. 39 Jahre (ATV/r+FTC/TDF) alt. Es wurden deutlich mehr Männer in die Studie eingeschlossen (EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF: 91,8% vs. 89,0%), was die höhere Prävalenz der HIV-1-Infektion unter Männern in Europa und den USA widerspiegelt. Der Baseline-BMI lag in der EVG/COBI/FTC/TDF-Gruppe im Mittel bei 25,4 kg/m² und in der ATV/r+FTC/TDF-Gruppe bei durchschnittlich 25,8 kg/m² (siehe Tabelle 4-13).

Der Großteil der Patienten in der Studie war kaukasischer Herkunft (EVG/COBI/FTC/TDF: 70,8%, ATV/r+FTC/TDF: 78,0%), wobei dieser Anteil im ATV/r+FTC/TDF-Arm etwas höher war als im EVG/COBI/FTC/TDF-Arm. Mit 20,4% der Patienten im EVG/COBI/FTC/TDF-Arm und 13,2% der Patienten im ATV/r+FTC/TDF-Arm bildeten schwarze Patienten und Patienten afrikanischer Abstammung die zweitgrößte Gruppe in dieser Studie. Die übrigen Patienten waren asiatischer (4,8% je Behandlungsarm) oder hawaiianischer (EVG/COBI/FTC/TDF: 0,3%, ATV/r+FTC/TDF: 0,6%) Herkunft, und ein kleiner Teil der Patienten waren Indianer oder Ureinwohner Alaskas (EVG/COBI/FTC/TDF: 0,6%, ATV/r+FTC/TDF: 0,8%). Außerdem gehörten im EVG/COBI/FTC/TDF-Arm 3,1% und im ATV/r+FTC/TDF-Arm 2,5% der Patienten einer anderen Rasse an (siehe Tabelle 4-14). Der Großteil der Patienten war nicht hispanischer/Latino-Abstammung (EVG/COBI/FTC/TDF: 80,5%, ATV/r+FTC/TDF: 83,9%). Außerdem war es bei 1,4% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 2,8% der Patienten im ATV/r+FTC/TDF-Arm nicht erlaubt nach der ethnischen Abstammung zu fragen (siehe Tabelle 4-15).

Die Ausgangsviruslast lag bei durchschnittlich 4,82 log₁₀ HIV-1-Kopien/mL im EVG/COBI/FTC/TDF-Arm und 4,80 log₁₀ Kopien/mL im ATV/r-FTC/TDF-Arm auf (siehe Tabelle 4-16). Auch hinsichtlich der kategorialen Viruslast waren die Patienten zwischen den beiden Behandlungsarmen balanciert (≤100.000/>100.000 Kopien/mL im EVG/COBI/FTC/TDF-Arm: 57,5%/42,5%; im ATV/r+FTC/TDF-Arm: 60,3%/39,7%). Die mittlere CD4-Zellzahl lag zu Baseline bei 364 Zellen/μL im EVG/COBI/FTC/TDF-Arm und 375 Zellen/μL im ATV/r+FTC/TDF-Arm. Bei dem Großteil der Patienten lag die CD4-Zellzahl zu Baseline zwischen 201 Zellen/μL und 500 Zellen/μL. Im EVG/COBI/FTC/TDF-Arm waren dies in den Gruppen 201-≤350 Zellen/μL und 351-≤500 Zellen/μL jeweils 34,6%. Im ATV/r-FTC/TDF-Arm lag dieser Anteil bei 34,9% in der 201-≤350 Zellen/μL-Gruppe und 34,4% in der 351-≤500 Zellen/μL-Gruppe (siehe Tabelle 4-17). Zu Baseline war ein Großteil der Patienten klinisch asymptomatisch (EVG/COBI/FTC/TDF: 80,7%, ATV/r+FTC/TDF:

82,5%), 10,2% im EVG/COBI/FTC/TDF-Arm und 10,7% im ATV/r+FTC/TDF-Arm hatten symptomatische HIV-Infektionen und 9,1% der Patienten im EVG/COBI/FTC/TDF-Arm und 6,8% der Patienten im ATV/r+FTC/TDF-Arm hatten zu Baseline Aids-definierende Erkrankungen (siehe Tabelle 4-18). Die eGFR nach Cockcroft-Gault lag zu Baseline im Mittel bei 119,6 mL/min im EVG/COBI/FTC/TDF-Arm und bei 199,4 mL/min im ATV/r+FTC/TDF-Arm.

Zur Berechnung der Adhärenz der Patienten wurde mit der Anzahl an Patienten gerechnet, die mindestens eine Flasche der Medikation wieder zurückgaben. Dies waren 99,7% je Behandlungsarm (siehe Tabelle 4-19). Davon wiesen 73,3% der Patienten im EVG/COBI/FTC/TDF-Arm und 78,0% der Patienten im ATV/r+FTC/TDF-Arm eine Adhärenzrate von $\geq 95\%$ auf. 15,9% der Patienten im EVG/COBI/FTC/TDF-Arm und 11,6% der Patienten im ATV/r+FTC/TDF-Arm wiesen eine Adhärenzrate zwischen $\geq 90\%$ und $< 95\%$ auf und lediglich ca. 10% pro Behandlungsarm wiesen eine Adhärenzrate unter 90% auf. Zu Woche 96 sind die Adhärenzraten sehr ähnlich.

In der Studie erhielten 353 Patienten EVG/COBI/FTC/TDF und 355 Patienten erhielten ATV/r+FTC/TDF. Davon beendeten 90,7% der Patienten im EVG/COBI/FTC/TDF-Arm und 88,7% der Patienten im ATV/r+FTC/TDF-Arm die ersten 48 Wochen der Behandlung in Behandlung (siehe Tabelle 4-182). Hauptgrund für den vorzeitigen Abbruch der Behandlung waren unerwünschte Ereignisse (EVG/COBI/FTC/TDF: 3,7%; ATV/r+FTC/TDF: 5,1%). Woche 96 beendeten 86,1% der EVG/COBI/FTC/TDF-Patienten und 84,5% der ATV/r+FTC/TDF-Patienten in Behandlung. Hauptabbruchgrund war dabei Lost-to-follow-up (EVG/COBI/FTC/TDF: 2,8%; ATV/r+FTC/TDF: 2,8%).

Insgesamt sind die Behandlungsarme untereinander sehr balanciert, es gibt keinerlei relevante Unterschiede zwischen den Behandlungsarmen.

Haupt-Analyse-Set für Wirksamkeitsendpunkte war das ITT-Set mit 353 Patienten im EVG/COBI/FTC/TDF-Arm und 355 Patienten im ATV/r+FTC/TDF-Arm. Auch in das Safety Analyse-Set (Haupt-Analyse-Set für Sicherheitsendpunkte) wurden 353 Patienten in den EVG/COBI/FTC/TDF-Arm und 355 Patienten in den ATV/r+FTC/TDF-Arm eingeschlossen. In das Per Protokoll-Set wurden zu Woche 48 318 Patienten in den EVG/COBI/FTC/TDF-Arm und 310 Patienten in den ATV/r+FTC/TDF-Arm eingeschlossen, zu Woche 96 waren dies 297 Patienten im EVG/COBI/FTC/TDF-Arm bzw. 294 Patienten im ATV/r+FTC/TDF-Arm.

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 4-21: 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			
S103	ja	ja	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In Studie S103 werden keine expliziten Angaben zur Erzeugung der Randomisierungsliste oder zur Verdeckung der Gruppenzuteilung gemacht. Da jedoch ein IVRS/IWRS zur Randomisierung verwendet wurde, wird davon ausgegangen dass die Erzeugung der Randomisierungsliste adäquat durchgeführt wurde und dass die Gruppenzuteilung verdeckt war.

Es wurde eine nach HIV-1-Viruslast zum Zeitpunkt des Screenings (≤ 100.000 oder > 100.000 Kopien/mL) stratifizierte Randomisierung (durchgeführt durch Bracket) durchgeführt. Die Patienten wurden im Verhältnis 1:1 den beiden Behandlungsarmen zugeteilt. Es waren sowohl die Patienten als auch die Prüfarzte gegenüber der Gruppenzugehörigkeit der Patienten verblindet.

Es können keine Hinweise auf ergebnisgesteuerte Berichterstattung gefunden werden und es gab auch keine sonstigen das Verzerrungspotenzial beeinflussenden Punkte. Insgesamt wird das Verzerrungspotenzial auf Studienebene für Studie S103 als niedrig eingestuft.

Für die unterschiedlichen Analysezeitpunkte, die im vorliegenden Nutzendossier dargestellt werden, wurden jeweils die zulassungsrelevanten Daten verwendet. D. h. für Endpunkte zu Woche 48 wurde der Daten Cut-Off zu Woche 48 verwendet und für Endpunkte zu Woche 96 wurde der Daten Cut-Off zu Woche 96 verwendet.

Für die hier eingeschlossene Studie wurden die Woche 48 und 96 Daten für die Veränderung der CD4-Zellzahl nachberechnet [98, 99].

Außerdem wurden Hazard Ratios für die Zeit bis zum Verlust des virologischen Ansprechens nachberechnet [100, 101].

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 4-22: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Virologisches Ansprechen	Zeit bis zum Verlust des virologischen Ansprechens	Veränderung der CD4-Zellzahl	Mortalität ^a
S103	ja	ja	ja	ja
^a Die Endpunkte Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse, unerwünschte Ereignisse von Interesse und häufigste unerwünschte Ereignisse (SOC) werden im Folgenden in einem gemeinsamen Abschnitt "Unerwünschte Ereignisse" dargestellt. Das Verzerrungspotenzial wird für diese Endpunkte pro Studie gemeinsam bewertet und unter Abschnitt 4.3.1.3.1.4 detailliert diskutiert. Es ist möglich diese Endpunkte pro Studie zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und pro Studie die gleiche Anzahl an Patienten in die Analyse einging.				

Tabelle 4-23: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Gesamtrate unerwünschter Ereignisse ^a	Schwerwiegende unerwünschte Ereignisse ^a	Unerwünschte Ereignisse Grad 3-4 ^a	Therapieabbrüche aufgrund unerwünschter Ereignisse ^a
S103	ja	ja	ja	ja
^a Die Endpunkte Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse, unerwünschte Ereignisse von Interesse und häufigste unerwünschte Ereignisse (SOC) werden im Folgenden in einem gemeinsamen Abschnitt "Unerwünschte Ereignisse" dargestellt. Das Verzerrungspotenzial wird für diese Endpunkte pro Studie gemeinsam bewertet und unter Abschnitt 4.3.1.3.1.4 detailliert diskutiert. Es ist möglich diese Endpunkte pro Studie zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und pro Studie die gleiche Anzahl an Patienten in die Analyse einging.				

Tabelle 4-24: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Wichtige unerwünschte Ereignisse ^a	Unerwünschte Ereignisse von Interesse ^a	Häufigste unerwünschte Ereignisse (SOC) ^a
S103	ja	ja	ja

^a Die Endpunkte Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse, unerwünschte Ereignisse von Interesse und häufigste unerwünschte Ereignisse (SOC) werden im Folgenden in einem gemeinsamen Abschnitt "Unerwünschte Ereignisse" dargestellt. Das Verzerrungspotenzial wird für diese Endpunkte pro Studie gemeinsam bewertet und unter Abschnitt 4.3.1.3.1.4 detailliert diskutiert. Es ist möglich diese Endpunkte pro Studie zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und pro Studie die gleiche Anzahl an Patienten in die Analyse einging.

4.3.1.3.1 Endpunkte– 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.

In den folgenden Abschnitten werden die Operationalisierung, das Verzerrungspotenzial, die Ergebnisse aus den einzelnen Studien und die Meta-Analysen der patientenrelevanten Endpunkte in separaten Abschnitten dargestellt.

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 Virologisches Ansprechen– RCT

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 virologisches Ansprechen

Studie	Operationalisierung
S103	Anteil an Patienten mit Viruslast <50 Kopien/mL zu Woche 48 bzw. 96. Zur Messung der HIV-1-RNA Plasma Konzentrationen wurde der COBAS Amplicor HIV-1 Monitor Test (Version 1.5) verwendet. Virologisches Ansprechen wurde mittels Snapshot-Algorithmus oder TLOVR-Algorithmus erfasst.
Studie	Operationalisierung
S103	Anteil an Patienten mit Viruslast <50 Kopien/mL zu Woche 48 bzw. 96. Zur Messung der HIV-1-RNA Plasma Konzentrationen wurde der COBAS Amplicor HIV-1 Monitor Test (Version 1.5) verwendet. Virologisches Ansprechen wurde mittels Snapshot-Algorithmus oder TLOVR-Algorithmus erfasst.

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 virologisches Ansprechen 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
S103	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Virologisches Ansprechen wurde definiert als Absinken der HIV-1-RNA-Viruslast unter die Nachweisgrenze von 50 Kopien/mL. Dabei wurde der FDA Snapshot-Algorithmus als primärer Analyse-Algorithmus gewählt, wobei das Analysefenster zu Woche 48 definiert war

von Tag 309 bis inklusive Tag 378. Zur Messung der Plasma Viruslast wurde der COBAS Amplicor HIV-1 Monitor-Test (Version 1.5) verwendet.

Snapshot-Algorithmus

Patienten wurden als **virologische Ansprechere** gemäß Snapshot definiert, wenn deren letzter im Analysefenster verfügbarer HIV-1-RNA-Wert bei <50 Kopien/mL lag, während sie sich in randomisierter Behandlung befanden.

Als **virologische Versager** wurden Patienten definiert,

- a) wenn deren letzter im Analysefenster verfügbarer HIV-1-RNA-Wert ≥ 50 Kopien/mL betrug oder
- b) die keine HIV-1-RNA-Daten im Woche 48-Analysefenster in Behandlung aufwiesen und
 - i. die Behandlung vor oder in dem Woche 48-Analysefenster aufgrund fehlender Wirksamkeit abbrechen oder
 - ii. die Behandlung vor oder in dem Woche 48 Analysefenster aus anderen Gründen als unerwünschte Ereignisse, Tod und fehlende Wirksamkeit abbrechen und einen letzten verfügbaren HIV-1-RNA-Wert ≥ 50 Kopien in Behandlung aufwiesen.

In die Kategorie „**Keine virologischen Daten im Woche 48 Analysefenster**“ wurden Patienten eingeordnet, die keine HIV-1-RNA-Werte im Woche 48-Analysefenster in Behandlung aufwiesen, aufgrund

- a) Abbruch der Studienbehandlung vor oder in dem Woche 48-Analysefenster aufgrund unerwünschter Ereignisse oder Tod (unabhängig davon ob der zuletzt verfügbare HIV-1-RNA-Wert vor oder zum Zeitpunkt des Abbruchs <50 Kopien/mL lag oder nicht) oder
- b) Abbruch der Studienbehandlung vor oder in dem Woche 48-Analysefenster aus anderen Gründen als unerwünschte Ereignisse/Tod und fehlende Wirksamkeit (d. h. Einverständnis zurückgezogen, lost-to-follow-up, Umzug, etc.) und der zuletzt verfügbare HIV-1-RNA-Wert in Behandlung lag <50 Kopien/mL oder
- c) fehlender Werte während des Analysefensters und in Behandlung.

Zu Woche 96 war diese Operationalisierung identisch, mit dem Unterschied dass das Analysefenster zu Woche 96 betrachtet wurde, welches von Tag 631 bis einschließlich Tag 714 definiert war.

Zusätzlich zum Snapshot-Algorithmus wurde der FDA TLOVR-Algorithmus verwendet.

TLOVR-Algorithmus

Die Definition des virologischen Ansprechens gemäß TLOVR-Algorithmus umfasste die folgenden drei Kriterien:

1. Erreichen eines bestätigten Abfalls der Viruslast HIV-1-RNA <50 Kopien/mL an zwei aufeinanderfolgenden Visiten vor oder an der oberen Grenze des Woche 48-Analysefensters.
2. Kein Eintritt der folgenden Ereignisse vor oder an der oberen Grenze des Woche 48-Analysefensters:
 - a) Tod
 - b) Vorzeitiger Therapieabbruch
3. Kein bestätigter virologischer Rebound (d. h. HIV-1-RNA ≥ 50 Kopien/mL an zwei aufeinanderfolgenden Visiten oder zuletzt verfügbarer HIV-1-RNA-Wert während der Studie ≥ 50 Kopien/mL, gefolgt von vorzeitigem Abbruch der Behandlung) bis zur oberen Grenze des Woche 48-Analysefensters, nachdem ein bestätigter Abfall der Viruslast unter die Nachweisgrenze von 50 HIV-1-RNA Kopien/mL erreicht wurde.

Für eine bestätigte Viruslast unter der Nachweisgrenze von 50 HIV-1-RNA Kopien/mL oder den bestätigten virologischen Rebound musste die erste HIV-1-RNA-Messung bei oder vor der oberen Grenze des Analysefensters erfolgen. Die bestätigende Messung (d. h. die zweite der aufeinanderfolgenden HIV-1-RNA-Messungen oder der vorzeitige Therapieabbruch) konnte nach der oberen Grenze des Woche 48-Analysefensters erfolgen.

Patienten, die mindestens eine der drei Kategorien nicht erfüllten, wurden als virologische Versager eingestuft und in die folgenden Kategorien eingeordnet:

- **Never suppressed:** Patienten, die bis zur oberen Grenze des Analysefensters niemals einen bestätigten Abfall (d. h. HIV-1-RNA <50 Kopien/mL an zwei aufeinanderfolgenden Visiten) erreichten.
- **Rebound:** Patienten, die bis zur oberen Grenze des Woche 48-Analysefensters einen bestätigten virologischen Rebound erlitten (d. h. HIV-1-RNA ≥ 50 Kopien/mL an zwei aufeinanderfolgenden Visiten oder zuletzt verfügbarer HIV-1-RNA-Wert ≥ 50 Kopien/mL gefolgt von vorzeitigem Therapieabbruch) nachdem sie einen bestätigten Abfall erreicht hatten.
- **Drug Discontinuation due to AE/Death and Other Reason:** Patienten, die die Behandlung bis zur oberen Grenze des Woche 48-Analysefensters aufgrund unerwünschter Ereignisse/Tod und anderer Gründe (auf dem electronic case report form [eCRF] dokumentiert) abbrachen.

Die Operationalisierung war zu Woche 96 identisch zu der zu Woche 48, mit der Ausnahme dass das Woche 96-Analysefenster betrachtet wurde, welches von Tag 631 bis einschließlich Tag 714 definiert wurde. Zu Woche 96 wurde der Therapieabbruch aufgrund fehlender Wirksamkeit ebenfalls unter virologischem Versagen betrachtet.

Zur Darstellung der Nichtunterlegenheit zur 12%-Grenze werden zusätzlich Risikodifferenzen dargestellt.

Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung. Da auch keine sonstigen das Verzerrungspotenzial beeinflussenden Faktoren gefunden werden können, wird das Verzerrungspotenzial als niedrig eingestuft.

Entsprechend der Auslegung des IQWiG in der Nutzenbewertung zu Rilpivirin (A12-04, Version 1.0, Stand 12.04.2012) wird in der vorliegenden Nutzenbewertung nur der Anteil an virologischen Ansprechern betrachtet, da davon ausgegangen wird, dass der Endpunkt virologisches Versagen (Wirksamkeit bzw. Resistenzen) hinreichend durch den Endpunkt virologisches Ansprechen abgedeckt ist [62].

Dessen ungeachtet werden hier als zusätzliche Information auch die Anteile der Patienten mit virologischem Versagen zu Woche 48 und 96 auf Endpunktebene dargestellt. Nutzenableitungen werden, basierend auf diesen Daten, jedoch nicht angestellt.

Für virologisches Versagen (Wirksamkeit) werden für den Snapshot-Algorithmus die folgenden Kriterien als relevant eingeschätzt:

1. letzter im Analysefenster verfügbarer HIV-1-RNA-Wert ≥ 50 Kopien/mL oder
2. keine HIV-1-RNA-Daten im Woche 48-Analysefenster in Behandlung und die Behandlung vor oder im Woche 48-Analysefenster aufgrund fehlender Wirksamkeit abgebrochen.

Für den TLOVR-Algorithmus werden für virologisches Versagen (Wirksamkeit) die folgenden Kriterien als relevant eingeschätzt:

1. bis zur oberen Grenze des Woche 48-Analysefensters einen virologischen Rebound erlitten (HIV-1 RNA ≥ 50 Kopien/mL an zwei aufeinander folgenden Visiten oder zuletzt verfügbarer HIV-1 RNA-Wert ≥ 50 Kopien/mL gefolgt von vorzeitigem Abbruch der Studienmedikation, nachdem virologisches Ansprechen erreicht wurde) (Rebound) oder
2. bis zur oberen Grenze des Woche 48-Analysefensters nie virologisches Ansprechen erreicht (definiert durch HIV-1 RNA < 50 Kopien/mL an zwei aufeinander folgenden Visiten) (Never suppressed).

Die Operationalisierungen in Studie S103 waren zu Woche 96 jeweils die gleichen.

Das Verzerrungspotenzial von virologischem Versagen (Wirksamkeit) ist vergleichbar mit dem des Endpunkts virologisches Ansprechen, was sowohl auf den Snapshot- als auch den TLOVR-Algorithmus zutrifft. Das ITT-Prinzip wurde adäquat umgesetzt, die Endpunkterheber waren verblindet, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung und auch sonstige das Verzerrungspotenzial beeinflussende Punkte können nicht gefunden 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 4-27: Ergebnisse für virologisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set)

Studie S103 Zeitpunkt Algorithmus	EVG/COBI/FTC/ TDF		ATV/r+FTC/ TDF		Behandlungseffekt	
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^b	RD [95%-KI] ^c
Woche 48						
<i>Snapshot</i>	353	316 (89,5)	355	308 (86,8)	1,30 [0,82;2,06], p=0,22	3,0 [-1,9;7,8] ^d
<i>TLOVR</i>	353	304 (86,1)	355	301 (84,8)	1,11 [0,73;1,69], p=0,55	1,6 [-3,6;6,8]
Woche 96						
<i>Snapshot</i>	353	294 (83,3)	355	292 (82,3)	1,08 [0,73; 1,59], p=0,70	1,1 [-4,5;6,7]
<i>TLOVR</i>	353	287 (81,3)	355	275 (77,5)	1,27 [0,88;1,82], p=0,19	4,0 [-2,0;9,9]
^a zusätzlich berechnet ^b p-Wert aus Cochran Mantel Haenszel Test stratifiziert nach Baseline HIV-1-RNA Schicht ^c basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen berechnet. ^d aufgrund zweier IDMC (Independent Data Monitoring Committee) Meetings, bei denen jeweils eine Wirksamkeitsanalyse zum Niveau $\alpha=0,001$ durchgeführt wurde, wurde, um das globale α -Niveau zu sichern, für den primären Endpunkt ein 95,2%-Konfidenzintervall (entsprechend: $\alpha=0,048$) berechnet.						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

Zu Woche 48 konnte im ITT-Set sowohl mittels des Snapshot-Algorithmus als auch mittels des TLOVR-Algorithmus die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze gezeigt werden. Die Risikodifferenz [95%-KI] bei Verwendung des Snapshot-Algorithmus ergibt sich zu 3,0 [-1,9;7,8]. Unter Anwendung des TLOVR-Algorithmus beträgt die Risikodifferenz 1,6 [-3,6;6,8]. Ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht bei keinem Algorithmus (OR [95%-KI]: Snapshot: 1,30 [0,82;2,06], p=0,22; TLOVR: 1,11 [0,73;1,69], p=0,55).

Auch zu Woche 96 konnte die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF jeweils unter Verwendung des Snapshot und des TLOVR-Algorithmus gezeigt werden. Die Untergrenzen der Konfidenzintervalle zu den jeweiligen Risikodifferenzen übersteigen die -12,0 deutlich (RD [95%-KI]: Snapshot: 1,1 [-4,5;6,7]; TLOVR: 4,0 [-2,0;9,9]). Das OR [95%-KI] unter Verwendung des Snapshot-Algorithmus ergibt sich zu 1,08 [0,73; 1,59] mit p=0,70 und unter Anwendung des TLOVR-Algorithmus beträgt das OR [95%-KI] 1,27 [0,88;1,82] mit p=0,19.

Tabelle 4-28: Ergebnisse für virologisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set)

Studie S103 Zeitpunkt Algorithmus	EVG/COBI/FTC/ TDF		ATV/r+FTC/TDF		Behandlungseffekt	
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^b	RD [95%-KI] ^c
Woche 48						
<i>Snapshot</i>	318	310 (97,5)	310	303 (97,7)	0,90 [0,32;2,50], p=0,95	-0,1 [-2,6;2,4] ^d
Woche 96						
<i>Snapshot</i>	297	288 (97,0)	294	283 (96,3)	1,24 [0,51;3,05], p=0,59	0,8 [-2,2;3,8]
^a zusätzlich berechnet ^b p-Wert aus Cochran Mantel Haenszel Test stratifiziert nach Baseline HIV-1-RNA Schicht ^c basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen berechnet. ^d aufgrund zweier IDMC (Independent Data Monitoring Committee) Meetings, bei denen jeweils eine Wirksamkeitsanalyse zum Niveau $\alpha=0,001$ durchgeführt wurde, wurde, um das globale α -Niveau zu sichern, für den primären Endpunkt ein 95,2%-Konfidenzintervall (entsprechend: $\alpha=0,048$) berechnet.						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

Im PP Analyse-Set konnte zu Woche 48 die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze gezeigt werden (RD [95%-KI]: -0,1 [-2,6;2,4]). Ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht nicht (OR [95%-KI]: 0,90 [0,32;2,50], p=0,95). Zu Woche 96 besteht ebenfalls kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,24 [0,51;3,05], p=0,59), doch auch hier kann die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze gezeigt werden (RD [95%-KI]: 0,8 [-2,2;3,8]).

Tabelle 4-29: Ergebnisse für virologisches Versagen (Wirksamkeit) aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 48)

Studie S103 Zeitpunkt Algorithmus	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^a
Woche 48					
<i>Snapshot</i>					
HIV-1 RNA ≥ 50 Kopien/mL	353	7 (2,0)	355	8 (2,3)	0,88 [0,31;2,45], p=0,8028
Therapieabbruch aufgrund fehlender Wirksamkeit	353	4 (1,1)	355	0 (0,0)	9,15 [0,49;170,67], p=0,1379
TLOVR					
Rebound	353	12 (3,4)	355	8 (2,3)	1,53 [0,62;3,78], p=0,3608
Never suppressed	353	7 (2,0)	355	7 (2,0)	1,01 [0,35;2,90], p=0,9915
Woche 96					
<i>Snapshot</i>					

Studie S103 Zeitpunkt Algorithmus	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^a
HIV-1 RNA \geq 50 Kopien/mL	353	7 (2,0)	355	11 (3,1)	0,63 [0,24;1,65], p=0,3496
Therapieabbruch aufgrund fehlender Wirksamkeit	353	4 (1,1)	355	1 (0,3)	4,06 [0,45;36,48], p=0,2114
TLOVR					
Rebound	353	24 (6,8)	355	28 (7,9)	0,85 [0,48;1,50], p=0,5791
Never suppressed	353	1 (0,3)	355	1 (0,3)	1,01 [0,06;16,14], p=0,9968
Therapieabbruch aufgrund fehlender Wirksamkeit	353	2 (0,6)	355	0 (0,0)	5,06 [0,24;105,71], p=0,2960
^a zusätzlich berechnet					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

Tabelle 4-29 zeigt, dass es hinsichtlich keines als relevant eingeschätzten Kriteriums für virologisches Versagen einen statistisch signifikanten Unterschied zwischen den Behandlungen gibt.

Virologisches Versagen aufgrund Wirksamkeit (Snapshot; HIV-1 RNA \geq 50 Kopien/mL und Therapieabbruch aufgrund fehlender Wirksamkeit kombiniert)

Hinsichtlich virologischen Versagens aufgrund Wirksamkeit (Snapshot), definiert durch die Anzahl an Patienten, die aufgrund eines HIV-1 RNA-Werts \geq 50 Kopien/mL oder eines Therapieabbruchs aufgrund fehlender Wirksamkeit, als virologische Versager eingestuft wurden (EVG/COBI/FTC/TDF: 11 [3,1%], ATV/r+FTC/TDF: 8 [2,3%]), besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,40 [0,55;3,51], p=0,4795). Auch zu Woche 96 (EVG/COBI/FTC/TDF: 11 [3,1%], ATV/r+FTC/TDF: 12 [3,4%]) kann hier kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (OR [95%-KI]: 0,92 [0,40;2,11], p=0,8429).

Virologisches Versagen aufgrund Wirksamkeit (TLOVR; Rebound und never suppressed kombiniert)

Mittels TLOVR-Algorithmus besteht hinsichtlich virologischen Versagens aufgrund Wirksamkeit zu Woche 48, definiert durch die Anzahl an Patienten mit virologischem Rebound oder Patienten, die nie virologisches Ansprechen erreicht haben (EVG/COBI/FTC/TDF: 19 [5,4%], ATV/r+FTC/TDF: 15 [4,2%]) kein statistisch signifikanter Unterschied (OR [95%-KI]: 1,29 [0,64;2,58], p=0,4726). Und auch zu Woche 96 (ohne Berücksichtigung des Kriteriums „Therapieabbruch aufgrund fehlender Wirksamkeit“: EVG/COBI/FTC/TDF: 25 [7,1%], ATV/r+FTC/TDF: 29 [8,2%]; unter Berücksichtigung des Kriteriums „Therapieabbruch aufgrund fehlender Wirksamkeit“: EVG/COBI/FTC/TDF: 27 [7,6%], ATV/r+FTC/TDF: 29 [8,2%]) kann hier kein statistisch signifikanter Unterschied zwischen den Behandlungen festgestellt werden (ohne Berücksichtigung des Kriteriums

„Therapieabbruch aufgrund fehlender Wirksamkeit“: OR [95%-KI]: 0,86 [0,49;1,49], p=0,5862; unter Berücksichtigung des Kriteriums „Therapieabbruch aufgrund fehlender Wirksamkeit“: OR [95%-KI]: 0,93 [0,54;1,61], p=0,7976).

Im Weiteren werden die Ergebnisse zum virologischen Versagen nicht weiter berücksichtigt, da sie als durch den Endpunkt virologisches Ansprechen abgedeckt angesehen werden. Demnach werden auch keine Subgruppenanalysen zu diesem Endpunkt durchgeführt.

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.

Da hier nur eine Studie in die Nutzenbewertung von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF eingeschlossen wurde, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.1.2 Zeit bis Verlust des virologischen Ansprechens – RCT

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

Tabelle 4-30: Operationalisierung von Zeit bis Verlust des virologischen Ansprechens

Studie	Operationalisierung
S103	Die Zeit bis zum Verlust des virologischen Ansprechens ist für Patienten, die ein virologisches Ansprechen (Abfall der Viruslast unter die Nachweisgrenze von 50 Kopien/mL) erreicht haben, operationalisiert als Zeit bis zum Eintritt des Todes oder des erstmaligen Abbruchs der Studienmedikation, Wechsels zu einem neuen Therapieregime, bestätigten HIV-1-RNA-Werts ≥ 50 Kopien/mL oder HIV-1-RNA-Werts ≥ 50 Kopien/mL gefolgt von einem dauerhaften Abbruch der Studienmedikation. Für Patienten, deren Viruslast nie unter die Nachweisgrenze gesenkt werden konnte, wird die Zeit bis zum Verlust des virologischen Ansprechens auf Studientag 1 gesetzt. Wenn ein Ereignis, basierend auf TLOVR, direkt gefolgt wurde von einem oder mehreren verpassten Untersuchungen, wurde die Zeit bis zum Verlust des virologischen Ansprechens ersetzt durch den Zeitpunkt des ersten der verpassten Untersuchungen.

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

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

Tabelle 4-31: Bewertung des Verzerrungspotenzials für Zeit bis Verlust des virologischen Ansprechens 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
S103	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Zeit bis zum Verlust des virologischen Ansprechens ist, für Patienten, die ein virologisches Ansprechen (Abfall der Viruslast unter die Nachweisgrenze von 50 Kopien/mL) erreicht haben, operationalisiert als Zeit bis zum Eintritt des Todes oder des erstmaligen Abbruchs der Studienmedikation, Wechsels zu einem neuen Therapieregime, bestätigter HIV-1-RNA-Wert ≥ 50 Kopien/mL oder HIV-1-RNA-Wert ≥ 50 Kopien/mL gefolgt von einem dauerhaften Abbruch der Studienmedikation. Für Patienten, deren Viruslast nie unter die Nachweisgrenze gesenkt werden konnte, wird die Zeit bis zum Verlust des virologischen Ansprechens auf Studientag 1 gesetzt.

Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt. Es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Faktoren. Das Verzerrungspotenzial wird als niedrig eingestuft.

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 4-32: Ergebnisse für Zeit bis Verlust des virologischen Ansprechens aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set)

Studie S103	Patienten unter Risiko	Patienten mit Ereignis	Kaplan-Meier Schätzer (%) [95%-KI]	Behandlungseffekt
				HR ^a [95%-KI] ^b , p-Wert ^c
Woche 48				
EVG/COBI/FTC/TDF	309	49	16 [11,1;20,3]	0,87 [0,59;1,27], p=0,48
ATV/r+FTC/TDF	306	54	17 [12,6;22,0]	
Woche 96				
EVG/COBI/FTC/TDF	290	66	21 [15,8;26,0]	0,81 [0,58;1,12], p=0,21
ATV/r+FTC/TDF	281	80	24 [19,2;29,4]	

^a Cox Proportional Hazards Modell, stratifiziert nach Ausgangsviruslast (≤ 100.000 vs. > 100.000 Kopien/mL)

Studie S103	Patienten unter Risiko	Patienten mit Ereignis	Kaplan-Meier Schätzer (%) [95%-KI]	Behandlungseffekt
				HR ^a [95%-KI] ^b , p-Wert ^c
^b Wald-Konfidenzgrenzen				
^c Log-Rank Test, stratifiziert nach Ausgangsviruslast (≤ 100.000 vs. > 100.000 Kopien/mL)				
KI = Konfidenzintervall				

Zu Woche 48 standen 309 Patienten im EVG/COBI/FTC/TDF-Arm unter Risiko, wovon bei 49 Patienten ein Ereignis eintrat. Der Kaplan-Meier Schätzer [95%-KI] beträgt 16% [11,1%;20,3%]. Im ATV/r+FTC/TDF-Arm standen zu Woche 48 noch 306 Patienten unter Risiko, wovon bei 54 Patienten ein Ereignis eintrat. Der zugehörige Kaplan-Meier Schätzer [95%-KI] ergibt sich zu 17% [12,6%;22,0%].

Das HR [95%-KI] zu Woche 48 ergibt sich zu 0,87 [0,59;1,27] mit $p=0,47$, was keinen statistisch signifikanten Unterschied zwischen den Behandlungen anzeigt.

Zu Woche 96 standen im EVG/COBI/FTC/TDF-Arm 290 Patienten unter Risiko, wovon bei 66 Patienten ein Ereignis eintrat. Der Kaplan-Meier Schätzer beträgt dabei 21% [15,8%;26,0%]. Im ATV/r+FTC/TDF-Arm standen zu Woche 96 281 Patienten unter Risiko, wovon bei 80 Patienten ein Ereignis eintrat. Der daraus resultierende Kaplan-Meier Schätzer ist 24% [19,2%;29,4%].

Auch nach 96 Wochen Behandlungen zeigt das HR [95%-KI] von 0,81 [0,58;1,12] mit $p=0,20$ keinen statistisch signifikanten Unterschied zwischen den Behandlungen.

In Abbildung 10 ist die Kaplan-Meier Kurve für die Zeit bis zum Verlust des virologischen Ansprechens abgebildet. Gleich ab Anfang der Behandlung ist die EVG/COBI/FTC/TDF (Quad)-Kurve deutlich flacher als die ATV/r+FTC/TDF-Kurve. Zwar nähert sie sich mit der Zeit etwas an die ATV/r+FTC/TDF Kurve an, bleibt jedoch stets flacher. Ab ca. Woche 72 wird die EVG/COBI/FTC/TDF-Kurve wieder etwas flacher, im Vergleich zur ATV/r+FTC/TDF-Kurve.

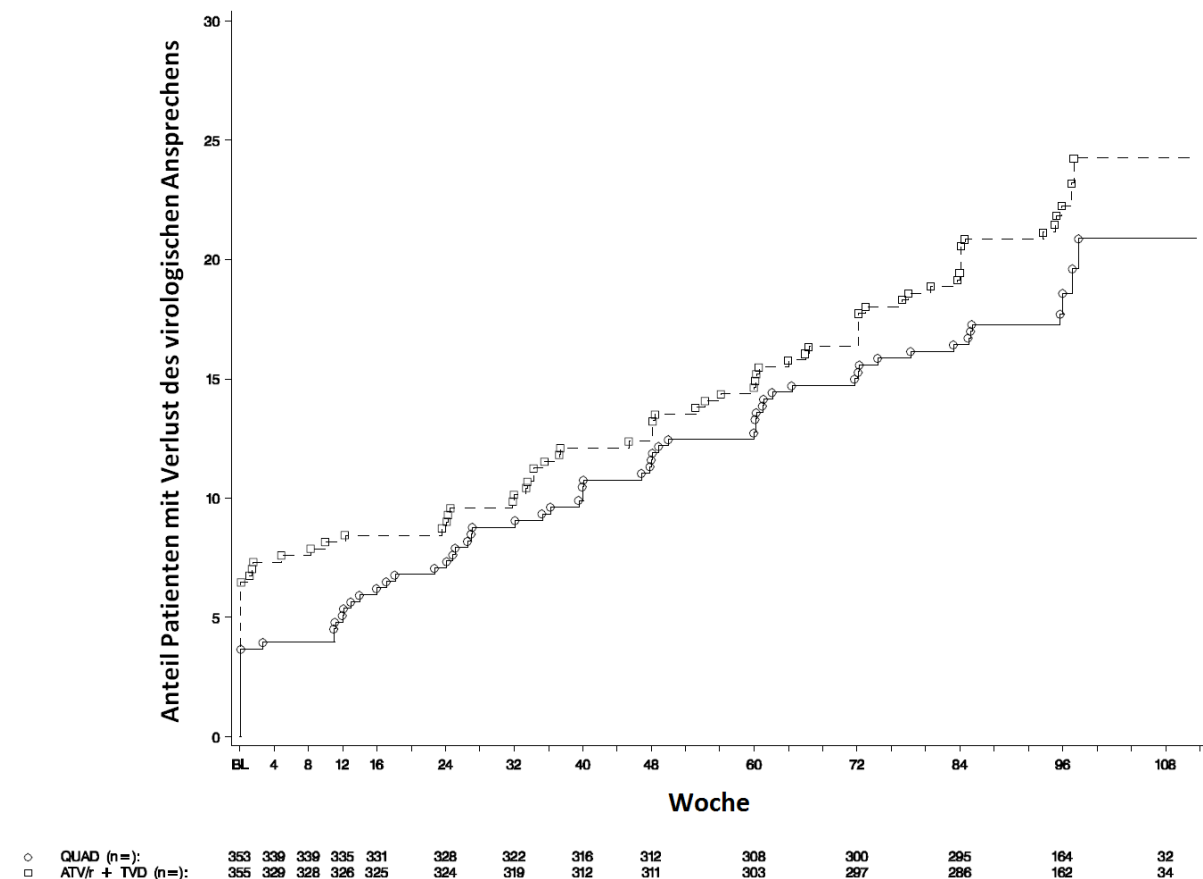


Abbildung 10: Kaplan-Meier Kurven für Zeit bis zum Verlust des virologischen Ansprechens, aus RCT; EVG/COBI/FTC/TDF versus ATV/r+FTC/TDF

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.

Da hier nur eine Studie in die Nutzenbewertung von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF eingeschlossen wurde, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.1.3 Veränderung der CD4-Zellzahl– RCT

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

Tabelle 4-33: Operationalisierung von Veränderung der CD4-Zellzahl

Studie	Operationalisierung
S103	Mittlere Veränderung der CD4-Zellzahl ab Baseline bis Woche 48 bzw. 96. Die CD4-Zellzahl wurde mittels Durchflusszytometrie erfasst.

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 4-34: Bewertung des Verzerrungspotenzials für Veränderung der CD4-Zellzahl 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
S103	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der Endpunkt Veränderung der CD4-Zellzahl ist operationalisiert als Veränderung der mittleren Anzahl der CD4-Zellzahl zwischen Baseline und Woche 48 und Woche 96. Die CD4-Zellzahl wurde an beiden Zeitpunkten mittels Durchflusszytometrie gemessen. Es werden die observed values betrachtet. Der Nichtberücksichtigungsanteil im EVG/COBI/FTC/TDF-Arm lag zu Woche 48 bei ca. 5,4% und im ATV/r+FTC/TDF-Arm lag dieser Anteil bei 9,6%. Zu Woche 96 lag der Nichtberücksichtigungsanteil im EVG/COBI/FTC/TDF-Arm bei 10,5% und im ATV/r+FTC/TDF-Arm bei 11,3%. Allerdings sind die Studienabbrecher detailliert beschrieben (siehe Tabelle 4-182), so dass eine Abschätzung des Einflusses möglich ist. Betrachtet werden die observed values.

Die Endpunkterheber waren verblindet. Es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Punkte. Das Verzerrungspotenzial wird somit als niedrig eingestuft.

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 4-35: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set)

Studie S103 Zeitpunkt Behandlung	Ausgangswert		Endwert		Veränderung		Behandlungseffekt
	N	MW (SD)	N	MW (SD)	N	MW (SD)	Adj. ^a MWD [95%-KI] p-Wert ^a
Woche 48							
EVG/COBI/ FTC/TDF	353	364 (180,6)	334	568 (230,0)	334	207 (164,2)	-6 [-31;18], p=0,61
ATV/r+FTC/ TDF	355	375 (158,9)	321	585 (210,9)	321	211 (160,3)	
Woche 96							
EVG/COBI/ FTC/TDF	353	364 (180,6)	316	615 (227,4)	316	256 (167,1)	-8 [-35;19], p=0,56
ATV/r+FTC/ TDF	355	375 (158,9)	315	636 (223,9)	315	261 (188,0)	
^a aus ANOVA adjustiert für Baseline HIV-1-RNA Level							
Adj. MWD = Adjustierte Mittelwertdifferenz, KI = Konfidenzintervall, MW = Mittelwert, N: Anzahl Patienten in der Auswertung, SD = Standardabweichung							

Die Patienten im EVG/COBI/FTC/TDF-Arm wiesen zu Baseline im Durchschnitt eine CD4-Zellzahl von 364 Zellen/ μ L (SD: 180,6) auf. Diese stieg bis Woche 48 um durchschnittlich 207 Zellen/ μ L (SD: 164,2) auf 568 Zellen/ μ L (SD: 230,0). Im ATV/r+FTC/TDF-Arm lag die durchschnittliche CD4-Zellzahl bei 375 Zellen/ μ L (SD: 158,9). Bis Woche 48 erhöhte sich diese im Mittel um 211 Zellen/ μ L mit einer Standardabweichung von 160,3 auf 585 Zellen/ μ L (SD: 210,9). Der Unterschied zwischen den beiden Behandlungen zu Woche 48 ist nicht statistisch signifikant (adj. MWD [95%-KI]: -6 [-31;18], p=0,61). Bis Woche 96 stieg die CD4-Zellzahl der Patienten im EVG/COBI/FTC/TDF-Arm im Mittel um 256 Zellen/ μ L (SD: 167,1) auf 615 Zellen/ μ L (SD: 227,4) an. Im ATV/r+FTC/TDF-Arm erhöhte sich die durchschnittliche CD4-Zellzahl um 261 Zellen/ μ L (SD: 188,0) auf 636 Zellen/ μ L (SD: 223,9). Auch hier besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (adj. MWD [95%-KI]: -8 [-35;19], p=0,56).

Tabelle 4-36: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set)

Studie S103 Zeitpunkt Behandlung	Ausgangswert		Endwert		Veränderung		Behandlungseffekt
	N	MW (SD)	N	MW (SD)	N	MW (SD)	Adj. ^a MWD [95%-KI] p-Wert ^a
Woche 48							
EVG/COBI/FTC/ TDF	318	359 (169,8)	315	575 (230,1)	315	214 (162,3)	0 [-25;24], p=0,97
ATV/r+FTC/TDF	310	373 (158,9)	308	583 (207,5)	308	212 (156,9)	
Woche 96							
EVG/COBI/FTC/ TDF	297	359 (169,8)	293	619 (223,2)	293	256 (163,9)	-6 [-34;21], p=0,65
ATV/r+FTC/TDF	294	373 (155,0)	293	632 (215,9)	293	260 (181,3)	
^a aus ANOVA adjustiert für Baseline HIV-1-RNA Level							
Adj. MWD = Adjustierte Mittelwertdifferenz, KI = Konfidenzintervall, MW = Mittelwert, N: Anzahl Patienten in der Auswertung, SD = Standardabweichung							

Die Patienten im EVG/COBI/FTC/TDF-Arm der PP-Population wiesen zu Baseline im Durchschnitt eine CD4-Zellzahl von 359 Zellen/ μ L (SD: 169,8) auf. Diese stieg bis Woche 48 um durchschnittlich 214 Zellen/ μ L (SD: 162,3) auf 575 Zellen/ μ L (SD: 230,1) an. Im ATV/r+FTC/TDF-Arm lag die durchschnittliche CD4-Zellzahl bei 373 Zellen/ μ L (SD: 158,9). Bis Woche 48 erhöhte sich diese im Mittel um 212 Zellen/ μ L mit einer Standardabweichung von 156,9 auf 583 Zellen/ μ L (SD: 207,5). Der Unterschied zwischen den beiden Behandlungen zu Woche 48 ist nicht statistisch signifikant (adj. MWD [95%-KI]: 0 [-25;24], p=0,97). Das Woche 96 PP Analyse-Set enthielt 297 EVG/COBI/FTC/TDF-Patienten, die eine CD4-Zellzahl zu Baseline von durchschnittlich 359 Zellen/ μ L (SD: 169,8) aufwiesen, und 294 ATV/r+FTC/TDF-Patienten, deren CD4-Zellzahl zu Baseline bei durchschnittlich 373 Zellen/ μ L (SD: 155,0) lag. Bis Woche 96 stieg die CD4-Zellzahl der Patienten im EVG/COBI/FTC/TDF-Arm im Mittel um 256 Zellen/ μ L (SD: 163,9) auf 619 Zellen/ μ L (SD: 223,2) an. Im ATV/r+FTC/TDF-Arm erhöhte sich die durchschnittliche CD4-Zellzahl um 260 Zellen/ μ L (SD: 181,3) auf 632 Zellen/ μ L (SD: 215,9). Auch hier besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (adj. MWD [95%-KI]: -6 [-34;21], p=0,65).

Die Ergebnisse aus dem PP-Analyse-Set sind somit konsistent zu den ITT-Ergebnissen.

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.

Da hier nur eine Studie in die Nutzenbewertung von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF eingeschlossen wurde, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.1.4 Unerwünschte Ereignisse– RCT

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

Tabelle 4-37: Operationalisierung von Unerwünschte Ereignisse

Studie	Operationalisierung
S103	<p><u>Mortalität</u> Anteil verstorbenen Patienten im Behandlungszeitraum bis Woche 48 und 96 (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage)</p> <p><u>Gesamtrate unerwünschter Ereignisse</u> Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96</p> <p><u>Schwerwiegende unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96</p> <p><u>Unerwünschte Ereignisse Grad 3-4</u> Anteil an Patienten mit mindestens einem unerwünschten Ereignis Grad 3-4 (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96</p> <p><u>Therapieabbrüche aufgrund unerwünschter Ereignisse</u> Anteil an Patienten, die die Studie aufgrund eines unerwünschten Ereignisses abgebrochen haben (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96</p> <p><u>Wichtige unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem wichtigen unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96. Wichtige unerwünschte Ereignisse von Interesse waren Hautausschläge. Zusätzlich werden neurologische & psychiatrische Ereignisse betrachtet.</p> <p><u>Unerwünschte Ereignisse von Interesse</u> Anteil an Patienten mit mindestens einem unerwünschten Ereignis von Interesse (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96. Unerwünschte Ereignisse von Interesse waren renale Ereignisse sowie Knochenfrakturen.</p> <p><u>Häufigste unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem häufigsten unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96. Es wurden die unerwünschten Ereignisse betrachtet, die bei mindestens 10% der Patienten auftraten. Häufigste unerwünschte Ereignisse, welche in der vorliegenden Fragestellung relevant sind, waren Erkrankungen des Gastrointestinaltrakts (SOC), allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), Erkrankungen des Nervensystems (SOC), psychiatrische Erkrankungen (SOC), Erkrankungen der Haut und des Unterhautzellgewebes (SOC) und Augenerkrankungen (SOC). Diese wurden jeweils präspezifiziert als SOC erfasst. Darunter werden die jeweils am häufigsten auftretenden unerwünschten Ereignisse (Auftreten bei $\geq 10\%$ der Patienten in einem der beiden Behandlungsarme) als PTs betrachtet.</p>

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

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

Tabelle 4-38: 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
S103	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In Studie S103 waren Patienten und Endpunkterheber zu jeder Zeit gegenüber der Gruppenzugehörigkeit verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung. Es können keine sonstigen das Verzerrungspotenzial beeinflussende Punkte gefunden werden. Das Verzerrungspotenzial auf Endpunktebene wird hier als niedrig bewertet.

Unter unerwünschten Ereignissen wurden verschiedene sicherheitsbezogene Ereignisse betrachtet. Dies waren die folgenden:

- Anteil verstorbener Patienten (Mortalität)
- Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art (Gesamtrate unerwünschter Ereignisse)
- Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Schwerwiegende unerwünschte Ereignisse)
- Anteil an Patienten mit mindestens einem unerwünschten Ereignis Grad 3-4 (Unerwünschte Ereignisse Grad 3-4)
- Anteil an Patienten, die die Therapie aufgrund unerwünschter Ereignisse vorzeitig abbrechen (Therapieabbrüche aufgrund unerwünschter Ereignisse)

Zusätzlich dazu wurden wichtige unerwünschte Ereignisse betrachtet. Dies sind unerwünschte Ereignisse die besonders in Zusammenhang mit ATV/r+FTC/TDF auftraten und in der Fachinformation von ATV/r+FTC/TDF als übliche Nebenwirkungen dieser Therapie aufgelistet sind. Folgende unerwünschte Ereignisse wurden als wichtig erachtet:

- Anteil an Patienten mit Hautausschlägen (Hautausschläge):

Diese präspezifizierte Gruppe von PT schloss die folgenden Terme ein: Dermatitis, Arzneimittelexanthem, Ekzem, exfoliativer Hautausschlag, Pruritus, generalisierter Pruritus, Hautausschlag, erythematöser Hautausschlag, follikulärer Hautausschlag, generalisierter Hautausschlag, makularer Hautausschlag, maculopapulöser Hautausschlag, morbilliformer Hautausschlag, papularer Hautausschlag, juckender Hautausschlag, pustulöser Hautausschlag, vesikulärer Hautausschlag, seborrhoische Dermatitis, seborrhoische Keratose, Hautreaktionen und Nesselausschlag.

Im Sinne der Vergleichbarkeit werden auch hier neurologische & psychiatrische Ereignisse betrachtet, auch wenn sie nicht als typische Nebenwirkungen von ATV/r+FTC/TDF in der Fachinformation aufgeführt sind. Folgender, nicht präspezifizierter, Endpunkt wird hier somit ebenso betrachtet:

- Anteil an Patienten mit unerwünschten Ereignissen aus dem neurologischen bzw. psychiatrischen Bereich (Neurologische & psychiatrische Ereignisse):
Diese präspezifizierte Gruppe von PT schloss die folgenden Terme ein: abnorme Träume, akute Psychose, affektive Erkrankung, Agitation, Amnesie, Angst, Angststörung, bipolare Störung, vollendeter Suizid, Verwirrheitszustände, depressive Stimmung, Depression, Desorientierung, Aufmerksamkeitsstörungen, Schwindel, euphorische Stimmung, Halluzination, Kopfschmerzen, Schlaflosigkeit, Reizbarkeit, verringerte Libido, starke Depressionen, veränderte Stimmung, Stimmungsschwankungen, Nervosität, Alpträume, Panikattacke, Paranoia, Unruhe, Schlafstörungen, Somnolenz, Suizidgedanken, Suizidversuch und Gleichgewichtsstörung.

Des Weiteren wurden unerwünschte Ereignisse, welche besonders in Zusammenhang mit Tenofoviridisoproxil (TDF), welches in beiden Behandlungsarmen der hier betrachteten Studie gegeben wurde, auftreten, betrachtet. Diese umfassen:

- Anteil an Patienten mit renalen unerwünschten Ereignissen (Renale Ereignisse):
Diese präspezifizierte Gruppe von PT schloss die folgenden Terme ein:
Fanconi-Syndrom, erworbenes Fanconi-Syndrome, Nierenversagen, akutes Nierenversagen und tubuläre Erkrankung.
- Anteil an Patienten mit Knochenfrakturen (Knochenfrakturen):
Dieser MedDRA high level group term (HLGT) „Frakturen“ schloss folgende Terme ein:

Hüftpfannenfraktur, Knöchelbruch, Abrissfraktur, Knochenfissur, Knochenfragmentierung, Halswirbelsäulenfraktur, Klavikula, komplizierter Bruch, Kompressionsfraktur, epiphysealer Bruch, Gesichtsknochenbruch, Schenkelhalsfraktur, Femurfraktur, Fibularfraktur, Flatterbrust, Fußfraktur, Unterarmfraktur, Fraktur, verzögerte Frakturheilung, Bruchverschiebung, Bruchfehlstellung, verzögerte Frakturheilung, Steißbeinfraktur, Sitzbeinfraktur, Kreuzbeinfraktur, Schädelimpressionsfraktur, Grünholzfraktur, Handfraktur, Hüftenfraktur, Humerusfraktur, Darmbeinfraktur, impaktierte Fraktur, Kieferfraktur, Fraktur der unteren Gliedmaßen, Lendenwirbelfraktur, multiple Frakturen, offene Fraktur, osteoporotische Fraktur, Patella Fraktur, pathologische Fraktur, Beckenfraktur, periprothetische Fraktur, Pseudarthrose, Schambeinfraktur, Radiusfraktur, Rippenfraktur, Skapulafraktur, skapulothorakale Dissoziation, Schädelbruch, Schädelbasisfraktur, Wirbelfraktur, sternale Fraktur, Ermüdungsfraktur, Synostose, Brustwirbelfraktur, Schienbeinfraktur, Torus Fraktur, traumatische Fraktur, Ellenfraktur, Fraktur der oberen Gliedmaßen, Handgelenkfraktur

Zu dem HLGTT „Frakturen“ wurde nach der Woche 48-Analyse zusätzlich der Term „Wirbelkompressionsfraktur“ hinzugefügt. Dieser ist also nicht in der Woche 48-Analyse enthalten, sondern erst in der Woche 96-Analyse.

Es werden die in der Studie am häufigsten beobachteten unerwünschten Ereignisse dargestellt. Dabei werden die folgenden SOC als für die vorliegende Fragestellung relevant erachtet:

- Erkrankungen des Gastrointestinaltrakts
- Allgemeine Erkrankungen und Beschwerden am Verabreichungsort
- Erkrankungen des Nervensystems
- Psychiatrische Erkrankungen
- Erkrankungen der Haut und des Unterhautzellgewebes
- Augenerkrankungen

Darunter werden jeweils die am häufigsten auftretenden PT beobachtet. Dies sind die Erkrankungen, die bei mindestens 10% der Patienten in einem der Behandlungsarme auftraten. Allgemein gilt hier, dem Vorgehen des IQWiG folgend [62], dass die einzelnen PT als durch die jeweilige SOC abgedeckt angesehen werden. Sie tragen somit nicht zusätzlich zur jeweiligen SOC zur Nutzenbewertung bei.

Da zu Woche 48 keine PT unter den SOC „Psychiatrische Erkrankungen“ und „Erkrankungen der Haut und des Unterhautzellgewebes“ bei mehr als 10% der Patienten in jeglichem Behandlungsarm auftraten, werden aus Vergleichbarkeitsgründen die PT dargestellt, die in Studie S102 (EVG/COBI/FTC/TDF-EFV/FTC/TDF-Vergleich) am häufigsten auftraten. Gleiches gilt für Erkrankungen der Haut und des Unterhautzellgewebes zu Woche 96.

Mit dem Begriff „Hautausschläge“ wird die Gruppe an PT bezeichnet, die im Rahmen der wichtigen unerwünschten Ereignisse betrachtet wird, wohingegen der Begriff „Hautausschlag“ einen bei mindestens 10% der Patienten auftretenden PT im Rahmen der häufigsten unerwünschten Ereignisse bezeichnet.

Die Endpunkterheber waren verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung. Auch gibt es keine sonstigen, das Verzerrungspotenzial beeinflussenden Punkte. Somit ist das Verzerrungspotenzial als niedrig einzustufen.

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.

Woche 48

Tabelle 4-39: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Studie S103 Endpunkt	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt ^a
	N	n (%)	N	n (%)	OR [95%-KI], p-Wert
Mortalität	353	0 (0,0)	355	3 (0,8)	0,14 [0,01;2,77], p=0,1980
Gesamtrate unerwünschter Ereignisse	353	323 (91,5)	355	333 (93,8)	0,71 [0,40;1,26], p=0,2423
Schwerwiegende unerwünschte Ereignisse	353	26 (7,4)	355	31 (8,7)	0,83 [0,48;1,43], p=0,5043
Therapieabbrüche aufgrund unerwünschter Ereignisse	353	13 (3,7)	355	18 (5,1)	0,72 [0,35;1,48], p=0,3689
Unerwünschte Ereignisse Grad 3-4	353	45 (12,7)	355	48 (13,5)	0,93 [0,60;1,45], p=0,7607
^a zusätzlich berechnet KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Tabelle 4-40: Ergebnisse für wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Studie S103 Endpunkt	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^b
Hautausschläge	353	63 (17,8)	355	64 (18,0)	0,99 [0,67;1,45], p=1,00
Neurologische & psychiatrische Ereignisse ^c	353	122 (34,6)	355	128 (36,1)	0,94 [0,69;1,27], p=0,69
^a zusätzlich berechnet ^b Fisher's exakter Test ^c Quelle: [102] KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Tabelle 4-41: Ergebnisse für unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Studie S103 Endpunkt	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^b
Renale Ereignisse	353	0 (0,0)	355	0 (0,0)	n.b.
Knochenfrakturen	353	3 (0,8)	355	6 (1,7)	0,50 [0,12;2,01], p=0,51

^a zusätzlich berechnet
^b Fisher's exakter Test
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio

Tabelle 4-42: Ergebnisse für die häufigsten unerwünschten Ereignisse (SOC, Grenze: ≥10%) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Studie S103 Endpunkt (SOC) Preferred term	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^a
Erkrankungen des Gastrointestinaltrakts (SOC)	353	189 (53,5)	355	201 (56,6)	0,88 [0,66;1,19], p=0,4103
<i>Diarrhoe</i>	353	77 (21,8)	355	97 (27,3)	0,74 [0,53;1,05], p=0,0891
<i>Übelkeit</i>	353	70 (19,8)	355	69 (19,4)	1,03 [0,71;1,49], p=0,8952
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)	353	96 (27,2)	355	94 (26,5)	1,04 [0,74;1,45], p=0,8297
<i>Fatigue</i>	353	50 (14,2)	355	45 (12,7)	1,14 [0,74;1,75], p=0,5615
Erkrankungen des Nervensystems (SOC)	353	91 (25,8)	355	93 (26,2)	0,98 [0,70;1,37], p=0,8991
<i>Kopfschmerzen</i>	353	53 (15,0)	355	44 (12,4)	1,25 [0,81;1,92], p=0,3114
Psychiatrische Erkrankungen ^b (SOC)	353	75 (21,2)	355	81 (22,8)	0,91 [0,64;1,30], p=0,6143
<i>Abnorme Träume</i>	353	12 (3,4)	355	14 (3,9)	0,86 [0,39;1,88], p=0,7005
<i>Schlaflosigkeit</i>	353	29 (8,2)	355	18 (5,1)	1,68 [0,91;3,08], p=0,0958
<i>Depression</i>	353	20 (5,7)	355	23 (6,5)	0,87 [0,47;1,61], p=0,6508
Erkrankungen der Haut und des Unterhautzellgewebes ^b (SOC)	353	98 (27,8)	355	102 (28,7)	0,95 [0,69;1,32], p=0,7743
<i>Hautausschlag</i>	353	25 (7,1)	355	22 (6,2)	1,15 [0,64;2,09], p=0,6365
Augenerkrankungen (SOC)	353	18 (5,1)	355	78 (22,0)	0,19 [0,11;0,33], p<0,00001
<i>Skleraler Ikterus</i>	353	2 (0,6)	355	51 (14,4)	0,03 [0,01;0,14], p<0,00001

^a zusätzlich berechnet
^b Da kein preferred term unter dieser SOC bei mehr als 10% der Patienten auftrat, werden die PTs aus dem EVG/COBI/FTC/TDF-EFV/FTC/TDF-Vergleich zu Woche 48 aus Vergleichsgründen hier dargestellt
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio

Unerwünschte Ereignisse

Kein Patient im EVG/COBI/FTC/TDF-Arm verstarb während der ersten 48 Wochen Behandlung, während im ATV/r+FTC/TDF-Arm drei Patienten (0,8%) der Patienten bis Woche 48 verstarben (OR [95%-KI]: 0,14 [0,01;2,77], p=0,1980).

91,5% der Patienten im EVG/COBI/FTC/TDF-Arm und 93,8% der Patienten im ATV/r+FTC/TDF-Arm hatten ein unerwünschtes Ereignis. Dieser Unterschied ist nicht statistisch signifikant (OR [95%-KI]: 0,71 [0,40;1,26], p=0,2423).

Schwerwiegende unerwünschte Ereignisse traten bei 7,4% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 8,7% der Patienten im ATV/r+FTC/TDF-Arm auf (OR [95%-KI]: 0,83 [0,48;1,43], p=0,5043).

3,7% der Patienten im EVG/COBI/FTC/TDF-Arm und 5,1% der Patienten im ATV/r+FTC/TDF-Arm brachen die Therapie aufgrund unerwünschter Ereignisse vorzeitig ab. Hier besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,72 [0,35;1,48], p=0,3689).

Auch hinsichtlich unerwünschter Ereignisse Grad 3-4 besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (EVG/COBI/FTC/TDF: 12,7%; ATV/r+FTC/TDF: 13,5%; OR [95%-KI]: 0,93 [0,60;1,45], p=0,7607).

Wichtige unerwünschte Ereignisse

Bei 17,8% der Patienten im EVG/COBI/FTC/TDF-Arm und 18,0% der Patienten im ATV/r+FTC/TDF-Arm wurden Hautausschläge beobachtet. Somit besteht hier kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,99 [0,67;1,45], p=1,00).

Hinsichtlich neurologischer & psychiatrischer Erkrankungen konnte kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen beobachtet werden (EVG/COBI/FTC/TDF: 34,6%, ATV/r+FTC/TDF: 36,1%). Das OR [95%-KI] beträgt 0,94 [0,69;1,27] mit p=0,69.

Unerwünschte Ereignisse von Interesse

In der Studie wurden bis Woche 48 keine renalen Ereignisse beobachtet. 0,8% der EVG/COBI/FTC/TDF-Patienten und 1,7% der ATV/r+FTC/TDF-Patienten wiesen bis Woche 48 Knochenfrakturen auf. Allerdings ist dieser Unterschied nicht statistisch signifikant (OR [95%-KI]: 0,50 [0,12;2,01], p=0,51).

Häufigste unerwünschte Ereignisse

Erkrankungen des Gastrointestinaltrakts (SOC) traten bei 53,5% der Patienten im EVG/COBI/FTC/TDF-Arm und 56,6% der Patienten im ATV/r+FTC/TDF-Arm auf. Der Unterschied war nicht statistisch signifikant (OR [95%-KI]: 0,88 [0,66;1,19], p=0,4103). Die am häufigsten darunter beobachteten Ereignisse waren Diarrhoe (EVG/COBI/FTC/TDF: 21,8%; ATV/r+FTC/TDF: 27,3%; OR [95%-KI]: 0,74 [0,53;1,05], p=0,0891) und Übelkeit

(EVG/COBI/FTC/TDF: 19,8%; ATV/r+FTC/TDF: 19,4%; OR [95%-KI]: 1,03 [0,71;1,49], $p=0,8952$). Hier bestehen allerdings keine statistisch signifikanten Unterschiede zwischen den Behandlungen.

Bei 27,2% der EVG/COBI/FTC/TDF-Patienten und 26,5% der ATV/r+FTC/TDF-Patienten wurden *allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)* beobachtet. Es besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,04 [0,74;1,45], $p=0,8297$). Das häufigste in dieser Kategorie beobachtete Ereignis war Fatigue, welches bei 14,2% der EVG/COBI/FTC/TDF-Patienten und 12,7% der ATV/r+FTC/TDF-Patienten beobachtet wurde (OR [95%-KI]: 1,14 [0,74;1,75], $p=0,5615$).

Erkrankungen des Nervensystems (SOC) traten bei 25,8% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 12,4% der Patienten im ATV/r+FTC/TDF-Arm auf. Es kann jedoch kein statistisch signifikanter Unterschied zwischen den Behandlungen gefunden werden (OR [95%-KI]: 0,98 [0,70;1,37], $p=0,8991$). Häufigstes Ereignis hier waren Kopfschmerzen, die bei 15,0% der EVG/COBI/FTC/TDF-Patienten und 12,4% der ATV/r+FTC/TDF-Patienten beobachtet wurden (OR [95%-KI]: 1,25 [0,81;1,92], $p=0,3114$).

Insgesamt gab es im EVG/COBI/FTC/TDF-Arm 75 (21,2%) Patienten mit *psychiatrischen Erkrankungen (SOC)*. Im ATV/r+FTC/TDF-Arm waren dies 81 (22,8%) Patienten. Das OR [95%-KI] mit 0,91 [0,64;1,30] und $p=0,6143$ zeigt jedoch keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen. 12 (3,4%) Patienten im EVG/COBI/FTC/TDF-Arm und 14 (3,9%) Patienten im ATV/r+FTC/TDF-Arm hatten abnorme Träume (OR [95%-KI]: 0,86 [0,39;1,88], $p=0,7005$) und 29 (8,2%) EVG/COBI/FTC/TDF-Patienten und 18 (5,1%) ATV/r+FTC/TDF-Patienten litten unter Schlaflosigkeit (OR [95%-KI]: 1,68 [0,91;3,08], $p=0,0958$). Bei 20 (5,7%) Patienten im EVG/COBI/FTC/TDF-Arm und 23 (6,5%) Patienten im ATV/r+FTC/TDF-Arm wurden Depressionen beobachtet. Der Unterschied zwischen den beiden Behandlungen ist nicht statistisch signifikant (OR [95%-KI]: 0,87 [0,47;1,61], $p=0,6508$).

Erkrankungen der Haut und des Unterhautzellgewebes (SOC) traten bei 98 (27,8%) Patienten im EVG/COBI/FTC/TDF-Arm und bei 102 (28,7%) Patienten im ATV/r+FTC/TDF-Arm auf (OR [95%-KI]: 0,95 [0,69;1,32], $p=0,7743$). Hautausschläge wurden bei 25 (7,1%) Patienten im EVG/COBI/FTC/TDF-Arm und bei 22 (6,2%) Patienten im ATV/r+FTC/TDF-Arm auf. Das OR [95%-KI] ergibt sich zu 1,15 [0,64;2,09] mit $p=0,6365$, was keinen statistisch signifikanten Unterschied zeigt.

Bei 5,1% der EVG/COBI/FTC/TDF-Patienten und 22,0% der ATV/r+FTC/TDF-Patienten wurden *Augenerkrankungen (SOC)* beobachtet. Dieser Unterschied ist statistisch signifikant zugunsten der EVG/COBI/FTC/TDF-Behandlung. Das OR [95%-KI] ergibt sich zu 0,19 [0,11;0,33] mit $p<0,00001$. Die häufigste Augenerkrankung war skleraler Ikterus, der bei 0,6% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 14,4% der Patienten der ATV/r+FTC/TDF-Patienten auftrat. Auch hier besteht ein statistisch signifikanter Vorteil der

EVG/COBI/FTC/TDF-Behandlung gegenüber der ATV/r+FTC/TDF-Behandlung (OR [95%-KI]: 0,03 [0,01;0,14], $p < 0,00001$).

Woche 96

Tabelle 4-43: Ergebnisse für Unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)

Studie S103 Endpunkt	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^a
Mortalität	353	0 (0,0)	355	3 (0,8)	0,14 [0,01;2,77], $p=0,1980$
Gesamtrate unerwünschter Ereignisse	353	336 (95,2)	355	343 (96,6)	0,69 [0,33;1,47], $p=0,3376$
Schwerwiegende unerwünschte Ereignisse	353	34 (9,6)	355	50 (14,1)	0,65 [0,41;1,03], $p=0,0684$
Therapieabbrüche aufgrund unerwünschter Ereignisse	353	15 (4,2)	355	21 (5,9)	0,71 [0,36;1,39], $p=0,3151$
Unerwünschte Ereignisse Grad 3-4	353	55 (15,6)	355	61 (17,2)	0,89 [0,60;1,33], $p=0,5648$

^a zusätzlich berechnet
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio

Tabelle 4-44: Ergebnisse für Wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)

Studie S103 Endpunkt	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^b
Hautausschläge	353	76 (21,5)	355	81 (22,8)	0,93 [0,65;1,32], $p=0,72$
Neurologische & psychiatrische Erkrankungen	353	146 (41,4)	355	151 (42,5)	0,95 [0,71;1,28], $p=0,76$

^a zusätzlich berechnet
^b Fisher's exakter Test
^c Quelle [103]
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio

Tabelle 4-45: Ergebnisse für Unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)

Studie S103 Endpunkt	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^b
Renale Ereignisse	353	4 (1,1)	355	0 (0,0)	9,15 [0,49;170,67], $p=0,061$
Knochenfrakturen	353	4 (1,1)	355	14 (3,9)	0,28 [0,09;0,86], $p=0,029$

^a zusätzlich berechnet
^b Fisher's exakter Test
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio

Tabelle 4-46: Ergebnisse für die häufigsten unerwünschten Ereignisse (SOC, Grenze: $\geq 10\%$) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)

Studie S103 Endpunkt (SOC) Preferred term	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^a
Erkrankungen des Gastrointestinaltrakts	353	205 (58,1)	355	226 (63,7)	0,79 [0,58;1,07], p=0,1279
<i>Diarrhoe</i>	353	88 (24,9)	355	111 (31,3)	0,73 [0,53;1,01], p=0,0611
<i>Übelkeit</i>	353	74 (21,0)	355	75 (21,1)	0,99 [0,69;1,42], p=0,9574
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	353	102 (28,9)	355	119 (33,5)	0,81 [0,59;1,11], p=0,1844
<i>Fatigue</i>	353	53 (15,0)	355	57 (16,1)	0,92 [0,61;1,39], p=0,7020
Erkrankungen des Nervensystems	353	112 (31,7)	355	120 (33,8)	0,91 [0,66;1,25], p=0,5565
<i>Kopfschmerzen</i>	353	59 (16,7)	355	53 (14,9)	1,14 [0,76;1,71], p=0,5156
Psychiatrische Erkrankungen	353	101 (28,6)	355	109 (30,7)	0,90 [0,66;1,25], p=0,5423
<i>Depression</i>	353	34 (9,6)	355	41 (11,5)	0,82 [0,50;1,32], p=0,4077
Erkrankungen der Haut und des Unterhautzellgewebes ^b	353	117 (33,1)	355	134 (37,7)	0,82 [0,60;1,11], p=0,2008
<i>Hautausschlag</i>	353	29 (8,2)	355	31 (8,7)	0,94 [0,55;1,59], p=0,8049
Augenerkrankungen	353	28 (7,9)	355	85 (23,9)	0,27 [0,17;0,43], p<0,00001
<i>Skleraler Ikterus</i>	353	2 (0,6)	355	51 (14,4)	0,03 [0,01;0,14], p<0,00001
^a zusätzlich berechnet					
^b Da kein preferred term unter dieser SOC bei mehr als 10% der Patienten auftrat, werden die PTs aus dem EVG/COBI/FTC/TDF-EFV/FTC/TDF-Vergleich zu Woche 48 hier dargestellt					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Unerwünschte Ereignisse

Zwischen Woche 48 und Woche 96 traten in keiner der beiden Behandlungsgruppen neue Todesfälle auf (OR [95%-KI]: 0,14 [0,01;2,77], p=0,1980). 95,2% der Patienten im EVG/COBI/FTC/TDF-Arm und 96,6% der Patienten im ATV/r+FTC/TDF hatten mindestens ein unerwünschtes Ereignis. Hier besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,69 [0,33;1,47], p=0,3376). Schwerwiegende unerwünschte Ereignisse traten bei 9,6% der EVG/COBI/FTC/TDF-Patienten und 14,1% der ATV/r+FTC/TDF-Patienten auf. Mit einem OR [95%-KI] von 0,65 [0,41;1,03] und p=0,0684 besteht hier also kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen. 4,2% der Patienten im EVG/COBI/FTC/TDF-Arm und 5,9% der Patienten im ATV/r+FTC/TDF-Arm brachen die Therapie aufgrund unerwünschter Ereignisse vorzeitig ab (OR [95%-KI]: 0,71 [0,36;1,39], p=0,3151) und 15,6% der EVG/COBI/FTC/TDF-Patienten

und 17,2% der ATV/r+FTC/TDF-Patienten erlebten unerwünschte Ereignisse Grad 3-4 (OR [95%-KI]: 0,89 [0,60;1,33], p=0,5648).

Wichtige unerwünschte Ereignisse

Bei 21,5% der EVG/COBI/FTC/TDF-Patienten und 22,8% der ATV/r+FTC/TDF-Patienten traten Hautausschläge auf, wobei hier kein statistisch signifikanter Unterschied besteht (OR [95%-KI]: 0,93 [0,65;1,32], p=0,72).

41,4% der Patienten im EVG/COBI/FTC/TDF-Arm und 42,5% der Patienten im ATV/r+FTC/TDF-Arm zeigten neurologische & psychiatrische Erkrankungen. Es besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,95 [0,71;1,28], p=0,76).

Unerwünschte Ereignisse von Interesse

Nach 96 Wochen traten im EVG/COBI/FTC/TDF-Arm bei 1,1% der Patienten renale Ereignisse auf, während im ATV/r+FTC/TDF-Arm kein Patient ein renales Ereignis erlebte (OR [95%-KI]: 9,15 [0,49;170,67], p=0,061). Knochenfrakturen wurden bei 1,1% der EVG/COBI/FTC/TDF-Patienten und 3,9% der ATV/r+FTC/TDF-Patienten beobachtet. Dieser Unterschied ist mit einem OR [95%-KI] von 0,28 [0,09;0,86] und p=0,029 statistisch signifikant zugunsten der EVG/COBI/FTC/TDF-Behandlung.

Pathologische Frakturen stellen die symptomatische Endstrecke Knochenmineraldichtereduzierender Zustandsbilder wie Osteopenie und Osteoporose dar. Hierbei handelt es je nach Genese um eine Erhöhung osteoklastischer Aktivität oder einer Verminderung osteoblastischer Aktivität mit entsprechender mechanisch-inadäquater Mineralfinstruktur des Knochens. Eine initiale leichtgradige Reduktion der Knochenmineraldichte mit Stabilisierung nach 24 oder 48 Wochen, je nach Messpunkt ist nach Einleitung einer antiretroviralen HIV Therapie für nahezu alle Therapieregime auch in kontrollierten klinischen Studien beschrieben. Ursächlich hierfür werden immunrekonstitutive Prozesse im Rahmen des Wiederanstiegs der CD4-Zellzahl beschrieben. Die beschriebenen Reduktionen bewegen sich im niedrigen einstelligen Prozentbereich und bewegen sich nicht in Dimensionen, in denen pathologische Frakturen zum Beispiel im Rahmen der postmenstruellen Osteoporose oder maligner Zustandsbilder regelmäßig beschrieben werden.

Der beschriebene Unterschied zwischen EVG/COBI/FTC/TDF und ATV/r+FTC/TDF nach 96 Wochen ist weder mit der o. g. Pathogenese noch aufgrund des kurzen Beobachtungszeitraumes schlüssig erklärbar. Eine für Frakturen ursächliche Knochendichtereduktion über 96 Wochen hätte zudem einen erheblichen Anstieg entsprechender Serumparameter, wie zum Beispiel der alkalischen Phosphatase und oder der Lactat-Dehydrogenase verursachen müssen, der in der Studie aber nicht beobachtet werden konnte.

Häufigste unerwünschte Ereignisse

Im EVG/COBI/FTC/TDF-Arm wurde bei 58,1% der Patienten *Erkrankungen des Gastrointestinaltrakts (SOC)* beobachtet. Im ATV/r+FTC/TDF-Arm waren dies 63,7% der

Patienten. Dieser Unterschied ist nicht statistisch signifikant (OR [95%-KI]: 0,79 [0,58;1,07], $p=0,1279$). Die häufigsten Erkrankungen des Gastrointestinaltrakts waren Diarrhoe (EVG/COBI/FTC/TDF: 24,9%; ATV/r+FTC/TDF: 31,3%; OR [95%-KI]: 0,73 [0,53;1,01], $p=0,0611$) und Übelkeit (EVG/COBI/FTC/TDF: 21,0%; ATV/r+FTC/TDF: 21,2%; OR [95%-KI]: 0,99 [0,69;1,42], $p=0,9574$).

Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) wurden bei 28,9% der EVG/COBI/FTC/TDF-Patienten und 33,5% der ATV/r+FTC/TDF-Patienten beobachtet. Es besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,81 [0,59;1,11], $p=0,1844$). Häufigstes Ereignis darunter war Fatigue, was bei 15,0% der EVG/COBI/FTC/TDF-Patienten und 16,1% der ATV/r+FTC/TDF-Patienten auftrat (OR [95%-KI]: 0,92 [0,61;1,39], $p=0,7020$).

31,7% der EVG/COBI/FTC/TDF-Patienten und 33,8% der ATV/r+FTC/TDF-Patienten hatten *Erkrankungen des Nervensystems (SOC)*. Mit einem OR [95%-KI] von 0,91 [0,66;1,25] und $p=0,5565$ besteht hier also kein statistisch signifikanter Unterschied. Am häufigsten darunter traten Kopfschmerzen auf, die bei 16,7% der EVG/COBI/FTC/TDF- und 14,9% der ATV/r+FTC/TDF-Patienten beobachtet wurden (OR [95%-KI]: 1,14 [0,76;1,71], $p=0,5156$).

Bis Woche 96 erlebten 28,6% der EVG/COBI/FTC/TDF-Patienten und 30,7% der ATV/r+FTC/TDF-Patienten *psychiatrische Erkrankungen (SOC)*, wobei hier kein statistisch signifikanter Unterschied besteht (OR [95%-KI]: 0,90 [0,66;1,25], $p=0,5423$). Am häufigsten trat hier das Ereignis Depression auf (EVG/COBI/FTC/TDF: 9,6%; ATV/r+FTC/TDF: 11,5%). Mit einem OR [95%-KI] von 0,82 [0,50;1,32] und $p=0,4077$ ist allerdings auch dieser Unterschied nicht statistisch signifikant.

Erkrankungen der Haut und des Unterhautzellgewebes (SOC) traten bei 117 (33,1%) Patienten im EVG/COBI/FTC/TDF-Arm und bei 134 (37,7%) Patienten im ATV/r+FTC/TDF-Arm auf. Der Unterschied ist jedoch nicht statistisch signifikant (OR [95%-KI]: 0,82 [0,60;1,11], $p=0,2008$). 29 (8,2%) Patienten im EVG/COBI/FTC/TDF-Arm und 31 (8,7%) Patienten im ATV/r+FTC/TDF-Arm hatten Hautausschläge. Allerdings besteht auch hier kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,94 [0,55;1,59], $p=0,8049$).

Augenerkrankungen (SOC) traten bei 7,9% der Patienten im EVG/COBI/FTC/TDF-Arm und 23,9% der ATV/r+FTC/TDF-Patienten auf. Das OR [95%-KI] ergibt sich hier zu 0,27 [0,17;0,43] mit $p<0,00001$. Im EVG/COBI/FTC/TDF-Arm sind also statistisch signifikant weniger Augenerkrankungen aufgetreten als im ATV/r+FTC/TDF-Arm. Die häufigste Augenerkrankung war skleraler Ikterus, der bei 0,6% der EVG/COBI/FTC/TDF-Patienten und 14,4% der ATV/r+FTC/TDF-Patienten beobachtet wurde. Auch hier besteht ein statistisch signifikanter Vorteil der EVG/COBI/FTC/TDF-Behandlung gegenüber der ATV/r+FTC/TDF-Behandlung (OR [95%-KI]: 0,03 [0,01;0,14], $p<0,00001$).

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.

Da hier nur eine Studie in die Nutzenbewertung von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF eingeschlossen wurde, wurde keine Meta-Analyse durchgeführt.

4.3.1.3.2 Subgruppenanalysen – RCT

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

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

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

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

A priori waren Subgruppenanalysen in Studie S103 nur für den primären Endpunkt virologisches Ansprechen (Snapshot) geplant. Die Subgruppenanalysen wurden für folgende Merkmale durchgeführt:

- Alter (<40 Jahre, ≥40 Jahre)
- Geschlecht (männlich, weiblich)
- Rasse (kaukasisch, nicht kaukasisch)
- Ausgangsviruslast (≤100.000 Kopien/mL, >100.000 Kopien/mL)
- CD4-Zellzahl zu Baseline (≤350 Zellen/μL, >350 Zellen/μL)
- Therapieadhärenz (<95%, ≥95%)

Zusätzlich dazu werden post-hoc Subgruppenanalysen nach Alter, Geschlecht, Rasse, Ausgangsviruslast, CD4-Zellzahl zu Baseline und Therapieadhärenz für alle in diese Nutzenbewertung eingeschlossenen Endpunkte durchgeführt [98-101]. Somit werden im Folgenden alle a priori spezifizierten Subgruppenanalysen für den primären Endpunkt und alle post-hoc durchgeführten Subgruppenanalysen für alle weiteren Endpunkte dargestellt.

Da dies die präspezifizierten Subgruppenvariablen sind, werden die Subgruppenanalysen auch im vorliegenden Dossier zur Nutzenbewertung stratifiziert nach diesen Variablen durchgeführt.

Bei der Interpretation und vor allem bei Nutzenableitung ist die Korrelation zwischen Ausgangsviruslast und CD4-Zellzahl zu Baseline zu berücksichtigen.

Subgruppenanalysen werden nur für das jeweilige primäre Analyse-Set durchgeführt.

Da die Ausprägung der Therapieadhärenz naturgemäß nicht zu Baseline, sondern erst im Nachgang erhoben wurde, ist dies keine Subgruppenanalyse im üblichen Sinne. Auch eine Stratifizierung ist bei Adhärenz nicht möglich, da sich die Subgruppen erst durch die Datenerhebung ergeben. Dennoch wird sie hier aus formalen Gründen dargestellt, um eine Interaktion zwischen der Behandlung und der Therapieadhärenz nach 48 bzw. 96 Wochen zu untersuchen. Da jedoch die Behandlung selbst und möglicherweise auch andere Faktoren einen Einfluss auf die Subgruppenbildung hatten, ist die Aussagekraft dieser Subgruppenanalyse als niedrig einzuschätzen. Darüber hinaus sind die in der vorliegenden Nutzenbewertung betrachteten Daten nicht praxisrelevant, da die Patienten aufgrund des double-dummy Designs der hier eingeschlossenen Studien, täglich jeweils eine Behandlung mit aktivem Wirkstoff und eine Behandlung mit Placebo einnehmen mussten. Alle Patienten nahmen somit mehrere Tabletten täglich ein, wodurch eine Übertragbarkeit der Adhärenz-Daten auf die Praxis, in der EVG/COBI/FTC/TDF-Patienten nur eine Tablette täglich und ATV/r+FTC/TDF-Patienten drei Tabletten täglich (ATV, RTV [Ritonavir] und FTC/TDF) einnehmen müssen, nicht gegeben. Aus diesem Grund werden die Daten hier vollständig dargestellt, jedoch nicht in die Bewertung und Nutzenableitung aufgenommen.

Jeweils ein Patient pro Behandlungsarm gab zu Woche 48 und 96 nicht mindestens eine Flasche der Studienmedikation wieder zurück, weswegen diese Patienten nicht in der Analyse der Therapieadhärenz berücksichtigt werden können.

Hinweise oder Belege werden primär unter Berücksichtigung des p-Wertes abgeleitet. Zusätzlich wird allerdings das I^2 -Maß dargestellt. Bedeutsame Heterogenität zwischen den Subgruppen wird ab einem p-Wert von $<0,20$ bzw. einem $I^2 > 50\%$ angenommen. Dabei besteht bei einem p-Wert von $\geq 0,05$ und $< 0,20$ ein Hinweis auf Effektmodifikation und ab $p < 0,05$ wird ein Beleg auf Effektmodifikation abgeleitet. Für den Fall dass ein p-Wert genau den Wert 0,20 oder 0,05 annimmt oder Inkonsistenzen zwischen p-Wert und Konfidenzintervall auftreten, wird der p-Wert mittels der Software R, unter Verwendung der gleichen Methode wie in RevMan 5.2, 4-stellig nachberechnet, um eine eindeutige Ableitung eines Hinweises oder Belegs zu gewährleisten.

Generell gilt zu beachten, dass die Wahrscheinlichkeit, mindestens einen signifikanten Interaktionstest zu erhalten, obwohl tatsächlich keine Interaktion besteht, bei der Vielzahl an durchgeführten Subgruppenanalysen, bei nahezu 1 liegt. Bei einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$, sind bei 294 durchgeführten Subgruppenanalysen 14 bis 15 fälschlicherweise als signifikant eingestufte Interaktionen zu erwarten.

Detaillierte Ergebnisse werden für alle präspezifizierte Subgruppenanalysen und darüber hinaus für die Subgruppenanalysen dargestellt, die in einen signifikanten Interaktionstest resultieren; alle weiteren Interaktionstests werden im Hauptteil lediglich mittels p-Wert in einer Übersichtstabelle dargestellt (Tabelle 4-47). Details zu allen durchgeführten Subgruppenanalysen werden aus Übersichtlichkeitsgründen in der Anlage [104] vollständig dargestellt.

In der folgenden Tabelle 4-47 werden alle präspezifizierten und post-hoc durchgeführten Interaktionstests mittels p-Wert zusammengefasst. Dabei werden die Analysezeitpunkte (Woche 48 und Woche 96) durch unterschiedliche Grau-Schattierungen differenziert dargestellt.

Tabelle 4-47: Übersicht über alle präspezifizierten (fett gedruckt) und post-hoc durchgeführten Interaktionstests

Endpunkt	p-Wert aus Interaktionstest					
	Alter	Geschlecht	Rasse	Ausgangs viruslast	CD4-Zellzahl zu Baseline	Therapieadhärenz
Woche 48						
Woche 96						
Virologisches Ansprechen (Snapshot-Algorithmus)	0,95	0,61	0,76	0,77	0,67	0,76
	0,12	0,10	0,63	0,85	0,62	0,092
Virologisches Ansprechen (TLOVR-Algorithmus)	0,91	0,64	0,71	0,55	0,25	0,87
	0,33	0,39	0,65	0,82	0,19	0,06
Zeit bis zum Verlust des virologischen Ansprechens	0,85	0,45	0,63	0,62	0,25	0,87
	0,47	0,27	0,73	0,86	0,19	0,04
Veränderung der CD4-Zellzahl	0,81	0,38	0,89	0,62	0,07	0,0530
	0,75	0,04	0,32	0,94	0,02	0,50
Mortalität	n.b.	n.b.	0,96	0,78	0,87	n.b.
	n.b.	n.b.	0,96	0,78	0,87	n.b.
Gesamt UE	0,79	0,64	0,55	0,17	0,90	0,91
	0,70	0,32	0,89	0,86	0,47	0,63
Gesamt SUE	0,0539	0,25	0,32	0,08	0,91	0,33
	0,40	0,87	0,53	0,19	0,34	0,99
Abbrüche aufgrund UE	0,95	0,95	0,10	0,53	0,60	0,80
	0,73	0,65	0,23	0,38	0,49	0,75
UE Grad 3-4	0,38	0,91	0,79	0,0002	0,75	0,15
	0,59	0,62	0,68	0,002	0,76	0,59
Neurologische & psychiatrische Ereignisse	0,58	0,15	0,60	0,75	0,75	0,12
	0,99	0,14	0,93	0,77	0,49	0,39

Endpunkt	p-Wert aus Interaktionstest					
	Alter	Geschlecht	Rasse	Ausgangs viruslast	CD4-Zellzahl zu Baseline	Therapieadhärenz
Woche 48						
Woche 96						
Hautausschläge	0,17	0,19	0,0494	0,94	0,17	0,39
	0,32	0,30	0,30	0,46	0,53	0,73
Renale Ereignisse	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.
	0,94	n.b.	n.b.	0,66	0,75	0,97
Knochenfrakturen	0,19	0,16	0,32	0,94	0,07	0,64
	0,50	0,09	0,81	0,63	0,23	0,55
Erkrankungen des Gastrointestinaltrakts (SOC)	0,66	0,63	0,01	0,25	0,09	0,44
	0,75	0,63	0,02	0,13	0,07	0,37
<i>Diarrhoe</i>	0,63	0,32	0,40	0,15	0,01	0,87
	0,43	0,30	0,41	0,39	0,18	0,44
<i>Übelkeit</i>	0,0541	0,31	0,40	0,60	0,14	0,82
	0,04	0,41	0,24	0,52	0,16	0,75
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)	0,87	0,49	0,75	0,34	0,58	1,00
	0,63	0,71	0,96	0,41	0,51	0,33
<i>Fatigue</i>	0,14	0,74	0,44	0,72	0,14	0,59
	0,04	0,51	0,32	0,56	0,43	0,90
Erkrankungen des Nervensystems (SOC)	0,11	0,91	0,17	0,96	0,57	0,44
	0,21	0,75	0,18	0,62	0,60	0,84
<i>Kopfschmerzen</i>	0,88	0,58	0,11	0,88	0,25	0,50
	0,70	0,46	0,45	0,94	0,18	0,79
Psychiatrische Erkrankungen (SOC)	0,54	0,004	0,76	0,12	0,58	0,01
	0,17	0,009	0,63	0,22	0,38	0,23
Depression	n.z.	n.z.	n.z.	n.z.	n.z.	n.z.
	0,19	0,13	0,13	0,99	0,51	0,91
Erkrankungen der Haut und des Unterhautzellgewebes (SOC)	0,10	0,27	0,66	0,85	0,25	0,40
	0,21	0,31	0,79	0,67	0,47	0,83
Augenerkrankungen (SOC)	0,02	0,65	0,15	0,16	0,59	0,18
	0,15	0,28	0,04	0,54	0,42	0,14
<i>Skleraler Ikterus</i>	0,97	0,34	0,75	0,87	0,82	0,36
	0,91	0,34	0,75	0,87	0,82	0,51
Präspezifizierte Subgruppenanalysen sind fett gedruckt						

Endpunkt	p-Wert aus Interaktionstest					
Woche 48	Alter	Geschlecht	Rasse	Ausgangsviruslast	CD4-Zellzahl zu Baseline	Therapieadhärenz
Woche 96						
n.b. = nicht berechenbar, n.z. = nicht zutreffend, SOC = System Organ Class, SUE = Schwerwiegendes unerwünschtes Ereignis, UE = Unerwünschtes Ereignis						

Virologisches Ansprechen (Snapshot)

Zu Woche 48 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,95$; $I^2=0\%$), Geschlecht ($p=0,61$; $I^2=0\%$), Rasse ($p=0,76$; $I^2=0\%$), Ausgangsviruslast ($p=0,77$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,67$; $I^2=0\%$) oder Therapieadhärenz ($p=0,76$; $I^2=0\%$).

Zu Woche 96 gibt es keine Hinweise oder Beleg auf Effektmodifikation durch Rasse ($p=0,63$; $I^2=0\%$), Ausgangsviruslast ($p=0,85$; $I^2=0\%$) oder CD4-Zellzahl zu Baseline ($p=0,62$; $I^2=0\%$). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Alter ($p=0,12$; $I^2=60,7\%$). In der Gruppe der unter 40-jährigen beträgt das OR [95%-KI] 1,40 [0,85;2,29] mit $p=0,1856$ und in der Gruppe der mindestens 40-jährigen ergibt sich das OR [95%-KI] zu 0,73 [0,38;1,40] mit $p=0,3170$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Die Konfidenzintervalle der Risikodifferenzen (RD [95%-KI]: <40 Jahre: 5,3 [-2,6;13,3]; ≥ 40 Jahre: -3,5 [-11,4;4,3]) zeigen außerdem, dass in beiden Subgruppen EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF nichtunterlegen zur 12%-Grenze ist. Des Weiteren zeigt sich ein Hinweis auf Effektmodifikation durch Geschlecht ($p=0,10$; $I^2=58,9\%$). In der Gruppe der Männer ergibt sich das OR [95%-KI] zu 1,20 [0,78;1,84] mit $p=0,4303$ und in der Gruppe der Frauen zu 0,47 [0,17;1,33] mit $p=0,1732$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch Therapieadhärenz ($p=0,092$; $I^2=61,7\%$). In der Gruppe der Patienten, die nach 96 Wochen eine Therapieadhärenz <95% aufwiesen, beträgt das OR [95%-KI] 0,70 [0,38;1,32] mit $p=0,2930$, und in der Gruppe der Patienten die nach 96 Wochen eine Therapieadhärenz $\geq 95\%$ aufwiesen, ergibt sich das OR [95%-KI] zu 1,42 [0,84;2,41] mit $p=0,2076$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Virologisches Ansprechen (TLOVR)

Es bestehen zu Woche 48 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,91$; $I^2=0\%$), Geschlecht ($p=0,64$; $I^2=0\%$), Rasse ($p=0,71$; $I^2=0\%$), Ausgangsviruslast ($p=0,55$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,25$; $I^2=25,5\%$) oder Therapieadhärenz ($p=0,87$; $I^2=0\%$).

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,33$; $I^2=0\%$), Geschlecht ($p=0,39$; $I^2=0\%$), Rasse ($p=0,65$; $I^2=0\%$) oder Ausgangsviruslast ($p=0,82$; $I^2=0\%$). Allerdings besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl

zu Baseline ($p=0,19$; $I^2=42,3\%$). In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl ≤ 350 Zellen/ μL aufwiesen, beträgt das OR [95%-KI] 0,99 [0,59;1,67] mit $p=0,9788$ und in der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl >350 Zellen/ μL aufwiesen, ist das OR [95%-KI] 1,63 [0,96;2,75] mit $p=0,0679$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Auch hinsichtlich Therapieadhärenz konnte ein Hinweis auf eine Effektmodifikation gefunden werden ($p=0,06$; $I^2=72,7\%$). In der Patientengruppe, die zu $<95\%$ adhärenz waren, beträgt das OR [95%-KI] 0,80 [0,44;1,45] mit $p=0,4568$ und in der Gruppe von Patienten, die eine Therapieadhärenz $\geq 95\%$ aufwiesen, ergibt sich das OR [95%-KI] zu 1,69 [1,04;2,74] mit $p=0,0337$. Es besteht somit in der Gruppe der Patienten, die das Medikament zu $\geq 95\%$ adhärenz einnahmen, ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF hinsichtlich des virologischen Ansprechens (TLOVR).

Zeit bis zum Verlust des virologischen Ansprechens

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,85$; $I^2=0\%$), Geschlecht ($p=0,45$; $I^2=0\%$), Rasse ($p=0,63$; $I^2=0\%$), Ausgangsviruslast ($p=0,62$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,25$; $I^2=24,5\%$) oder Therapieadhärenz ($p=0,87$; $I^2=0\%$).

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,47$; $I^2=0\%$), Geschlecht ($p=0,27$; $I^2=18,9\%$), Rasse ($p=0,73$; $I^2=0\%$) oder Ausgangsviruslast ($p=0,86$; $I^2=0\%$). Es besteht jedoch ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,19$; $I^2=42,2\%$). Für Patienten, die zu Baseline eine CD4-Zellzahl ≤ 350 Zellen/ μL aufwiesen, beträgt das HR [95%-KI] 1,01 [0,64;1,60] mit $p=0,96$ und für Patienten mit höherer CD4-Zellzahl zu Baseline liegt das HR [95%-KI] bei 0,65 [0,41;1,04] mit $p=0,074$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Außerdem besteht ein Beleg für Effektmodifikation durch Therapieadhärenz ($p=0,04$; $I^2=75,2\%$). Für Patienten, die die Studienmedikation $<95\%$ adhärenz einnahmen, beträgt das HR [95%-KI] 1,19 [0,72;1,95] mit $p=0,49$. Für Patienten, die eine Therapieadhärenz von $\geq 95\%$ aufwiesen, ergibt sich das HR [95%-KI] zu 0,60 [0,39;0,94] mit $p=0,025$. Es besteht somit für Patienten mit hoher Therapieadhärenz ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF.

Veränderung der CD4-Zellzahl

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,81$; $I^2=0\%$), Geschlecht ($p=0,38$; $I^2=0\%$), Rasse ($p=0,89$; $I^2=0\%$) oder Ausgangsviruslast ($p=0,62$; $I^2=0\%$). Allerdings besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,07$; $I^2=68,5\%$). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline von ≤ 350 Zellen/ μL , beträgt die MWD [95%-KI] -28,00 [-60,04;4,04] mit $p=0,0871$. In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL ergibt sich eine MWD [95%-KI] von 17,00 [-20,71;54,71] mit $p=0,3769$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf

Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch Therapieadhärenz ($p=0,0530$; $I^2=73,3\%$). Für Patienten, die zu $<95\%$ adhären waren ergibt sich eine MWD [95%-KI] von $-51,00$ $[-106,09;4,09]$ mit $p=0,0660$. Für Patienten, die zu $\geq 95\%$ adhären waren, ergibt sich die MWD [95%-KI] zu $10,00$ $[-17,99;37,99]$ mit $p=0,4838$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, ist dieser Hinweis auf Effektmodifikation nicht relevant für das Gesamtfazit.

Zu Woche 96 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,75$; $I^2=0\%$), Rasse ($p=0,32$; $I^2=0\%$), Ausgangsviruslast ($p=0,94$; $I^2=0\%$) oder Therapieadhärenz ($p=0,50$; $I^2=0\%$). Es besteht jedoch ein Beleg für eine Effektmodifikation durch Geschlecht ($p=0,04$; $I^2=76,6\%$). In der Gruppe der männlichen Patienten, beträgt die MWD [95%-KI] $7,00$ $[-21,43;35,43]$ mit $p=0,6290$ und unter den Frauen beträgt sie $-107,00$ $[-211,21;-2,79]$ mit $p=0,0442$ (Hedges' g [95%-KI]: $-0,51$ $[-1,08;0,05]$). Es besteht somit unter den Frauen ein statistisch signifikanter, klinisch jedoch nicht relevanter, Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF. Da jedoch die Subgruppenanalysen, stratifiziert nach Geschlecht, aufgrund der Heterogenität der Subgruppen (primär durch den Einfluss weiblicher Sexualhormone [Estrogen] unterscheidet sich das weibliche Immunsystem hinsichtlich der Immunantwort von dem männlichen; es wurde außerdem in Studie S103 nicht nach Geschlecht stratifiziert) und durch die sehr geringe Anzahl an Frauen in der Studie, mit einer geringen Aussagekraft behaftet sind, ist diese Subgruppenanalyse mit Vorsicht zu interpretieren. Ein weiterer Beleg für Effektmodifikation besteht durch CD4-Zellzahl zu Baseline ($p=0,02$; $I^2=81,5\%$). Bei Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL ergibt sich die MWD [95%-KI] zu $-40,00$ $[-77,16;-2,84]$ mit $p=0,0350$ (Hedges' g [95%-KI]: $-0,24$ $[-0,47;-0,02]$), was einen statistisch signifikanten, jedoch klinisch nicht relevanten, Unterschied zuungunsten von EVG/COBI/FTC/TDF zeigt. In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl >350 Zellen/ μL aufwiesen, beträgt die MWD [95%-KI] $25,00$ $[-15,32;65,32]$, $p=0,2274$.

Mortalität

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter (Interaktionstest aufgrund zu geringer Ereignisrate nicht berechenbar), Geschlecht (Interaktionstest aufgrund zu geringer Ereignisrate nicht berechenbar), Rasse ($p=0,96$; $I^2=0\%$), Ausgangsviruslast ($p=0,78$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,87$; $I^2=0\%$) oder Therapieadhärenz (Interaktionstest aufgrund zu geringer Ereignisrate nicht berechenbar).

Auch zu Woche 96 können keine Hinweise oder Belege auf Effektmodifikation durch Alter (Interaktionstest aufgrund zu geringer Ereignisrate nicht berechenbar), Geschlecht (Interaktionstest aufgrund zu geringer Ereignisrate nicht berechenbar), Rasse ($p=0,96$; $I^2=0\%$), Ausgangsviruslast ($p=0,78$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,87$; $I^2=0\%$) oder Therapieadhärenz (Interaktionstest aufgrund zu geringer Ereignisrate nicht berechenbar) gefunden werden.

Gesamtrate unerwünschter Ereignisse

Es bestehen zu Woche 48 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,79$; $I^2=0\%$), Geschlecht ($p=0,64$; $I^2=0\%$), Rasse ($p=0,55$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,90$; $I^2=0\%$) oder Therapieadhärenz ($p=0,91$; $I^2=0\%$). Allerdings besteht ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,17$; $I^2=47,1\%$). In der Gruppe der Patienten, die zu Baseline eine Viruslast ≤ 100.000 Kopien/mL aufwiesen, beträgt das OR [95%-KI] 0,94 [0,47;1,90] mit $p=0,8725$. In der Gruppe der Patienten, die zu Baseline eine Viruslast 100.000 Kopien/mL aufwiesen, ergibt sich das OR [95%-KI] zu 0,39 [0,13;1,12] mit $p=0,0791$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,70$; $I^2=0\%$), Geschlecht ($p=0,32$; $I^2=0,2\%$), Rasse ($p=0,89$; $I^2=0\%$), Ausgangsviruslast ($p=0,86$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,47$; $I^2=0\%$) oder Therapieadhärenz ($p=0,63$; $I^2=0\%$).

Gesamtrate schwerwiegender unerwünschter Ereignisse

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,25$; $I^2=25,4\%$), Rasse ($p=0,32$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,91$; $I^2=0\%$) oder Therapieadhärenz ($p=0,33$; $I^2=0\%$). Allerdings besteht ein Hinweis auf Effektmodifikation durch Alter ($p=0,0539$; $I^2=73,1\%$). In der Gruppe der unter 40-jährigen beträgt das OR [95%-KI] 0,57 [0,29;1,10] mit $p=0,0913$ und in der Gruppe der ≥ 40 -jährigen Patienten ergibt sich das OR [95%-KI] zu 1,90 [0,67;5,36] mit $p=0,2263$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Des Weiteren besteht ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,08$; $I^2=67,8\%$). Für Patienten mit Ausgangsviruslast von ≤ 100.000 Kopien/mL beläuft sich das OR [95%-KI] auf 1,32 [0,62;2,82] mit $p=0,4696$ und in der Gruppe der Patienten mit Ausgangsviruslast > 100.000 Kopien/mL ist das OR [95%-KI] 0,49 [0,22;1,10] mit $p=0,0827$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Es bestehen zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,40$; $I^2=0\%$), Geschlecht ($p=0,87$; $I^2=0\%$), Rasse ($p=0,53$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,34$; $I^2=0\%$) oder Therapieadhärenz ($p=0,99$; $I^2=0\%$). Es besteht allerdings ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,19$; $I^2=41,5\%$). In der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL beträgt das OR [95%-KI] 0,84 [0,46;1,53] mit $p=0,5731$ und für Patienten mit Ausgangsviruslast > 100.000 Kopien/mL ergibt sich das OR [95%-KI] zu 0,45 [0,21;0,94] mit $p=0,0325$. Somit besteht in der Gruppe der Ausgangsviruslast > 100.000 Kopien/mL ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF.

Therapieabbrüche aufgrund unerwünschter Ereignisse

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,95$; $I^2=0\%$), Geschlecht ($p=0,95$; $I^2=0\%$), Ausgangsviruslast ($p=0,53$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,60$; $I^2=0\%$) oder Therapieadhärenz ($p=0,80$; $I^2=0\%$). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Rasse ($p=0,10$; $I^2=64,1\%$). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 1,02 [0,46;2,29] mit $p=0,9540$ und unter den nicht-kaukasischen Patienten ergibt sich ein OR [95%-KI] von 0,14 [0,02;1,25] mit $p=0,0788$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,73$; $I^2=0\%$), Geschlecht ($p=0,65$; $I^2=0\%$), Rasse ($p=0,23$; $I^2=31,8\%$), Ausgangsviruslast ($p=0,38$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,49$; $I^2=0\%$) oder Therapieadhärenz ($p=0,75$; $I^2=0\%$).

Unerwünschte Ereignisse Grad 3-4

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,38$; $I^2=0\%$), Geschlecht ($p=0,91$; $I^2=0\%$), Rasse ($p=0,79$; $I^2=0\%$) oder CD4-Zellzahl zu Baseline ($p=0,75$; $I^2=0\%$). Es kann jedoch ein Beleg für Effektmodifikation durch Ausgangsviruslast gefunden werden ($p=0,0002$; $I^2=92,9\%$). In der Gruppe der Patienten, die zu Baseline eine Viruslast ≤ 100.000 Kopien/mL aufwiesen, beträgt das OR [95%-KI] 1,95 [1,08;3,52] mit $p=0,0262$, was einen statistisch signifikanten Unterschied zuungunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL zeigt sich jedoch mit einem OR [95%-KI] von 0,32 [0,15;0,67] und $p=0,0025$ ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF. Es besteht außerdem ein Hinweis auf Effektmodifikation durch Therapieadhärenz ($p=0,15$; $I^2=52,0\%$). Unter den Patienten, die zu $<95\%$ adhären waren, ergibt sich ein OR [95%-KI] von 0,57 [0,25;1,28] mit $p=0,1741$ und für Patienten mit einer Therapieadhärenz von $\geq 95\%$ beträgt das OR [95%-KI] 1,16 [0,69;1,95] mit $p=0,5797$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Es können zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,59$; $I^2=0\%$), Geschlecht ($p=0,62$; $I^2=0\%$), Rasse ($p=0,68$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,76$; $I^2=0\%$) oder Therapieadhärenz ($p=0,59$; $I^2=0\%$) gefunden werden. Es besteht jedoch ein Beleg für Effektmodifikation durch Ausgangsviruslast ($p=0,002$; $I^2=90,1\%$). In der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL beträgt das OR [95%-KI] 1,48 [0,89;2,46] mit $p=0,1268$ und für Patienten mit Ausgangsviruslast >100.000 Kopien/mL ergibt sich das OR [95%-KI] zu 0,37 [0,18;0,74] mit $p=0,0050$. Es besteht somit in der Gruppe der Patienten mit höherer Ausgangsviruslast ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF.

Hautausschläge

Zu Woche 48 können keine Hinweise oder Belege auf Effektmodifikation durch Ausgangsviruslast ($p=0,94$; $I^2=0\%$) oder Therapieadhärenz ($p=0,39$; $I^2=0\%$) gefunden werden. Allerdings gibt es einen Hinweis auf Effektmodifikation durch Alter ($p=0,17$; $I^2=47,6\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,77 [0,46;1,31] mit $p=0,3366$ und für die Gruppe der mindestens 40-jährigen Patienten ergibt sich das OR [95%-KI] zu 1,33 [0,76;2,34] mit $p=0,3205$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch Geschlecht ($p=0,19$; $I^2=41,6\%$). In der Gruppe der Männer beträgt das OR [95%-KI] 1,05 [0,71;1,57] mit $p=0,7969$ und in der Gruppe der weiblichen Patienten ergibt sich ein OR [95%-KI] von 0,34 [0,06;1,77] mit $p=0,1991$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Es besteht außerdem ein Beleg für Effektmodifikation durch Rasse ($p=0,0494$; $I^2=74,1\%$). In der Gruppe der kaukasischen Patienten berechnet sich ein OR [95%-KI] von 0,82 [0,53;1,28] und $p=0,3862$ und in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 2,29 [0,91;5,77] mit $p=0,0777$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation konnte durch CD4-Zellzahl zu Baseline gefunden werden ($p=0,17$; $I^2=46,6\%$). Für Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL beträgt das OR [95%-KI] 0,76 [0,45;1,29] mit $p=0,3028$ und in der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl >350 Zellen/ μL aufwiesen, ergibt sich das OR [95%-KI] zu 1,30 [0,74;2,28] mit $p=0,3636$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Zu Woche 96 können keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,32$; $I^2=0\%$), Geschlecht ($p=0,30$; $I^2=5,4\%$), Rasse ($p=0,30$; $I^2=5,2\%$), Ausgangsviruslast ($p=0,46$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,53$; $I^2=0\%$) oder Therapieadhärenz ($p=0,73$; $I^2=0\%$) gefunden werden.

Neurologische & psychiatrische Ereignisse

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,58$; $I^2=0\%$), Rasse ($p=0,60$; $I^2=0\%$), Ausgangsviruslast ($p=0,75$; $I^2=0\%$) oder CD4-Zellzahl zu Baseline ($p=0,75$; $I^2=0\%$). Es gibt jedoch einen Hinweis auf Effektmodifikation durch Geschlecht ($p=0,15$; $I^2=51,2\%$). In der Gruppe der Männer beträgt das OR [95%-KI] 0,87 [0,63;1,21] und $p=0,4169$ und für die weiblichen Patienten ergibt sich ein OR [95%-KI] von 1,87 [0,70;5,01] mit $p=0,2151$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht für Therapieadhärenz ($p=0,12$; $I^2=58,1\%$). Für Patienten, die zu $<95\%$ adhärenz waren, ist das OR [95%-KI] 0,61 [0,32;1,14] mit $p=0,1204$ und für Patienten mit einer $\geq 95\%$ -igen Therapieadhärenz ergibt sich das OR [95%-KI] zu 1,07 [0,75;1,53] mit

$p=0,6942$. Da aber auch hier in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Es bestehen zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,99$; $I^2=0\%$), Rasse ($p=0,93$; $I^2=0\%$), Ausgangsviruslast ($p=0,77$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,49$; $I^2=0\%$) oder Therapieadhärenz ($p=0,39$; $I^2=0\%$). Allerdings besteht ein Hinweis auf Effektmodifikation durch Geschlecht ($p=0,14$; $I^2=53,4\%$). In der Gruppe der Männer beträgt das OR [95%-KI] 0,89 [0,65, 1,21] mit $p=0,4538$ und für die Gruppe der Frauen ist das OR [95%-KI] 1,91 [0,72, 5,09] mit $p=0,1940$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Renale Ereignisse

Bis Woche 48 wurden keine renalen Ereignisse beobachtet. Die Interaktionstests sind somit alle aufgrund zu geringer Ereignisraten nicht berechenbar.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,94$; $I^2=0\%$), Geschlecht (Interaktionstest aufgrund zu geringer Ereignisraten nicht berechenbar), Rasse (Interaktionstest aufgrund zu geringer Ereignisraten nicht berechenbar), Ausgangsviruslast ($p=0,66$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,75$; $I^2=0\%$) oder Therapieadhärenz ($p=0,97$; $I^2=0\%$).

Knochenfrakturen

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Rasse ($p=0,32$; $I^2=0,8\%$), Ausgangsviruslast ($p=0,94$; $I^2=0\%$) oder Therapieadhärenz ($p=0,64$; $I^2=0\%$). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Alter ($p=0,19$; $I^2=42,1\%$). In der Gruppe der unter 40-jährigen beträgt das OR [95%-KI] 1,88 [0,17;20,93] mit $p=0,6067$ und in der Gruppe der mindestens 40-jährigen Patienten beträgt das OR [95%-KI] 0,22 [0,02;1,86] mit $p=0,1629$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch Geschlecht ($p=0,16$; $I^2=48,2\%$). In der Gruppe der Männer beträgt das OR [95%-KI] 0,32 [0,06;1,60] mit $p=0,1659$ und in der Gruppe der Frauen ergibt sich das OR [95%-KI] zu 4,16 [0,16;105,81] mit $p=0,3882$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Außerdem besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,07$; $I^2=69,0\%$). In der Gruppe der Patienten, die eine CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL aufwiesen, beträgt das OR [95%-KI] 2,81 [0,29;27,28] mit $p=0,3732$, und in der Gruppe der Patienten mit einer höheren CD4-Zellzahl zu Baseline ergibt sich das OR [95%-KI] zu 0,10 [0,01;1,75] mit $p=0,1136$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Es bestehen zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,50$; $I^2=0\%$), Rasse ($p=0,81$; $I^2=0\%$), Ausgangsviruslast ($p=0,63$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,23$; $I^2=31,2\%$) oder Therapieadhärenz ($p=0,55$; $I^2=0\%$). Jedoch besteht ein Hinweis auf Effektmodifikation durch Geschlecht ($p=0,09$; $I^2=65,7\%$). In der Gruppe der männlichen Patienten ergibt sich ein OR [95%-KI] von 0,20 [0,06;0,71] mit $p=0,0125$, was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF anzeigt, und in der Gruppe der Frauen beträgt das OR [95%-KI] 4,16 [0,16;105,81] mit $p=0,3882$. Da in der Gruppe der Frauen nur sehr wenig Patienten eingeschlossen sind und lediglich ein Patient eine Knochenfraktur erlitt, ist diese Interaktion mit Vorsicht zu interpretieren.

Der beschriebene Unterschied zwischen EVG/COBI/FTC/TDF und ATV/r+FTC/TDF nach 96 Wochen ist weder mit der in Abschnitt 4.3.1.3.1.4 beschriebenen Pathogenese noch aufgrund des kurzen Beobachtungszeitraumes schlüssig erklärbar. Eine für Frakturen ursächliche Knochendichtereduktion über 96 Wochen hätte zudem einen erheblichen Anstieg entsprechender Serumparameter, wie zum Beispiel der alkalischen Phosphatase und oder der Lactat-Dehydrogenase verursachen müssen, der in der Studie aber nicht beobachtet werden konnte.

Erkrankungen des Gastrointestinaltrakts (SOC)

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,66$; $I^2=0\%$), Geschlecht ($p=0,63$; $I^2=0\%$), Ausgangsviruslast ($p=0,25$; $I^2=25,2\%$) oder Therapieadhärenz ($p=0,44$; $I^2=0\%$). Es besteht ein Beleg für Effektmodifikation durch Rasse ($p=0,01$; $I^2=84,1\%$). In der Gruppe der Kaukasier beträgt das OR [95%-KI] 1,12 [0,79;1,58] mit $p=0,5288$ und in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,46 [0,25;0,84] mit $p=0,0113$. Es besteht somit unter den nicht-kaukasischen Patienten hinsichtlich der Vermeidung von Erkrankungen des Gastrointestinaltrakts ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF. Ein Hinweis auf Effektmodifikation besteht durch CD4-Zellzahl zu Baseline ($p=0,09$; $I^2=64,4\%$). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL beträgt das OR [95%-KI] 1,15 [0,75;1,77] mit $p=0,5142$ und in der Gruppe der Patienten mit einer CD4-Zellzahl zu Baseline >350 Zellen/ μL ergibt sich ein OR [95%-KI] von 0,69 [0,46;1,05] mit $p=0,0821$.

Nach 96 Wochen bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,75$; $I^2=0\%$), Geschlecht ($p=0,63$; $I^2=0\%$) oder Therapieadhärenz ($p=0,37$; $I^2=0\%$). Es besteht jedoch ein Beleg für Effektmodifikation durch Rasse ($p=0,02$; $I^2=82,5\%$). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 1,00 [0,70;1,42] mit $p=0,9823$ und in der Gruppe der nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,42 [0,23;0,78] mit $p=0,0057$. Somit besteht ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF unter den nicht-kaukasischen Patienten. Hinsichtlich Ausgangsviruslast kann ein Hinweis auf Effektmodifikation gefunden werden ($p=0,13$; $I^2=55,8$). In der Gruppe der Patienten, die zu Baseline eine Viruslast ≤ 100.000 Kopien/mL aufwiesen, beträgt das OR [95%-KI] 0,65 [0,44;0,97] mit $p=0,0328$, was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF hinsichtlich der Vermeidung gastrointestinaler Erkrankungen anzeigt. In der Gruppe der Patienten mit Ausgangsviruslast

>100.000 Kopien/mL besteht kein Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,04 [0,65;1,67], $p=0,8599$). Außerdem besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,07$; $I^2=69,2\%$). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL beträgt das OR [95%-KI] 1,06 [0,68;1,63] mit $p=0,8029$ und in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL ergibt sich das OR [95%-KI] zu 0,61 [0,40;0,92] mit $p=0,0202$. Es besteht also in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF hinsichtlich der Vermeidung gastrointestinaler Erkrankungen.

Da Viruslast und CD4-Zellzahl in der Regel miteinander korrelieren, ist es nicht verwunderlich, bei signifikanter Interaktion durch Ausgangsviruslast auch eine signifikante Interaktion durch CD4-Zellzahl zu Baseline zu finden.

Daten zeigen, dass nicht nur die antiretrovirale Medikation – insbesondere geboosterte Protease-Inhibitoren –, sondern auch die HIV-Infektion selbst und der immunologische Status des Patienten einen Einfluss auf den Gastrointestinaltrakt haben [47]. Das könnte zur Folge haben, dass die Verträglichkeitsunterschiede zwischen den Behandlungen bei den kränkeren Patienten (Ausgangsviruslast >100.000 Kopien/mL, bzw. CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL) durch den zusätzlichen Einfluss der HIV-Infektion selbst überlagert werden, und sich die eigentlichen Unterschiede der Medikation erst bei einem besseren immunologischen Status zeigen, da der additive Effekt der HIV-Infektion gewissermaßen wegfällt. Dies korreliert mit den gefundenen Ergebnissen: Bei den Erkrankungen des Gastrointestinaltrakts zeigt sich ein Anhaltspunkt für einen geringen Zusatznutzen für EVG/COBI/FTC/TDF bei Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL, bzw. CD4-Zellzahl zu Baseline >350 Zellen/ μL .

Diarrhoe

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,63$; $I^2=0\%$), Geschlecht ($p=0,32$; $I^2=0,5\%$), Rasse ($p=0,40$; $I^2=0\%$) oder Therapieadhärenz ($p=0,87$; $I^2=0\%$). Allerdings gibt es einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,15$; $I^2=51,9\%$). In der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL beträgt das OR [95%-KI] 0,59 [0,37;0,94] mit $p=0,0257$, was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,99 [0,59;1,67], $p=0,9627$). Außerdem besteht ein Beleg für Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,01$; $I^2=83,2\%$). In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl ≤ 350 Zellen/ μL aufwiesen, beträgt das OR [95%-KI] 1,18 [0,71;1,95] mit $p=0,5280$. In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,49 [0,30;0,80], $p=0,0042$).

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,43$; $I^2=0\%$), Geschlecht ($p=0,30$; $I^2=7,4\%$), Rasse ($p=0,41$; $I^2=0\%$), Ausgangsviruslast ($p=0,39$; $I^2=0\%$) oder Therapieadhärenz ($p=0,44$; $I^2=0\%$). Allerdings besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,18$; $I^2=44,3\%$). In der Gruppe der Patienten, die eine CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL aufwiesen, beträgt das OR [95%-KI] 0,93 [0,57;1,49] mit $p=0,7546$. In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL ergibt sich das OR [95%-KI] zu 0,59 [0,37;0,93] mit $p=0,0240$. Es besteht also ein statistisch signifikanter Unterschied zwischen den Behandlungen, der zugunsten von EVG/COBI/FTC/TDF ausgeprägt ist.

Gastrointestinale Nebenwirkungen und Diarrhoe stellen eine häufige Nebenwirkung bei allen Protease-Inhibitoren dar. Geht man von einem kombinierten negativen Effekt von HIV-Infektion und ART auf gastrointestinale Nebenwirkungen aus, könnte der zusätzliche gastrointestinale Effekt durch den ungünstigeren immunologischen Status einen Unterschied zwischen EVG/COBI/FTC/TDF und ATV/r+FTC/TDF verschleiern. Basierend auf dieser Hypothese würde sich ein günstigeres gastrointestinales Profil von EVG/COBI/FTC/TDF bei den Patienten mit besserem Immunstatus zeigen, was den gefundenen Daten entspricht.

Übelkeit

Es bestehen zu Woche 48 keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,31$; $I^2=3,8\%$), Rasse ($p=0,40$; $I^2=0\%$), Ausgangsviruslast ($p=0,60$; $I^2=0\%$) oder Therapieadhärenz ($p=0,82$; $I^2=0\%$). Es gibt jedoch einen Hinweis auf Effektmodifikation durch Alter ($p=0,0541$; $I^2=73,1\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 1,33 [0,83;2,14] mit $p=0,2317$, und in der Gruppe der mindestens 40-jährigen Patienten beträgt das OR [95%-KI] 0,61 [0,33;1,16] mit $p=0,1308$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch CD4-Zellzahl zu Baseline ($p=0,14$; $I^2=53,5\%$). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL beträgt das OR [95%-KI] 1,33 [0,79;2,25] mit $p=0,2791$, und in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL beträgt das OR [95%-KI] 0,76 [0,44;1,30] mit $p=0,3207$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,41$; $I^2=0\%$), Rasse ($p=0,24$; $I^2=27,4\%$), Ausgangsviruslast ($p=0,52$; $I^2=0\%$) oder Therapieadhärenz ($p=0,75$; $I^2=0\%$). Allerdings konnte ein Beleg für Effektmodifikation durch Alter gefunden werden ($p=0,04$; $I^2=76,8\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 1,31 [0,83;2,08] mit $p=0,2512$, und für die Gruppe der mindestens 40-jährigen Patienten ist das OR [95%-KI] 0,58 [0,31;1,07] mit $p=0,0829$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Beleg für Effektmodifikation nicht auf das Gesamtfazit aus. Ein Hinweis auf Effektmodifikation kann durch CD4-Zellzahl zu Baseline gefunden werden ($p=0,16$; $I^2=48,5\%$). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL beträgt

das OR [95%-KI] 1,27 [0,76;2,13] mit $p=0,3563$ und in der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl >350 Zellen/ μL aufwiesen, beträgt das OR [95%-KI] 0,76 [0,45;1,27] mit $p=0,2943$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,87$; $I^2=0\%$), Geschlecht ($p=0,49$; $I^2=0\%$), Rasse ($p=0,75$; $I^2=0\%$), Ausgangsviruslast ($p=0,34$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,58$; $I^2=0\%$) oder Therapieadhärenz ($p=1,00$; $I^2=0\%$).

Auch zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,63$; $I^2=0\%$), Geschlecht ($p=0,71$; $I^2=0\%$), Rasse ($p=0,96$; $I^2=0\%$), Ausgangsviruslast ($p=0,41$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,51$; $I^2=0\%$) oder Therapieadhärenz ($p=0,33$; $I^2=0\%$).

Fatigue

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,74$; $I^2=0\%$), Rasse ($p=0,44$; $I^2=0\%$), Ausgangsviruslast ($p=0,72$; $I^2=0\%$) oder Therapieadhärenz ($p=0,59$; $I^2=0\%$). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Alter ($p=0,14$; $I^2=55,2\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,85 [0,48;1,51] mit $p=0,5823$ und für die Gruppe der mindestens 40-jährigen Patienten ist das OR [95%-KI] 1,66 [0,85;3,24] mit $p=0,1339$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch CD4-Zellzahl zu Baseline ($p=0,14$; $I^2=53,7\%$). In der Gruppe der Patienten mit CD4-Zellzahl ≤ 350 Zellen/ μL beträgt das OR [95%-KI] 1,59 [0,85;2,99] mit $p=0,1504$ und in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL beträgt das OR [95%-KI] 0,82 [0,45;1,52] mit $p=0,5310$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,51$; $I^2=0\%$), Rasse ($p=0,32$; $I^2=0\%$), Ausgangsviruslast ($p=0,56$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,43$; $I^2=0\%$) oder Therapieadhärenz ($p=0,90$; $I^2=0\%$). Es besteht jedoch ein Beleg für Effektmodifikation durch Alter ($p=0,04$; $I^2=76,7\%$). In der Gruppe der Patienten unter 40 Jahre beträgt das OR [95%-KI] 0,63 [0,36;1,09] mit $p=0,0983$ und in der Gruppe der mindestens 40-jährigen Patienten beträgt das OR [95%-KI] 1,51 [0,81;2,81] mit $p=0,1922$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Beleg für Effektmodifikation nicht auf das Gesamtfazit aus.

Erkrankungen des Nervensystems (SOC)

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,91$; $I^2=0\%$), Ausgangsviruslast ($p=0,96$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,57$; $I^2=0\%$) oder Therapieadhärenz ($p=0,44$; $I^2=0\%$). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Alter ($p=0,11$; $I^2=61,8\%$). In der Gruppe der Patienten unter 40 Jahren beträgt das OR [95%-KI] 1,26 [0,80;1,99] mit $p=0,3128$, und in der Gruppe der Patienten, die mindestens 40 Jahre alt sind, beträgt das OR [95%-KI] 0,72 [0,43;1,20] mit $p=0,2048$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch Rasse ($p=0,17$; $I^2=47,8\%$). In der Gruppe der kaukasischen Patienten ergibt sich ein OR [95%-KI] von 1,13 [0,76;1,66] mit $p=0,5447$ und für die nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,65 [0,33;1,28] mit $p=0,2118$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Beleg für Effektmodifikation nicht auf das Gesamtfazit aus.

Es bestehen zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,21$; $I^2=35,9\%$), Geschlecht ($p=0,75$; $I^2=0\%$), Ausgangsviruslast ($p=0,62$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,60$; $I^2=0\%$) oder Therapieadhärenz ($p=0,84$; $I^2=0\%$). Allerdings besteht ein Hinweis auf Effektmodifikation durch Rasse ($p=0,18$; $I^2=43,5\%$). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 1,03 [0,72;1,49] mit $p=0,8555$ und für die nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,63 [0,34;1,19] mit $p=0,1530$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Kopfschmerzen

Es bestehen zu Woche 48 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,88$; $I^2=0\%$), Geschlecht ($p=0,58$; $I^2=0\%$), Ausgangsviruslast ($p=0,88$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,25$; $I^2=24,3\%$) oder Therapieadhärenz ($p=0,50$; $I^2=0\%$). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Rasse ($p=0,11$; $I^2=60,3\%$). In der Gruppe der Kaukasier beträgt das OR [95%-KI] 1,53 [0,91;2,57] mit $p=0,1053$ und in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,72 [0,33;1,57] mit $p=0,4043$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Zu Woche 96 können keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,70$; $I^2=0\%$), Geschlecht ($p=0,46$; $I^2=0\%$), Rasse ($p=0,45$; $I^2=0\%$), Ausgangsviruslast ($p=0,94$; $I^2=0\%$) oder Therapieadhärenz ($p=0,79$; $I^2=0\%$) gefunden werden. Allerdings gibt es einen Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,18$; $I^2=45,2\%$). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL beträgt das OR [95%-KI] 1,45 [0,84;2,52] mit $p=0,1863$ und in der Gruppe der Patienten, die an Baseline eine CD4-Zellzahl >350 Zellen/ μL aufwiesen, ergibt sich das OR [95%-KI] zu 0,82 [0,45;1,52] mit $p=0,5310$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied

zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Psychiatrische Erkrankungen (SOC)

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,54$; $I^2=0\%$), Rasse ($p=0,76$; $I^2=0\%$) oder CD4-Zellzahl zu Baseline ($p=0,58$; $I^2=0\%$). Es besteht ein Beleg für Effektmodifikation durch Geschlecht ($p=0,004$; $I^2=87,8\%$). In der Gruppe der Männer beträgt das OR [95%-KI] 0,76 [0,53;1,11] mit $p=0,1597$ und in der Gruppe der weiblichen Patienten ergibt sich das OR [95%-KI] zu 5,35 [1,49;19,19] mit $p=0,0101$. In der Gruppe der weiblichen Patienten besteht somit ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF. Da jedoch die Subgruppenanalysen, stratifiziert nach Geschlecht, aufgrund der Heterogenität der Subgruppen (primär durch den Einfluss weibliche Sexualhormone [Estrogen] unterscheidet sich das weibliche Immunsystem hinsichtlich der Immunantwort von dem männlichen ; es wurde außerdem in Studie S103 nicht nach Geschlecht stratifiziert) und durch die sehr geringe Anzahl an Frauen in der Studie, mit einer geringen Aussagekraft behaftet sind, ist diese Subgruppenanalyse mit Vorsicht zu interpretieren. Ein Hinweis auf Effektmodifikation besteht durch Ausgangsviruslast ($p=0,12$; $I^2=57,5\%$). In der Gruppe der Patienten, die zu Baseline eine Viruslast ≤ 100.000 Kopien/mL aufwiesen, beträgt das OR [95%-KI] 0,71 [0,44;1,15] mit $p=0,1619$ und in der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL beträgt das OR [95%-KI] 1,26 [0,73;2,17] mit $p=0,4142$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Des Weiteren besteht ein Beleg für Effektmodifikation durch Therapieadhärenz ($p=0,01$; $I^2=84,2\%$). In der Gruppe der Patienten, die zu $<95\%$ adhären waren, beträgt das OR [95%-KI] 0,42 [0,21;0,84] mit $p=0,0147$ und in der Gruppe der zu $\geq 95\%$ adhären Patienten ergibt sich das OR [95%-KI] zu 1,19 [0,78;1,82] mit $p=0,4096$. In der Gruppe der Patienten, die weniger adhären waren, besteht also ein statistisch signifikanter Vorteil von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF. Während im ATV/r+FTC/TDF-Arm die Anzahl psychiatrischer Erkrankungen bei geringerer Therapieadhärenz gegenüber $\geq 95\%$ iger Therapieadhärenz deutlich größer ist, kann im EVG/COBI/FTC/TDF-Arm bei geringerer Therapieadhärenz keine Erhöhung der Inzidenz psychiatrischer Erkrankungen gegenüber der stärker adhären Gruppe beobachtet werden.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Rasse ($p=0,63$; $I^2=0\%$), Ausgangsviruslast ($p=0,22$; $I^2=34,9\%$), CD4-Zellzahl zu Baseline ($p=0,38$; $I^2=0\%$) oder Therapieadhärenz ($p=0,23$; $I^2=29,8\%$). Allerdings besteht ein Hinweis auf Effektmodifikation durch Alter ($p=0,17$; $I^2=47,1\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,74 [0,49;1,12] mit $p=0,1571$ und in der Gruppe der Patienten ≥ 40 Jahre ergibt sich das OR [95%-KI] zu 1,18 [0,70;1,97] mit $p=0,5329$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Des Weiteren besteht ein Beleg für Effektmodifikation durch Geschlecht ($p=0,009$; $I^2=85,4\%$). In der Gruppe der männlichen Patienten beträgt das OR [95%-KI] 0,78 [0,56;1,10] mit $p=0,1581$

und für die weiblichen Patienten ergibt sich ein OR [95%-KI] von 4,16 [1,25;13,82] mit $p=0,0202$. In der Gruppe der Frauen besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF. Allerdings ist diese Subgruppenanalyse aufgrund der sehr geringen Anzahl an Frauen in der Studie mit Vorsicht zu interpretieren.

Depression

Zu Woche 48 traten Depressionen nicht bei mindestens 10% der Patienten in einem der Behandlungsarme auf, weswegen zu diesem Zeitpunkt keine Subgruppenanalysen durchgeführt werden.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Ausgangsviruslast ($p=0,99$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,51$; $I^2=0\%$) oder Therapieadhärenz ($p=0,91$; $I^2=0\%$). Es besteht ein Hinweis auf Effektmodifikation durch Alter ($p=0,19$; $I^2=42,1\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 1,09 [0,57;2,09] mit $p=0,7837$, und in der Gruppe der mindestens 40-jährigen Patienten ergibt sich das OR [95%-KI] zu 0,57 [0,27;1,19] mit $p=0,1323$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch Geschlecht ($p=0,13$; $I^2=55,4\%$). In der Gruppe der männlichen Patienten ergibt sich das OR [95%-KI] zu 0,72 [0,44;1,20] mit $p=0,2101$ und für die Frauen beträgt das OR [95%-KI] 2,96 [0,50;17,41] mit $p=0,2299$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Außerdem gibt es einen Hinweis auf Effektmodifikation durch Rasse ($p=0,13$; $I^2=55,8\%$). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 0,68 [0,39;1,18] mit $p=0,1694$ und unter den nicht-kaukasischen Patienten ergibt sich ein OR [95%-KI] 1,75 [0,58;5,25] mit $p=0,3212$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Erkrankungen der Haut und des Unterhautzellgewebes (SOC)

Es bestehen zu Woche 48 keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,27$; $I^2=18,7\%$), Rasse ($p=0,66$; $I^2=0\%$), Ausgangsviruslast ($p=0,85$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,25$; $I^2=24,8\%$) oder Therapieadhärenz ($p=0,40$; $I^2=0\%$). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Alter ($p=0,10$; $I^2=63,6\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,74 [0,48;1,16] mit $p=0,1946$ und für Patienten ≥ 40 Jahre beträgt das OR [95%-KI] 1,30 [0,80;2,11] mit $p=0,2896$. Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,21$; $I^2=37,0$), Geschlecht ($p=0,31$; $I^2=2,7$), Rasse ($p=0,79$; $I^2=0\%$), Ausgangsviruslast

($p=0,67$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,47$; $I^2=0\%$) oder Therapieadhärenz ($p=0,83$; $I^2=0\%$).

Augenerkrankungen (SOC)

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,65$; $I^2=0\%$) oder CD4-Zellzahl zu Baseline ($p=0,59$; $I^2=0\%$). Es gibt jedoch einen Beleg für Effektmodifikation durch Alter ($p=0,02$; $I^2=81,9\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] mit 0,11 [0,05;0,23] mit $p<0,00001$, und unter den älteren Patienten ergibt sich das OR [95%-KI] zu 0,42 [0,18;0,93] mit $p=0,0334$. Dies ist lediglich eine quantitative Interaktion, da in beiden Subgruppen ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF vorliegt. Des Weiteren besteht ein Hinweis auf Effektmodifikation durch Rasse ($p=0,15$; $I^2=51,5\%$). In der Gruppe der kaukasischen Patienten ist das OR [95%-KI] 0,24 [0,13;0,44] mit $p<0,00001$, und für die nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,09 [0,02;0,31] mit $p=0,0001$. Auch hier liegt eine quantitative Interaktion vor und in beiden Subgruppen besteht ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF. Auch hinsichtlich Ausgangsviruslast besteht ein Hinweis auf Effektmodifikation ($p=0,16$; $I^2=50,1\%$). Doch auch hier besteht eine quantitative Interaktion, da sowohl in der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL (OR [95%-KI]: 0,13 [0,06;0,29], $p<0,00001$) als auch in der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL (OR [95%-KI]: 0,28 [0,14;0,58], $p=0,0007$) ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF besteht. Außerdem gibt es einen Hinweis auf Effektmodifikation durch Therapieadhärenz ($p=0,18$; $I^2=43,7\%$). Und auch hier ist die Interaktion rein quantitativer Natur (OR [95%-KI]: Adhärenz $<95\%$: 0,08 [0,02;0,35], $p=0,0009$; Adhärenz $\geq 95\%$: 0,23 [0,13;0,42], $p<0,00001$) mit statistisch signifikanten Unterschieden zugunsten von EVG/COBI/FTC/TDF in beiden Subgruppen.

Insgesamt ist zu Woche 48 zu beobachten, dass EVG/COBI/FTC/TDF über nahezu alle Subgruppen hinweg einen statistisch signifikanten Vorteil gegenüber ATV/r+FTC/TDF zeigt.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ($p=0,28$; $I^2=13,1\%$), Ausgangsviruslast ($p=0,54$; $I^2=0\%$) oder CD4-Zellzahl zu Baseline ($p=0,42$; $I^2=0\%$). Allerdings besteht hier - im Gegensatz zu Woche 48, wo sich ein Beleg für Effektmodifikation durch Alter zeigt - lediglich ein Hinweis auf Effektmodifikation durch Alter ($p=0,15$; $I^2=51,6\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,21 [0,11;0,37] mit $p<0,00001$ und in der Gruppe der mindestens 40-jährigen Patienten ergibt sich das OR [95%-KI] zu 0,41 [0,20;0,83] mit $p=0,0138$. Da in beiden Subgruppen ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF besteht, liegt auch hier nur eine quantitative Interaktion vor. Ein Beleg für Effektmodifikation besteht für Rasse ($p=0,04$; $I^2=76,7\%$). In der Gruppe der kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,36 [0,22;0,60] mit $p=0,0001$, und unter den nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,09 [0,02;0,31] mit $p=0,0001$. Da in beiden Subgruppen ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF vorliegt, ist dies lediglich eine quantitative Interaktion. Ein weiterer Hinweis auf Effektmodifikation besteht

durch Therapieadhärenz ($p=0,14$; $I^2=53,1\%$). Für Patienten, die zu $<95\%$ adhären waren, beträgt das OR [95%-KI] 0,13 [0,04;0,40] mit $p=0,0003$ und für Patienten, die eine Therapieadhärenz $\geq 95\%$ aufwiesen ergibt sich das OR [95%-KI] zu 0,33 [0,20;0,55] mit $p<0,0001$. Da in beiden Subgruppen ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF vorliegt, ist auch dies eine quantitative Interaktion.

Auch zu Woche 96 zeigen sich hinsichtlich der SOC Augenerkrankungen in nahezu allen Subgruppen statistisch signifikante Vorteile von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF.

Bei der Bewertung der signifikanten Interaktionstests gilt zu beachten, dass die Wahrscheinlichkeit, mindestens eine fälschlicherweise als signifikant eingestufte Interaktion zu erhalten, bei der Menge an durchgeführten Subgruppenanalysen sehr hoch ist.

Skleraler Ikterus

Es bestehen zu Woche 48 keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,97$; $I^2=0\%$), Geschlecht ($p=0,34$; $I^2=0\%$), Rasse ($p=0,75$; $I^2=0\%$), Ausgangsviruslast ($p=0,87$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,82$; $I^2=0\%$) oder Therapieadhärenz ($p=0,36$; $I^2=0\%$).

Und auch zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ($p=0,91$; $I^2=0\%$), Geschlecht ($p=0,34$; $I^2=0\%$), Rasse ($p=0,75$; $I^2=0\%$), Ausgangsviruslast ($p=0,87$; $I^2=0\%$), CD4-Zellzahl zu Baseline ($p=0,82$; $I^2=0\%$) oder Therapieadhärenz ($p=0,51$; $I^2=0\%$).

Generell sind Subgruppenanalysen, stratifiziert nach Therapieadhärenz, kritisch zu betrachten, da naturgemäß nicht zu Baseline stratifiziert werden konnte, sondern sich die Zuteilung zu den Subgruppen erst durch die erhobenen Daten ergeben hat. Da die Behandlung und möglicherweise auch andere Faktoren einen Einfluss auf die erst retrospektiv erfolgte Subgruppenbildung hatten, ist die Aussagekraft dieser Subgruppenanalyse als niedrig einzuschätzen. Die gefundenen Ergebnisse zeigen zumeist nur, dass eine höhere Therapieadhärenz zu mehr Nebenwirkungen führt, womit gegebenenfalls vorliegende Unterschiede im Nebenwirkungsprofil signifikant werden können. Da die Therapieadhärenz vor dem Hintergrund einer möglichen Resistenzentwicklung in jedem Fall unabhängig vom antiretroviralen Regime bei idealerweise 100% liegen muss, sind die gefundenen Daten klinisch von minderer Bedeutung. Darüber hinaus sind die in der vorliegenden Nutzenbewertung betrachteten Daten auf die klinische Praxis und den tatsächlichen Einsatz von EVG/COBI/FTC/TDF nicht übertragbar, da die Patienten aufgrund des double-dummy Designs der hier eingeschlossenen Studien täglich jeweils eine Behandlung mit aktivem Wirkstoff und eine Behandlung mit Placebo einnehmen mussten. Alle Patienten nahmen somit mehrere Tabletten täglich ein, wodurch eine Übertragbarkeit der Adhärenz-Daten auf die Praxis, in der EVG/COBI/FTC/TDF-Patienten nur eine Tablette täglich und ATV/r+FTC/TDF-Patienten drei Tabletten täglich (ATV, RTV und FTC/TDF) einnehmen

müssen, nicht gegeben. Aus diesem Grund werden die Daten hier vollständig dargestellt, jedoch nicht in die Bewertung und Nutzenableitung aufgenommen.

Im Folgenden werden die Ergebnisse der signifikanten Subgruppenanalysen detailliert dargestellt.

4.3.1.3.2.1 Virologisches Ansprechen (Snapshot-Algorithmus; präspezifizierte Subgruppenanalysen)

4.3.1.3.2.1.1 Virologisches Ansprechen (Snapshot) zu Woche 48

4.3.1.3.2.1.1.1 Virologisches Ansprechen (Snapshot) stratifiziert nach Alter

Tabelle 4-48: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p- Wert ^b	RD (%) [95%- KI] ^c
	N	n (%)	N	n (%)		
<40 Jahre	206	181 (87,9)	193	163 (84,5)	1,34 [0,75;2,39], p=0,3248	3,4 [-3,5;10,3]
≥40 Jahre	147	135 (91,8)	162	145 (89,5)	1,38 [0,63;3,02], p=0,4840	3,1 [-3,4;9,7]
p-Wert-Interaktionstest^d:					0,95	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.1.2 Virologisches Ansprechen (Snapshot) stratifiziert nach Geschlecht

Tabelle 4-49: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Geschlecht (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p- Wert ^b	RD (%) [95%- KI] ^c
	N	n (%)	N	n (%)		
Männlich	324	292 (90,1)	316	276 (87,3)	1,37 [0,83;2,25], p=0,2666	3,1 [-1,9;8,0]
Weiblich	29	24 (82,8)	39	32 (82,1)	0,96 [0,27;3,43], p=0,9397	-1,2 [-20,6;18,3]
p-Wert-Interaktionstest^d:					0,61	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.1.3 Virologisches Ansprechen (Snapshot) stratifiziert nach Rasse

Tabelle 4-50: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Rasse (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p- Wert ^b	RD (%) [95%- KI] ^c
	N	n (%)	N	n (%)		
Kaukasisch	250	225 (90,0)	277	240 (86,6)	1,41 [0,82;2,42], p=0,2337	3,4 [-2,1;9,0]
Nicht- kaukasisch	103	91 (88,3)	78	68 (87,2)	1,19 [0,48;2,95], p=0,8115	2,1 [-7,9;12,1]
p-Wert-Interaktionstest^d:					0,76	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.1.4 Virologisches Ansprechen (Snapshot) stratifiziert nach Ausgangsviruslast

Tabelle 4-51: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Ausgangsviruslast (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p- Wert ^b	RD (%) [95%- KI] ^c
	N	n (%)	N	n (%)		
≤100.000 Kopien/mL	203	188 (92,6)	214	192 (89,7)	1,44 [0,72;2,85], p=0,3014	2,9 [-2,5;8,3]
≥100.000 Kopien/mL	150	128 (85,3)	141	116 (82,3)	1,25 [0,67;2,34], p=0,4784	3,1 [-5,4;11,5]
p-Wert-Interaktionstest^d:					0,77	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.1.5 Virologisches Ansprechen (Snapshot) stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-52: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p- Wert ^b	RD (%) [95%- KI] ^c
	N	n (%)	N	n (%)		
≤350 Zellen/μL	176	157 (89,2)	163	143 (87,7)	1,19 [0,61;2,33], p=0,6709	1,7 [-5,3;8,7]
>350 Zellen/μL	177	159 (89,8)	192	165 (85,9)	1,46 [0,77;2,76], p=0,2554	3,9 [-2,8;10,6]
p-Wert-Interaktionstest^d:					0,67	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.1.6 Virologisches Ansprechen (Snapshot) stratifiziert nach Therapieadhärenz

Tabelle 4-53: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Therapieadhärenz (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^b	RD (%) [95%-KI] ^c
	N	n (%)	N	n (%)		
<95%	94	76 (80,9)	78	59 (75,6)	1,30 [0,62;2,73], p=0,4087	4,3 [-8,2;16,8]
≥95%	258	240 (93,0)	276	249 (90,2)	1,52 [0,81;2,84], p=0,2455	3,1 [-1,6;7,8]
p-Wert-Interaktionstest^d:					0,76	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.2 Virologisches Ansprechen (Snapshot) zu Woche 96**4.3.1.3.2.1.2.1 Virologisches Ansprechen (Snapshot) stratifiziert nach Alter**

Tabelle 4-54: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Alter (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^b	RD (%) [95%-KI] ^c
	N	n (%)	N	n (%)		
<40 Jahre	206	170 (82,5)	193	149 (77,2)	1,40 [0,85;2,29], p=0,1856	5,3 [-2,6;13,3]
≥40 Jahre	147	124 (84,4)	162	143 (88,3)	0,73 [0,38;1,40], p=0,3170	-3,5 [-11,4;4,3]
p-Wert-Interaktionstest^d:					0,12	
I² (zusätzlich berechnet):					60,7%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.2.2 Virologisches Ansprechen (Snapshot) stratifiziert nach Geschlecht

Tabelle 4-55: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Geschlecht (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p- Wert ^b	RD (%) [95%- KI] ^c
	N	n (%)	N	n (%)		
Männlich	324	277 (85,5)	316	263 (83,2)	1,20 [0,78;1,84], p=0,4303	2,4 [-3,3;8,0]
Weiblich	29	17 (58,6)	39	29 (74,4)	0,47 [0,17;1,33], p=0,1732	-17,8 [-40,6;5,0]
p-Wert-Interaktionstest^d:					0,10	
I² (zusätzlich berechnet):					58,9%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.2.3 Virologisches Ansprechen (Snapshot) stratifiziert nach Rasse

Tabelle 4-56: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Rasse (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p- Wert ^b	RD (%) [95%- KI] ^c
	N	n (%)	N	n (%)		
Kaukasisch	250	212 (84,8)	277	229 (82,7)	1,17 [0,74;1,87], p=0,5093	2,2 [-4,2;8,5]
Nicht- kaukasisch	103	82 (79,6)	78	63 (80,8)	0,95 [0,45;1,99], p=0,8468	-0,7 [-12,8;11,4]
p-Wert-Interaktionstest^d:					0,63	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.2.4 Virologisches Ansprechen (Snapshot) stratifiziert nach Ausgangsviruslast

Tabelle 4-57: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Ausgangsviruslast (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^b	RD (%) [95%-KI] ^c
	N	n (%)	N	n (%)		
≤100.000 Kopien/mL	203	171 (84,2)	214	179 (83,6)	1,04 [0,62;1,76], p=0,8694	0,6 [-6,5;7,6]
≥100.000 Kopien/mL	150	123 (82,0)	141	113 (80,1)	1,13 [0,63;2,03], p=0,6859	1,9 [-7,2;10,9]
p-Wert-Interaktionstest^d:					0,85	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.2.5 Virologisches Ansprechen (Snapshot) stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-58: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^b	RD (%) [95%-KI] ^c
	N	n (%)	N	n (%)		
≤350 Zellen/μL	176	144 (81,8)	163	134 (82,2)	0,98 [0,56;1,71], p=0,9255	-0,4 [-8,7;7,9]
>350 Zellen/μL	177	150 (84,7)	192	158 (82,3)	1,20 [0,69;2,08], p=0,5264	2,5 [-5,2;10,1]
p-Wert-Interaktionstest^d:					0,62	
I² (zusätzlich berechnet):					0%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.1.2.6 Virologisches Ansprechen (Snapshot) stratifiziert nach Therapieadhärenz

Tabelle 4-59: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Therapieadhärenz (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p- Wert ^b	RD (%) [95%- KI] ^c
	N	n (%)	N	n (%)		
<95%	97	67 (69,1)	103	78 (75,7)	0,70 [0,38;1,32], p=0,2930	-6,3 [-18,9;6,3]
≥95%	255	227 (89,0)	251	214 (85,3)	1,42 [0,84;2,41], p=0,2076	4,2 [-1,7;10,1]
p-Wert-Interaktionstest^d:					0,092	
I² (zusätzlich berechnet):					61,7%	
^a Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe ^b zusätzlich berechnet (Chi-Quadrat Test) ^c RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet ^d Wald Test						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

4.3.1.3.2.2 Virologisches Ansprechen (TLOVR-Algorithmus)**4.3.1.3.2.2.1 Virologisches Ansprechen (TLOVR) zu Woche 96****4.3.1.3.2.2.1.1 Virologisches Ansprechen (TLOVR) stratifiziert nach CD4-Zellzahl zu Baseline**

Tabelle 4-60: Subgruppenanalyse für virologisches Ansprechen (TLOVR), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	138 (78,4)	163	128 (78,5)	0,99 [0,59;1,67], p=0,9788
>350 Zellen/μL	177	149 (84,2)	192	147 (76,6)	1,63 [0,96;2,75], p=0,0679
p-Wert Interaktionstest^a:					0,19
I² (zusätzlich berechnet):					42,3%
^a zusätzlich berechnet ^b Chi-Quadrat Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz					

4.3.1.3.2.1.2 Virologisches Ansprechen (TLOVR) stratifiziert nach Therapieadhärenz

Tabelle 4-61: Subgruppenanalyse für virologisches Ansprechen (TLOVR), stratifiziert nach Therapieadhärenz (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<95%	97	64 (66,0)	103	73 (70,9)	0,80 [0,44;1,45], p=0,4568
≥95%	255	223 (87,5)	251	202 (80,5)	1,69 [1,04;2,74], p=0,0337
p-Wert Interaktionstest^a:					0,06
I² (zusätzlich berechnet):					72,7%
^a zusätzlich berechnet ^b Chi-Quadrat Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz					

4.3.1.3.2.3 Zeit bis zum Verlust des virologischen Ansprechens**4.3.1.3.2.4 Zeit bis zum Verlust des virologischen Ansprechens zu Woche 96****4.3.1.3.2.4.1 Zeit bis zum Verlust des virologischen Ansprechens stratifiziert nach CD4-Zellzahl zu Baseline**

Tabelle 4-62: Subgruppenanalyse für Zeit bis zum Verlust des virologischen Ansprechens, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		HR ^a [95%-KI] ^b , p - Wert ^a
	N	n (% ^c)	N	n (% ^c)	
≤350 Zellen/μL	140	38 (26,0)	130	35 (23,0)	1,01 [0,64;1,60], p=0,96
>350 Zellen/μL	150	28 (16,0)	151	45 (25,0)	0,65 [0,41;1,04], p=0,074
p-Wert-Interaktionstest^d:					0,19
I² (zusätzlich berechnet):					42,2%
^a Cox-Proportional Hazards Modell ^b Wald Konfidenzintervall ^c Kaplan-Meier Schätzer (am Ende des Woche 96-Analysefensters) ^d zusätzlich berechnet KI = Konfidenzintervall, N: Anzahl der Patienten unter Risiko am Anfang des Woche 96-Analysefensters, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz					

4.3.1.3.2.4.2 Zeit bis zum Verlust des virologischen Ansprechens stratifiziert nach Therapieadhärenz

Tabelle 4-63: Subgruppenanalyse für Zeit bis zum Verlust des virologischen Ansprechens, stratifiziert nach Therapieadhärenz (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		HR ^a [95%-KI] ^b , p - Wert ^a
	N	n (% ^c)	N	n (% ^c)	
<95%	67	33 (39,0)	76	30 (33,0)	1,19 [0,72;1,95], p=0,49
≥95%	223	32 (14,0)	205	49 (20,0)	0,60 [0,39;0,94], p=0,025
p-Wert-Interaktionstest^d:					0,04
I² (zusätzlich berechnet):					75,2%
^a Cox-Proportional Hazards Modell ^b Wald Konfidenzintervall ^c Kaplan-Meier Schätzer (am Ende des Woche 96-Analysefensters) ^d zusätzlich berechnet KI = Konfidenzintervall, N: Anzahl der Patienten unter Risiko am Anfang des Woche 96-Analysefensters, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz					

4.3.1.3.2.5 Veränderung der CD4-Zellzahl

4.3.1.3.2.5.1 Veränderung der CD4-Zellzahl zu Woche 48

4.3.1.3.2.5.1.1 Veränderung der CD4-Zellzahl stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-64: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		MWD ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	MW der Veränderung (SD)	N	MW der Veränderung (SD)	
≤350 Zellen/μL	168	199 (142,9)	149	227 (147,3)	-28,00 [-60,04;4,04], p=0,0871
>350 Zellen/μL	166	215 (183,3)	172	198 (169,9)	17,00 [-20,71;54,71], p=0,3769
p-Wert Interaktionstest^a:					0,07
I² (zusätzlich berechnet):					68,5%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, SD = Standardabweichung					

4.3.1.3.2.5.1.2 Veränderung der CD4-Zellzahl stratifiziert nach Therapieadhärenz

Tabelle 4-65: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach Therapieadhärenz (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		MWD ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	MW der Veränderung (SD)	N	MW der Veränderung (SD)	
<95%	83	171 (153,6)	63	222 (178,5)	-51,00 [-106,09;4,09], p=0,0660
≥95%	251	219 (166,2)	258	209 (155,7)	10,00 [-17,99;37,99], p=0,4838
p-Wert Interaktionstest^a:					0,0530
I² (zusätzlich berechnet):					73,3%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, SD = Standardabweichung					

4.3.1.3.2.5.2 Veränderung der CD4-Zellzahl zu Woche 96**4.3.1.3.2.5.2.1 Veränderung der CD4-Zellzahl stratifiziert nach Geschlecht**

Tabelle 4-66: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach Geschlecht (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		MWD ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	MW der Veränderung (SD)	N	MW der Veränderung (SD)	
Männlich	296	257 (167,9)	282	250 (180,2)	7,00 [-21,43;35,43], p=0,6290
Weiblich	20	245 (158,3)	33	352 (227,9)	-107,00 [-211,21;-2,79], p=0,0442
p-Wert Interaktionstest^a:					0,04
I² (zusätzlich berechnet):					76,6%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, SD = Standardabweichung					

4.3.1.3.2.5.2.2 Veränderung der CD4-Zellzahl stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-67: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		MWD ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	MW der Veränderung (SD)	N	MW der Veränderung (SD)	
≤350 Zellen/μL	158	256 (158,0)	143	296 (169,7)	-40,00 [-77,16;-2,84], p=0,0350
>350 Zellen/μL	158	256 (176,1)	172	231 (197,6)	25,00 [-15,32;65,32], p=0,2274
p-Wert Interaktionstest^a:					0,02
I² (zusätzlich berechnet):					81,5%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, SD = Standardabweichung					

4.3.1.3.2.6 Gesamtrate unerwünschter Ereignisse

4.3.1.3.2.6.1 Gesamtrate unerwünschter Ereignisse zu Woche 48

4.3.1.3.2.6.1.1 Gesamtrate unerwünschter Ereignisse stratifiziert nach Ausgangsviruslast

Tabelle 4-68: Subgruppenanalyse für Gesamtrate unerwünschter Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	186 (91,6)	214	197 (92,1)	0,94 [0,47;1,90], p=0,8725
>100.000 Kopien/mL	150	137 (91,3)	141	136 (96,5)	0,39 [0,13;1,12], p=0,0791
p-Wert Interaktionstest^a:					0,17
I² (zusätzlich berechnet):					47,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.7 Gesamtrate schwerwiegender unerwünschter Ereignisse**4.3.1.3.2.7.1 Gesamtrate schwerwiegender unerwünschter Ereignisse zu Woche 48****4.3.1.3.2.7.1.1 Gesamtrate schwerwiegender unerwünschter Ereignisse stratifiziert nach Alter**

Tabelle 4-69: Subgruppenanalyse für Gesamtrate schwerwiegender unerwünschter Ereignisse, stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	16 (7,8)	193	25 (13,0)	0,57 [0,29;1,10], p=0,0913
≥40 Jahre	147	10 (6,8)	162	6 (3,7)	1,90 [0,67;5,36], p=0,2263
p-Wert Interaktionstest^a:					0,0539
I² (zusätzlich berechnet):					73,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.7.1.2 Gesamtrate schwerwiegender unerwünschter Ereignisse stratifiziert nach Ausgangsviruslast

Tabelle 4-70: Subgruppenanalyse für Gesamtrate schwerwiegender unerwünschter Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	16 (7,9)	214	13 (6,1)	1,32 [0,62;2,82], p=0,4696
>100.000 Kopien/mL	150	10 (6,7)	141	18 (12,8)	0,49 [0,22;1,10], p=0,0827
p-Wert Interaktionstest^a:					0,08
I² (zusätzlich berechnet):					67,8%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.7.2 Gesamtrate schwerwiegender unerwünschter Ereignisse zu Woche 96**4.3.1.3.2.7.2.1 Gesamtrate schwerwiegender unerwünschter Ereignisse stratifiziert nach Ausgangsviruslast**

Tabelle 4-71: Subgruppenanalyse für Gesamtrate schwerwiegender unerwünschter Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	22 (10,8)	214	27 (12,6)	0,84 [0,46;1,53], p=0,5731
>100.000 Kopien/mL	150	12 (8,0)	141	23 (16,3)	0,45 [0,21;0,94], p=0,0325
p-Wert Interaktionstest^a:					0,19
I² (zusätzlich berechnet):					41,5%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar ,OR = Odds Ratio					

4.3.1.3.2.8 Therapieabbrüche aufgrund unerwünschter Ereignisse**4.3.1.3.2.8.1 Therapieabbrüche aufgrund unerwünschter Ereignisse zu Woche 48****4.3.1.3.2.8.1.1 Therapieabbrüche aufgrund unerwünschter Ereignisse stratifiziert nach Rasse**

Tabelle 4-72: Subgruppenanalyse für Therapieabbrüche aufgrund unerwünschter Ereignisse, stratifiziert nach Rasse (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	12 (4,8)	277	13 (4,7)	1,02 [0,46;2,29], p=0,9540
Nicht-kaucasisch	103	1 (1,0)	78	5 (6,4)	0,14 [0,02;1,25], p=0,0788
p-Wert Interaktionstest^a:					0,10
I² (zusätzlich berechnet):					64,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.9 Unerwünschte Ereignisse Grad 3-4**4.3.1.3.2.9.1 Unerwünschte Ereignisse Grad 3-4 zu Woche 48****4.3.1.3.2.9.1.1 Unerwünschte Ereignisse Grad 3-4 stratifiziert nach Ausgangsviruslast**

Tabelle 4-73: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Ausgangsviruslast (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	34 (16,7)	214	20 (9,3)	1,95 [1,08;3,52], p=0,0262
>100.000 Kopien/mL	150	11 (7,3)	141	28 (19,9)	0,32 [0,15;0,67], p=0,0025
p-Wert Interaktionstest^a:					0,0002
I² (zusätzlich berechnet):					92,9%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.9.1.2 Unerwünschte Ereignisse Grad 3-4 stratifiziert nach Therapieadhärenz

Tabelle 4-74: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Therapieadhärenz (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<95%	94	12 (12,8)	78	16 (20,5)	0,57 [0,25;1,28], p=0,1741
≥95%	258	33 (12,8)	276	31 (11,2)	1,16 [0,69;1,95], p=0,5797
p-Wert Interaktionstest^a:					0,15
I² (zusätzlich berechnet):					52,0%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.9.2 Unerwünschte Ereignisse Grad 3-4 zu Woche 96**4.3.1.3.2.9.2.1 Unerwünschte Ereignisse Grad 3-4 stratifiziert nach Ausgangsviruslast**

Tabelle 4-75: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Ausgangsviruslast (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	42 (20,7)	214	32 (15,0)	1,48 [0,89;2,46], p=0,1268
>100.000 Kopien/mL	150	13 (8,7)	141	29 (20,6)	0,37 [0,18;0,74], p=0,0050
p-Wert Interaktionstest^a:					0,002
I² (zusätzlich berechnet):					90,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.10 Neurologische & psychiatrische Ereignisse**4.3.1.3.2.10.1 Neurologische & psychiatrische Ereignisse zu Woche 48****4.3.1.3.2.10.1.1 Neurologische & psychiatrische Ereignisse stratifiziert nach Geschlecht**

Tabelle 4-76: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Geschlecht (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Männlich	324	108 (33,3)	316	115 (36,4)	0,87 [0,63;1,21], p=0,4169
Weiblich	29	14 (48,3)	39	13 (33,3)	1,87 [0,70;5,01], p=0,2151
p-Wert Interaktionstest^a:					0,15
I² (zusätzlich berechnet):					51,2%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.10.1.2 Neurologische & psychiatrische Ereignisse stratifiziert nach Therapieadhärenz

Tabelle 4-77: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Therapieadhärenz (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<95%	94	29 (30,9)	78	33 (42,3)	0,61 [0,32;1,14], p=0,1204
≥95%	258	93 (36,0)	276	95 (34,4)	1,07 [0,75;1,53], p=0,6942
p-Wert Interaktionstest^a:					0,12
I² (zusätzlich berechnet):					58,1%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.10.2 Neurologische & psychiatrische Ereignisse zu Woche 96

4.3.1.3.2.10.2.1 Neurologische & psychiatrische Ereignisse stratifiziert nach Geschlecht

Tabelle 4-78: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Geschlecht (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Männlich	324	131 (40,4)	316	137 (43,4)	0,89 [0,65, 1,21], p=0,4538
Weiblich	29	15 (51,7)	39	14 (35,9)	1,91 [0,72, 5,09], p=0,1940
p-Wert Interaktionstest^a:					0,14
I² (zusätzlich berechnet):					53,4%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.11 Hautausschläge**4.3.1.3.2.11.1 Hautausschläge zu Woche 48****4.3.1.3.2.11.1.1 Hautausschläge stratifiziert nach Alter**

Tabelle 4-79: Subgruppenanalyse für Hautausschläge, stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	31 (15,0)	193	36 (18,7)	0,77 [0,46;1,31], p=0,3366
≥40 Jahre	147	32 (21,8)	162	28 (17,3)	1,33 [0,76;2,34], p=0,3205
p-Wert Interaktionstest^a:					0,17
I² (zusätzlich berechnet):					47,6%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.11.1.2 Hautausschläge stratifiziert nach Geschlecht

Tabelle 4-80: Subgruppenanalyse für Hautausschläge, stratifiziert nach Geschlecht (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Männlich	324	61 (18,8)	316	57 (18,0)	1,05 [0,71;1,57], p=0,7969
Weiblich	29	2 (6,9)	39	7 (17,9)	0,34 [0,06;1,77], p=0,1991
p-Wert Interaktionstest^a:					0,19
I² (zusätzlich berechnet):					41,6%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.11.1.3 Hautausschläge stratifiziert nach Rasse

Tabelle 4-81: Subgruppenanalyse für Hautausschläge, stratifiziert nach Rasse (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	44 (17,6)	277	57 (20,6)	0,82 [0,53;1,28], p=0,3862
Nicht-kaukasisch	103	19 (18,4)	78	7 (9,0)	2,29 [0,91;5,77], p=0,0777
p-Wert Interaktionstest^a:					0,0494
I² (zusätzlich berechnet):					74,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.11.1.4 Hautausschläge stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-82: Subgruppenanalyse für Hautausschläge, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	32 (18,2)	163	37 (22,7)	0,76 [0,45;1,29], p=0,3028
>350 Zellen/μL	177	31 (17,5)	192	27 (14,1)	1,30 [0,74;2,28], p=0,3636
p-Wert Interaktionstest^a:					0,17
I² (zusätzlich berechnet):					46,6%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.12 Knochenfrakturen**4.3.1.3.2.12.1 Knochenfrakturen zu Woche 48****4.3.1.3.2.12.1.1 Knochenfrakturen stratifiziert nach Alter**

Tabelle 4-83: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	2 (1,0)	193	1 (0,5)	1,88 [0,17;20,93], p=0,6067
≥40 Jahre	147	1 (0,7)	162	5 (3,1)	0,22 [0,02;1,86], p=0,1629
p-Wert Interaktionstest^a:					0,19
I² (zusätzlich berechnet):					42,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

4.3.1.3.2.12.1.2 Knochenfrakturen stratifiziert nach Geschlecht

Tabelle 4-84: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach Geschlecht (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Männlich	324	2 (0,6)	316	6 (1,9)	0,32 [0,06;1,60], p=0,1659
Weiblich	29	1 (3,4)	39	0 (0,0)	4,16 [0,16;105,81], p=0,3882
p-Wert Interaktionstest^a:					0,16
I² (zusätzlich berechnet):					48,2%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.12.1.3 Knochenfrakturen stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-85: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	3 (1,7)	163	1 (0,6)	2,81 [0,29;27,28], p=0,3732
>350 Zellen/μL	177	0 (0,0)	192	5 (2,6)	0,10 [0,01;1,75], p=0,1136
p-Wert Interaktionstest^a:					0,07
I² (zusätzlich berechnet):					69,0
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.12.2 Knochenfrakturen zu Woche 96**4.3.1.3.2.12.2.1 Knochenfrakturen stratifiziert nach Geschlecht**

Tabelle 4-86: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach Geschlecht (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Männlich	324	3 (0,9)	316	14 (4,4)	0,20 [0,06;0,71], p=0,0125
Weiblich	29	1 (3,4)	39	0 (0,0)	4,16 [0,16;105,81], p=0,3882
p-Wert Interaktionstest^a:					0,09
I² (zusätzlich berechnet):					65,7%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.13 Erkrankungen des Gastrointestinaltrakts (SOC)**4.3.1.3.2.13.1 Erkrankungen des Gastrointestinaltrakts (SOC) zu Woche 48****4.3.1.3.2.13.1.1 Erkrankungen des Gastrointestinaltrakts (SOC) stratifiziert nach Rasse**

Tabelle 4-87: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Rasse (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	144 (57,6)	277	152 (54,9)	1,12 [0,79;1,58], p=0,5288
Nicht-kaukasisch	103	45 (43,7)	78	49 (62,8)	0,46 [0,25;0,84], p=0,0113
p-Wert Interaktionstest^a:					0,01
I² (zusätzlich berechnet):					84,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.13.1.2 Erkrankungen des Gastrointestinaltrakts (SOC) stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-88: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	98 (55,7)	163	85 (52,1)	1,15 [0,75;1,77], p=0,5142
>350 Zellen/μL	177	91 (51,4)	192	116 (60,4)	0,69 [0,46;1,05], p=0,0821
p-Wert Interaktionstest^a:					0,09
I² (zusätzlich berechnet):					64,4%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.13.2 Erkrankungen des Gastrointestinaltrakts (SOC) zu Woche 96**4.3.1.3.2.13.2.1 Erkrankungen des Gastrointestinaltrakts (SOC) stratifiziert nach Rasse**

Tabelle 4-89: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Rasse (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	155 (62,0)	277	172 (62,1)	1,00 [0,70;1,42], p=0,9823
Nicht-kaukasisch	103	50 (48,5)	78	54 (69,2)	0,42 [0,23;0,78], p=0,0057
p-Wert Interaktionstest^a:					0,02
I² (zusätzlich berechnet):					82,5%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.13.2.2 Erkrankungen des Gastrointestinaltrakts (SOC) stratifiziert nach Ausgangsviruslast

Tabelle 4-90: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Ausgangsviruslast (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	112 (55,2)	214	140 (65,4)	0,65 [0,44;0,97], p=0,0328
>100.000 Kopien/mL	150	93 (62,0)	141	86 (61,0)	1,04 [0,65;1,67], p=0,8599
p-Wert Interaktionstest^a:					0,13
I² (zusätzlich berechnet):					55,8%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.13.2.3 Erkrankungen des Gastrointestinaltrakts (SOC) stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-91: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	106 (60,2)	163	96 (58,9)	1,06 [0,68;1,63], p=0,8029
>350 Zellen/μL	177	99 (55,9)	192	130 (67,7)	0,61 [0,40;0,92], p=0,0202
p-Wert Interaktionstest^a:					0,07
I² (zusätzlich berechnet):					69,2%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.14 Diarrhoe

4.3.1.3.2.14.1 Diarrhoe zu Woche 48

4.3.1.3.2.14.1.1 Diarrhoe stratifiziert nach Ausgangsviruslast

Tabelle 4-92: Subgruppenanalyse für Diarrhoe, stratifiziert nach Ausgangsviruslast (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	38 (18,7)	214	60 (28,0)	0,59 [0,37;0,94], p=0,0257
>100.000 Kopien/mL	150	39 (26,0)	141	37 (26,2)	0,99 [0,59;1,67], p=0,9627
p-Wert Interaktionstest^a:					0,15
I² (zusätzlich berechnet):					51,9%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.14.1.2 Diarrhoe stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-93: Subgruppenanalyse für Diarrhoe, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	44 (25,0)	163	36 (22,1)	1,18 [0,71;1,95], p=0,5280
>350 Zellen/μL	177	33 (18,6)	192	61 (31,8)	0,49 [0,30;0,80], p=0,0042
p-Wert Interaktionstest^a:					0,01
I² (zusätzlich berechnet):					83,2%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.14.1.3 Diarrhoe stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-94: Subgruppenanalyse für Diarrhoe, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	47 (26,7)	163	46 (28,2)	0,93 [0,57;1,49], p=0,7546
>350 Zellen/μL	177	41 (23,2)	192	65 (33,9)	0,59 [0,37;0,93], p=0,0240
p-Wert Interaktionstest^a:					0,18
I² (zusätzlich berechnet):					44,3%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.15 Übelkeit**4.3.1.3.2.15.1 Übelkeit zu Woche 48****4.3.1.3.2.15.1.1 Übelkeit stratifiziert nach Alter**

Tabelle 4-95: Subgruppenanalyse für Übelkeit, stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	52 (25,2)	193	39 (20,2)	1,33 [0,83;2,14], p=0,2317
≥40 Jahre	147	18 (12,2)	162	30 (18,5)	0,61 [0,33;1,16], p=0,1308
p-Wert Interaktionstest^a:					0,0541
I² (zusätzlich berechnet):					73,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.15.1.2 Übelkeit stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-96: Subgruppenanalyse für Übelkeit, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	42 (23,9)	163	31 (19,0)	1,33 [0,79;2,25], p=0,2791
>350 Zellen/μL	177	28 (15,8)	192	38 (19,8)	0,76 [0,44;1,30], p=0,3207
p-Wert Interaktionstest^a:					0,14
I² (zusätzlich berechnet):					53,5%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.15.2 Übelkeit zu Woche 96**4.3.1.3.2.15.2.1 Übelkeit stratifiziert nach Alter**

Tabelle 4-97: Subgruppenanalyse für Übelkeit, stratifiziert nach Alter (Woche 60)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	55 (26,7)	193	42 (21,8)	1,31 [0,83;2,08], p=0,2512
≥40 Jahre	147	19 (12,9)	162	33 (20,4)	0,58 [0,31;1,07], p=0,0829
p-Wert Interaktionstest^a:					0,04
I² (zusätzlich berechnet):					76,8%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.15.2.2 Übelkeit stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-98: Subgruppenanalyse für Übelkeit, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	43 (24,4)	163	33 (20,2)	1,27 [0,76;2,13], p=0,3563
>350 Zellen/μL	177	31 (17,5)	192	42 (21,9)	0,76 [0,45;1,27], p=0,2943
p-Wert Interaktionstest^a:					0,16
I² (zusätzlich berechnet):					48,5%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.16 Fatigue**4.3.1.3.2.16.1 Fatigue zu Woche 48****4.3.1.3.2.16.1.1 Fatigue stratifiziert nach Alter**

Tabelle 4-99: Subgruppenanalyse für Fatigue, stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	26 (12,6)	193	28 (14,5)	0,85 [0,48;1,51], p=0,5823
≥40 Jahre	147	24 (16,3)	162	17 (10,5)	1,66 [0,85;3,24], p=0,1339
p-Wert Interaktionstest^a:					0,14
I² (zusätzlich berechnet):					55,2%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.16.1.2 Fatigue stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-100: Subgruppenanalyse für Fatigue, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	29 (16,5)	163	18 (11,0)	1,59 [0,85;2,99], p=0,1504
>350 Zellen/μL	177	21 (11,9)	192	27 (14,1)	0,82 [0,45;1,52], p=0,5310
p-Wert Interaktionstest^a:					0,14
I² (zusätzlich berechnet):					53,7%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.16.2 Fatigue zu Woche 96**4.3.1.3.2.16.2.1 Fatigue stratifiziert nach Alter**

Tabelle 4-101: Subgruppenanalyse für Fatigue, stratifiziert nach Alter (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	26 (12,6)	193	36 (18,7)	0,63 [0,36;1,09], p=0,0983
≥40 Jahre	147	27 (18,4)	162	21 (13,0)	1,51 [0,81;2,81], p=0,1922
p-Wert Interaktionstest^a:					0,04
I² (zusätzlich berechnet):					76,7%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.17 Erkrankungen des Nervensystems (SOC)**4.3.1.3.2.17.1 Erkrankungen des Nervensystems (SOC) zu Woche 48****4.3.1.3.2.17.1.1 Erkrankungen des Nervensystems (SOC) stratifiziert nach Alter**

Tabelle 4-102: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	56 (27,2)	193	44 (22,8)	1,26 [0,80;1,99], p=0,3128
≥40 Jahre	147	35 (23,8)	162	49 (30,2)	0,72 [0,43;1,20], p=0,2048
p-Wert Interaktionstest^a:					0,11
I² (zusätzlich berechnet):					61,8%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.17.1.2 Erkrankungen des Nervensystems (SOC) stratifiziert nach Rasse

Tabelle 4-103: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Rasse (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	69 (27,6)	277	70 (25,3)	1,13 [0,76;1,66], p=0,5447
Nicht-kaukasisch	103	22 (21,4)	78	23 (29,5)	0,65 [0,33;1,28], p=0,2118
p-Wert Interaktionstest^a:					0,17
I² (zusätzlich berechnet):					47,8%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.17.2 Erkrankungen des Nervensystems (SOC) zu Woche 96**4.3.1.3.2.17.2.1 Erkrankungen des Nervensystems (SOC) stratifiziert nach Rasse**

Tabelle 4-104: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Rasse (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	84 (33,6)	277	91 (32,9)	1,03 [0,72;1,49], p=0,8555
Nicht-kaukasisch	103	28 (27,2)	78	29 (37,2)	0,63 [0,34;1,19], p=0,1530
p-Wert Interaktionstest^a:					0,18
I² (zusätzlich berechnet):					43,5%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.18 Kopfschmerzen**4.3.1.3.2.18.1 Kopfschmerzen zu Woche 48****4.3.1.3.2.18.1.1 Kopfschmerzen stratifiziert nach Rasse**

Tabelle 4-105: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Rasse (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	38 (15,2)	277	29 (10,5)	1,53 [0,91;2,57], p=0,1053
Nicht-kaukasisch	103	15 (14,6)	78	15 (19,2)	0,72 [0,33;1,57], p=0,4043
p-Wert Interaktionstest^a:					0,11
I² (zusätzlich berechnet):					60,3%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.18.2 Kopfschmerzen zu Woche 96**4.3.1.3.2.18.2.1 Kopfschmerzen stratifiziert nach CD4-Zellzahl zu Baseline**

Tabelle 4-106: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤350 Zellen/μL	176	38 (21,6)	163	26 (16,0)	1,45 [0,84;2,52], p=0,1863
>350 Zellen/μL	177	21 (11,9)	192	27 (14,1)	0,82 [0,45;1,52], p=0,5310
p-Wert Interaktionstest^a:					0,18
I² (zusätzlich berechnet):					45,2%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.19 Psychiatrische Erkrankungen (SOC)**4.3.1.3.2.19.1 Psychiatrische Erkrankungen (SOC) zu Woche 48****4.3.1.3.2.19.1.1 Psychiatrische Erkrankungen (SOC) stratifiziert nach Geschlecht**

Tabelle 4-107: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Geschlecht (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Männlich	324	64 (19,8)	316	77 (24,4)	0,76 [0,53;1,11], p=0,1597
Weiblich	29	11 (37,9)	39	4 (10,3)	5,35 [1,49;19,19], p=0,0101
p-Wert Interaktionstest^a:					0,004
I² (zusätzlich berechnet):					87,8%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.19.1.2 Psychiatrische Erkrankungen (SOC) stratifiziert nach Ausgangsviruslast

Tabelle 4-108: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Ausgangsviruslast (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	37 (18,2)	214	51 (23,8)	0,71 [0,44;1,15], p=0,1619
>100.000 Kopien/mL	150	38 (25,3)	141	30 (21,3)	1,26 [0,73;2,17], p=0,4142
p-Wert Interaktionstest^a:					0,12
I² (zusätzlich berechnet):					57,5%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.19.1.3 Psychiatrische Erkrankungen (SOC) stratifiziert nach Therapieadhärenz

Tabelle 4-109: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Therapieadhärenz (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<95%	94	18 (19,1)	78	28 (35,9)	0,42 [0,21;0,84], p=0,0147
≥95%	258	57 (22,1)	276	53 (19,2)	1,19 [0,78;1,82], p=0,4096
p-Wert Interaktionstest^a:					0,01
I² (zusätzlich berechnet):					84,2%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.19.2 Psychiatrische Erkrankungen (SOC) zu Woche 96

4.3.1.3.2.19.2.1 Psychiatrische Erkrankungen (SOC) stratifiziert nach Alter

Tabelle 4-110: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Alter (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	62 (30,1)	193	71 (36,8)	0,74 [0,49;1,12], p=0,1571
≥40 Jahre	147	39 (26,5)	162	38 (23,5)	1,18 [0,70;1,97], p=0,5329
p-Wert Interaktionstest^a:					0,17
I² (zusätzlich berechnet):					47,1%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.19.2.2 Psychiatrische Erkrankungen (SOC) stratifiziert nach Geschlecht

Tabelle 4-111: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Geschlecht (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Männlich	324	90 (27,8)	316	104 (32,9)	0,78 [0,56;1,10], p=0,1581
Weiblich	29	11 (37,9)	39	5 (12,8)	4,16 [1,25;13,82], p=0,0202
p-Wert Interaktionstest^a:					0,009
I² (zusätzlich berechnet):					85,4%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.19.2.3 Psychiatrische Erkrankungen (SOC) stratifiziert nach Therapieadhärenz

Tabelle 4-112: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Therapieadhärenz (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<95%	97	30 (30,9)	103	41 (39,8)	0,68 [0,38;1,21], p=0,1906
≥95%	255	71 (27,8)	251	68 (27,1)	1,04 [0,70;1,53], p=0,8498
p-Wert Interaktionstest^a:					0,23
I² (zusätzlich berechnet):					29,8%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.20 Abnorme Träume

Da weder zu Woche 48 noch zu Woche 96 bei mindestens 10% der Patienten abnorme Träume auftraten, werden hier keine Subgruppenanalysen dazu dargestellt.

4.3.1.3.2.21 Schlaflosigkeit

Da weder zu Woche 48 noch zu Woche 96 bei mindestens 10% der Patienten Schlaflosigkeit auftrat, werden hier keine Subgruppenanalysen dazu dargestellt.

4.3.1.3.2.22 Depression

4.3.1.3.2.22.1 Depression zu Woche 48

Da zu Woche 48 nicht bei mindestens 10% der Patienten Depressionen auftraten, werden hier für Depressionen keine Subgruppenanalysen dargestellt.

4.3.1.3.2.22.2 Depression zu Woche 96

4.3.1.3.2.22.2.1 Depression stratifiziert nach Alter

Tabelle 4-113: Subgruppenanalyse für Depression, stratifiziert nach Alter (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	22 (10,7)	193	19 (9,8)	1,09 [0,57;2,09], p=0,7837
≥40 Jahre	147	12 (8,2)	162	22 (13,6)	0,57 [0,27;1,19], p=0,1323
p-Wert Interaktionstest^a:					0,19
I² (zusätzlich berechnet):					42,1%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.22.2.2 Depression stratifiziert nach Geschlecht

Tabelle 4-114: Subgruppenanalyse für Depression, stratifiziert nach Geschlecht (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Männlich	324	30 (9,3)	316	39 (12,3)	0,72 [0,44;1,20], p=0,2101
Weiblich	29	4 (13,8)	39	2 (5,1)	2,96 [0,50;17,41], p=0,2299
p-Wert Interaktionstest^a:					0,13
I² (zusätzlich berechnet):					55,4%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.22.2.3 Depression stratifiziert nach Rasse

Tabelle 4-115: Subgruppenanalyse für Depression, stratifiziert nach Rasse (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	23 (9,2)	277	36 (13,0)	0,68 [0,39;1,18], p=0,1694
Nicht-kaukasisch	103	11 (10,7)	78	5 (6,4)	1,75 [0,58;5,25], p=0,3212
p-Wert Interaktionstest^a:					0,13
I² (zusätzlich berechnet):					55,8%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.23 Erkrankungen der Haut und des Unterhautzellgewebes (SOC)**4.3.1.3.2.23.1 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) zu Woche 48****4.3.1.3.2.23.1.1 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) stratifiziert nach Alter**

Tabelle 4-116: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	49 (23,8)	193	57 (29,5)	0,74 [0,48;1,16], p=0,1946
≥40 Jahre	147	49 (33,3)	162	45 (27,8)	1,30 [0,80;2,11], p=0,2896
p-Wert Interaktionstest^a:					0,10
I² (zusätzlich berechnet):					63,6%
^a zusätzlich berechnet ^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.24 Hautausschlag

Da weder zu Woche 48 noch zu Woche 96 bei mindestens 10% der Patienten Hautausschlag auftrat, werden hier keine Subgruppenanalysen dazu dargestellt.

4.3.1.3.2.25 Augenerkrankungen (SOC)**4.3.1.3.2.25.1 Augenerkrankungen (SOC) zu Woche 48****4.3.1.3.2.25.1.1 Augenerkrankungen (SOC) stratifiziert nach Alter**

Tabelle 4-117: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Alter (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	9 (4,4)	193	56 (29,0)	0,11 [0,05;0,23], p<0,00001
≥40 Jahre	147	9 (6,1)	162	22 (13,6)	0,42 [0,18;0,93], p=0,0334
p-Wert Interaktionstest^a:					0,02
I² (zusätzlich berechnet):					81,9%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.25.1.2 Augenerkrankungen (SOC) stratifiziert nach Rasse

Tabelle 4-118: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Rasse (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	15 (6,0)	277	58 (20,9)	0,24 [0,13;0,44], p<0,00001
Nicht-kaukasisch	103	3 (2,9)	78	20 (25,6)	0,09 [0,02;0,31], p=0,0001
p-Wert Interaktionstest^a:					0,15
I² (zusätzlich berechnet):					51,5%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.25.1.3 Augenerkrankungen (SOC) stratifiziert nach Ausgangsviruslast

Tabelle 4-119: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Ausgangsviruslast (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	203	7 (3,4)	214	47 (22,0)	0,13 [0,06;0,29], p<0,00001
>100.000 Kopien/mL	150	11 (7,3)	141	31 (22,0)	0,28 [0,14;0,58], p=0,0007
p-Wert Interaktionstest^a:					0,16
I² (zusätzlich berechnet):					50,1%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.25.1.4 Augenerkrankungen (SOC) stratifiziert nach Therapieadhärenz

Tabelle 4-120: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Therapieadhärenz (Woche 48)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<95%	94	2 (2,1)	78	17 (21,8)	0,08 [0,02;0,35], p=0,0009
≥95%	258	16 (6,2)	276	61 (22,1)	0,23 [0,13;0,42], p<0,00001
p-Wert Interaktionstest^a:					0,18
I² (zusätzlich berechnet):					43,7%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.25.2 Augenerkrankungen (SOC) zu Woche 96**4.3.1.3.2.25.2.1 Augenerkrankungen (SOC) stratifiziert nach Alter**

Tabelle 4-121: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Alter (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<40 Jahre	206	16 (7,8)	193	56 (29,0)	0,21 [0,11;0,37], p<0,00001
≥40 Jahre	147	12 (8,2)	162	29 (17,9)	0,41 [0,20;0,83], p=0,0138
p-Wert Interaktionstest^a:					0,15
I² (zusätzlich berechnet):					51,6%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.25.2.2 Augenerkrankungen (SOC) stratifiziert nach Rasse

Tabelle 4-122: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Rasse (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
Kaukasisch	250	25 (10,0)	277	65 (23,5)	0,36 [0,22;0,60], p=0,0001
Nicht-kaucasisch	103	3 (2,9)	78	20 (25,6)	0,09 [0,02;0,31], p=0,0001
p-Wert Interaktionstest^a:					0,04
I² (zusätzlich berechnet):					76,7%
^a zusätzlich berechnet ^b Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

4.3.1.3.2.25.2.3 Augenerkrankungen (SOC) stratifiziert nach Therapieadhärenz

Tabelle 4-123: Subgruppenanalyse für Augenerkrankungen (SOC), stratifiziert nach Therapieadhärenz (Woche 96)

Studie S103 Subgruppe	EVG/COBI/FTC/TDF		ATV/r+FTC/TDF		OR ^a [95%-KI] ^a , p-Wert ^{a,b}
	N	n (%)	N	n (%)	
<95%	97	4 (4,1)	103	25 (24,3)	0,13 [0,04;0,40], p=0,0003
≥95%	255	24 (9,4)	251	60 (23,9)	0,33 [0,20;0,55], p<0,0001
p-Wert Interaktionstest^a:					0,14
I² (zusätzlich berechnet):					53,1%
^a zusätzlich berechnet					
^b Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

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.

Der Nutzen bzw. Zusatznutzen der Behandlung mit Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil (EVG/COBI/FTC/TDF) gegenüber der zweckmäßigen Vergleichstherapie ATV/r+FTC/TDF wird hier anhand einer randomisierten, aktivkontrollierten Studie gezeigt, welche EVG/COBI/FTC/TDF in der zugelassenen Dosierung mit der zweckmäßigen Vergleichstherapie, Ritonavir-geboostertes Atazanavir in Kombination mit FTC/TDF, vergleicht.

Die in der eingeschlossenen Studie untersuchte Studienpopulation besteht aus antiretroviral nicht-vorbehandelten erwachsenen Patienten mit einer HIV-1-Infektion. Aufgrund der Übertragbarkeit der Daten dieser Patienten auf vorbehandelte HIV-1 infizierte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, treffen die Ergebnisse dieser Studie auch auf die im vorliegenden Modul 4B behandelte Population zu.

In die Bewertung des Zusatznutzens sind folgende Endpunkte eingeflossen: Virologisches Ansprechen (Snapshot und TLOVR), Zeit bis zum Verlust des virologischen Ansprechens, Veränderung der CD4-Zellzahl, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschter Ereignisse, Studienabbrüche aufgrund unerwünschter Ereignisse, unerwünschte Ereignisse Grad 3-4, wichtige unerwünschte Ereignisse (Hautausschläge,

neurologische & psychiatrische Ereignisse), unerwünschte Ereignisse von Interesse (renale Ereignisse, Knochenfrakturen) und häufigste unerwünschte Ereignisse (Erkrankungen des Gastrointestinaltrakts [SOC], Allgemeine Erkrankungen und Beschwerden am Verabreichungsort [SOC], Erkrankungen des Nervensystems [SOC], Psychiatrische Erkrankungen [SOC], Erkrankungen der Haut und des Unterhautzellgewebes [SOC], Augenerkrankungen [SOC]).

Virologisches Ansprechen

Snapshot-Algorithmus

Unter Verwendung des Snapshot-Algorithmus und der ITT-Population kann zu Woche 48 für den primären Endpunkt „**Virologisches Ansprechen (Snapshot)**“ die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze gezeigt werden (RD [95%-KI]: 3,0 [-1,9;7,8]). Diese Nichtunterlegenheit gegenüber ATV/r+FTC/TDF besteht auch noch nach 96 Wochen der Behandlung. Die RD [95%-KI] beträgt zu Woche 96 1,1 [-4,5;6,7]. Statistisch signifikante Unterschiede zwischen den beiden Behandlungen bestehen an keinem der beiden Zeitpunkte (OR [95%-KI]: Woche 48: 1,30 [0,82;2,06], $p=0,22$; Woche 96: 1,08 [0,73; 1,59], $p=0,70$).

In der PP-Population kann zu Woche 48 konsistent zur ITT-Population die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze gezeigt werden (RD [95%-KI]: -0,1 [-2,6;2,4]). Und mit einer RD [95%-KI] von 0,8 [-2,2;3,8] besteht diese Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF auch noch nach 96 Wochen Behandlung.

Es besteht zu Woche 96 jeweils ein Hinweis auf Effektmodifikation durch Alter ($p=0,12$; $I^2=60,7\%$), Geschlecht ($p=0,10$; $I^2=58,9\%$) und Therapieadhärenz ($p=0,092$; $I^2=61,7\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

TLOVR-Algorithmus

Mittels TLOVR-Algorithmus zeigen sich ähnliche Ergebnisse. Auch hier kann für den Endpunkt „**Virologisches Ansprechen (TLOVR)**“ die Nichtunterlegenheit zur 12%-Grenze von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zu beiden Zeitpunkten gezeigt werden. Die RD [95%-KI] zu Woche 48 beträgt 1,6 [-3,6;6,8] und zu Woche 96 ist die RD [95%-KI] 4,0 [-2,0;9,9]. Statistisch signifikante Unterschiede zwischen den beiden Behandlungen können jedoch auch hier zu keinem der beiden Zeitpunkte beobachtet werden (OR [95%-KI]: Woche 48: 1,11 [0,73;1,69], $p=0,55$; Woche 96: 1,27 [0,88;1,82], $p=0,19$).

Zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,19$; $I^2=42,3\%$). Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch Therapieadhärenz ($p=0,06$; $I^2=72,7\%$). In der Patientengruppe, die zu $<95\%$ adhärenz waren,

beträgt das OR [95%-KI] 0,80 [0,44;1,45] mit $p=0,4568$ und in der Gruppe von Patienten, die eine Therapieadhärenz $\geq 95\%$ aufwiesen, ergibt sich das OR [95%-KI] zu 1,69 [1,04;2,74] mit $p=0,0337$. Es besteht somit in der Gruppe der Patienten, die das Medikament zu $\geq 95\%$ adhärenz einnahmen, ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF hinsichtlich des virologischen Ansprechens (TLOVR).

Zeit bis zum Verlust des virologischen Ansprechens

Zu Woche 48 zeigt sich hinsichtlich **Zeit bis zum Verlust des virologischen Ansprechens** kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen. Das HR [95%-KI] ergibt sich zu 0,87 [0,59;1,27] mit $p=0,48$. Auch an Woche 96 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (HR [95%-KI]: 0,81 [0,58;1,12], $p=0,21$).

Es besteht zu Woche 48 ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,19$; $I^2=42,2\%$). Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Außerdem besteht ein Beleg für Effektmodifikation durch Therapieadhärenz ($p=0,04$; $I^2=75,2\%$). Für Patienten, die die Studienmedikation zu $< 95\%$ adhärenz einnahmen, beträgt das HR [95%-KI] 1,19 [0,72;1,95] mit $p=0,49$. Für Patienten, die eine Therapieadhärenz von $\geq 95\%$ aufwiesen, ergibt sich das HR [95%-KI] zu 0,60 [0,39;0,94] mit $p=0,025$. Es besteht somit für Patienten mit hoher Therapieadhärenz ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF.

Veränderung der CD4-Zellzahl

Zu Woche 48 zeigt sich, unter Verwendung der ITT-Population, hinsichtlich der **Veränderung der CD4-Zellzahl** kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen. Die adj. MWD [95%-KI] beträgt -6 [-31;18] mit $p=0,61$. Auch zu Woche 96 besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (adj. MWD [95%-KI]: -8 [-35;19], $p=0,56$).

Entsprechend kann in der PP-Population weder zu Woche 48 (adj. MWD [95%-KI]: 0 [-25;24], $p=0,97$) noch zu Woche 96 (adj. MWD: -6 [-34;21], $p=0,65$) ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen gezeigt werden.

Es bestehen zu Woche 48 Hinweise auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,07$; $I^2=68,5\%$) und Therapieadhärenz ($p=0,0530$; $I^2=73,3\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus. Zu Woche 96 gibt es Belege für Effektmodifikation durch Geschlecht ($p=0,04$; $I^2=76,6\%$) und CD4-Zellzahl zu Baseline ($p=0,02$; $I^2=81,5\%$). In der Gruppe der männlichen Patienten, beträgt die MWD [95%-KI] 7,00 [-21,43;35,43] mit $p=0,6290$ und unter den Frauen beträgt sie -107,00 [-211,21;-2,79] mit $p=0,0442$. Es besteht somit unter den Frauen ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF. In der Subgruppenanalyse, stratifiziert nach CD4-Zellzahl zu Baseline, ergibt sich bei Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL

die MWD [95%-KI] zu -40,00 [-77,16;-2,84] mit $p=0,0350$, was einen statistisch signifikanten Unterschied zuungunsten von EVG/COBI/FTC/TDF zeigt. In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl >350 Zellen/ μL aufwiesen, beträgt die MWD [95%-KI] 25,00 [-15,32;65,32], $p=0,2274$.

Mortalität

Bis Woche 48 starben in der vorliegenden Studie S103 insgesamt 3 Patienten, welche alle im ATV/r+FTC/TDF-Arm waren. Das OR [95%-KI] beträgt 0,14 [0,01;2,77] mit $p=0,1980$, woraus sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen hinsichtlich der **Mortalitätsrate** ableiten lässt. Da bis Woche 96 kein weiterer Patient verstarb, ergeben sich zu diesem Zeitpunkt identische Ergebnisse.

Subgruppenanalysen liefern keine Hinweise oder Belege auf Effektmodifikation.

Gesamtrate unerwünschter Ereignisse

Zu Woche 48 besteht hinsichtlich der **Gesamtrate unerwünschter Ereignisse** kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,71 [0,40;1,26], $p=0,2423$). Und auch zu Woche 96 kann kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (OR [95%-KI]: 0,69 [0,33;1,47], $p=0,3376$).

Es besteht zu Woche 48 ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,17$; $I^2=47,1\%$). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Gesamtrate schwerwiegender unerwünschter Ereignisse

Zu Woche 48 besteht hinsichtlich des Auftretens **schwerwiegender unerwünschter Ereignisse** kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,83 [0,48;1,43], $p=0,5043$). Zu Woche 96 kann trotz numerisch geringerer Rate an schwerwiegenden unerwünschten Ereignissen im EVG/COBI/FTC/TDF-Arm kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen abgeleitet werden (OR [95%-KI]: 0,65 [0,41;1,03], $p=0,0684$).

Subgruppenanalysen zu Woche 48 zeigen Hinweise auf Effektmodifikation durch Alter ($p=0,0539$; $I^2=73,1\%$) und Ausgangsviruslast ($p=0,08$; $I^2=67,8\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus. Zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,19$; $I^2=41,5\%$). In der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL beträgt das OR [95%-KI] 0,84 [0,46;1,53] mit $p=0,5731$ und für Patienten mit Ausgangsviruslast >100.000 Kopien/mL ergibt sich das OR [95%-KI] zu 0,45 [0,21;0,94] mit $p=0,0325$. Somit besteht in der Gruppe mit Ausgangsviruslast >100.000 Kopien/mL ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF.

Therapieabbrüche aufgrund unerwünschter Ereignisse

Es besteht hinsichtlich **Therapieabbrüchen aufgrund unerwünschter Ereignisse** zu Woche 48 kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,72 [0,35;1,48], $p=0,3689$). Und auch zu Woche 96 sind die Abbruchraten der Therapie aufgrund unerwünschter Ereignisse in den beiden Behandlungsarmen ähnlich (OR [95%-KI]: 0,71 [0,36;1,39], $p=0,3151$).

Eine Subgruppenanalyse zu Woche 48 zeigt einen Hinweis auf Effektmodifikation durch Rasse ($p=0,10$; $I^2=64,1\%$). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Unerwünschte Ereignisse Grad 3-4

Zu Woche 48 besteht hinsichtlich **unerwünschter Ereignisse Grad 3-4** kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,93 [0,60;1,45], $p=0,7607$). Und auch zu Woche 96 lässt sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen ableiten (OR [95%-KI]: 0,89 [0,60;1,33], $p=0,5648$).

Zu Woche 48 besteht ein Beleg für Effektmodifikation durch Ausgangsviruslast ($p=0,0002$; $I^2=92,9\%$). In der Gruppe der Patienten, die zu Baseline eine Viruslast ≤ 100.000 Kopien/mL aufwiesen, beträgt das OR [95%-KI] 1,95 [1,08;3,52] mit $p=0,0262$, was einen statistisch signifikanten Unterschied zuungunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL zeigt sich jedoch mit einem OR [95%-KI] von 0,32 [0,15;0,67] und $p=0,0025$ ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF. Außerdem besteht ein Hinweis auf Effektmodifikation durch Therapieadhärenz ($p=0,15$; $I^2=52,0\%$). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,002$; $I^2=90,1\%$). In der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL beträgt das OR [95%-KI] 1,48 [0,89;2,46] mit $p=0,1268$ und für Patienten mit Ausgangsviruslast >100.000 Kopien/mL ergibt sich das OR [95%-KI] zu 0,37 [0,18;0,74] mit $p=0,0050$. Es besteht somit in der Gruppe der Patienten mit höherer Ausgangsviruslast ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF.

Hautausschläge

Zu Woche 48 besteht hinsichtlich des Auftretens von **Hautausschlägen** kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,99 [0,67;1,45], $p=1,00$). Und auch zu Woche 96 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,93 [0,65;1,32], $p=0,72$).

Zu Woche 48 bestehen Hinweise auf Effektmodifikation durch Alter ($p=0,17$; $I^2=47,6\%$), Geschlecht ($p=0,19$; $I^2=41,6\%$) und CD4-Zellzahl zu Baseline ($p=0,17$; $I^2=46,6\%$). Außerdem

gibt es einen Beleg für Effektmodifikation durch Rasse ($p=0,0494$; $I^2=74,1\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

Neurologische & psychiatrische Ereignisse

Es besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen den Behandlungen hinsichtlich des Auftretens **neurologischer & psychiatrischer Ereignisse** (OR [95%-KI]: 0,94 [0,69;1,27], $p=0,69$). Zu Woche 96 kann auch kein statistisch signifikanter Unterschied zwischen den Behandlungen festgestellt werden (OR [95%-KI]: 0,95 [0,71;1,28], $p=0,76$).

Zu Woche 48 ergeben sich Hinweise auf Effektmodifikation durch Geschlecht ($p=0,15$; $I^2=51,2\%$) und Therapieadhärenz ($p=0,12$; $I^2=58,1\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

Es besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Geschlecht ($p=0,14$; $I^2=53,4\%$). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Renale Ereignisse

Zu Woche 48 traten keine **renalen Ereignisse** auf, weswegen keine Schätzer berechenbar sind. Bis Woche 96 traten vier renale Ereignisse im EVG/COBI/FTC/TDF-Arm auf. Das daraus resultierende OR [95%-KI] von 9,15 [0,49;170,67] mit $p=0,061$ zeigt jedoch keine statistische Signifikanz.

Es bestehen keine Hinweise oder Belege auf Effektmodifikation.

Knochenfrakturen

Zu Woche 48 besteht zwischen den Behandlungen kein statistisch signifikanter Unterschied hinsichtlich des Auftretens von **Knochenfrakturen** (OR [95%-KI]: 0,50 [0,12;2,01], $p=0,51$). Zu Woche 96 ist der Unterschied zwischen den beiden Behandlungen allerdings statistisch signifikant zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,28 [0,09;0,86], $p=0,029$).

Zu Woche 48 bestehen Hinweise auf Effektmodifikation durch Alter ($p=0,19$; $I^2=42,1\%$), Geschlecht ($p=0,16$; $I^2=48,2\%$) und CD4-Zellzahl zu Baseline ($p=0,07$; $I^2=69,0\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus. Zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch Geschlecht ($p=0,09$; $I^2=65,7\%$). In der Gruppe der männlichen Patienten ergibt sich ein OR [95%-KI] von 0,20 [0,06;0,71] mit $p=0,0125$, was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF anzeigt, und in der

Gruppe der Frauen beträgt das OR [95%-KI] 4,16 [0,16;105,81] mit $p=0,3882$. Da in der Gruppe der Frauen nur sehr wenig Patienten eingeschlossen sind und lediglich ein Patient eine Knochenfraktur erlitt, ist diese Interaktion mit Vorsicht zu interpretieren.

Erkrankungen des Gastrointestinaltrakts (SOC)

Hinsichtlich des Auftretens von **Erkrankungen des Gastrointestinaltrakts (SOC)** besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,88 [0,66;1,19], $p=0,4103$). Auch zu Woche 96 kann kein statistisch signifikanter Unterschied zwischen den Behandlungen gezeigt werden (OR [95%-KI]: 0,79 [0,58;1,07], $p=0,1279$).

Zu Woche 48 gibt es einen Beleg für Effektmodifikation durch Rasse ($p=0,01$; $I^2=84,1\%$). In der Gruppe der Kaukasier beträgt das OR [95%-KI] 1,12 [0,79;1,58] mit $p=0,5288$ und in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,46 [0,25;0,84] mit $p=0,0113$. Es besteht somit unter den nicht-kaukasischen Patienten hinsichtlich der Vermeidung von Erkrankungen des Gastrointestinaltrakts ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF. Zu Woche 96 besteht ein Beleg für Effektmodifikation durch Rasse ($p=0,02$; $I^2=82,5\%$). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 1,00 [0,70;1,42] mit $p=0,9823$ und in der Gruppe der nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,42 [0,23;0,78] mit $p=0,0057$. Somit besteht ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF unter den nicht-kaukasischen Patienten. Ein weiterer Hinweis auf Effektmodifikation kann durch Ausgangsviruslast gefunden werden ($p=0,13$; $I^2=55,8$). In der Gruppe der Patienten, die zu Baseline eine Viruslast ≤ 100.000 Kopien/mL aufwiesen, beträgt das OR [95%-KI] 0,65 [0,44;0,97] mit $p=0,0328$, was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF hinsichtlich der Vermeidung gastrointestinaler Erkrankungen anzeigt. In der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL besteht kein Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,04 [0,65;1,67], $p=0,8599$). Außerdem besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,07$; $I^2=69,2\%$). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL beträgt das OR [95%-KI] 1,06 [0,68;1,63] mit $p=0,8029$ und in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL ergibt sich das OR [95%-KI] zu 0,61 [0,40;0,92] mit $p=0,0202$. Es besteht also in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF hinsichtlich der Vermeidung gastrointestinaler Erkrankungen.

Diarrhoe

Diarrhoe trat im EVG/COBI/FTC/TDF-Arm numerisch geringer auf als im ATV/r+FTC/TDF-Arm. Allerdings besteht keine statistische Signifikanz (OR [95%-KI]: 0,74 [0,53;1,05], $p=0,0891$). Der Unterschied wird zu Woche 96 noch etwas größer, es lässt sich jedoch auch hier keine statistische Signifikanz ableiten (OR [95%-KI]: 0,73 [0,53;1,01], $p=0,0611$).

Es besteht zu Woche 48 ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,15$; $I^2=51,9\%$). In der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL beträgt das OR [95%-KI] 0,59 [0,37;0,94] mit $p=0,0257$, was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,99 [0,59;1,67], $p=0,9627$). Außerdem besteht ein Beleg für Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,01$; $I^2=83,2\%$). In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl ≤ 350 Zellen/ μL aufwiesen, beträgt das OR [95%-KI] 1,18 [0,71;1,95] mit $p=0,5280$. In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,49 [0,30;0,80], $p=0,0042$). Zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,18$; $I^2=44,3\%$). In der Gruppe der Patienten, die eine CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL aufwiesen, beträgt das OR [95%-KI] 0,93 [0,57;1,49] mit $p=0,7546$. In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL ergibt sich das OR [95%-KI] zu 0,59 [0,37;0,93] mit $p=0,0240$. Es besteht also ein statistisch signifikanter Unterschied zwischen den Behandlungen, der zugunsten von EVG/COBI/FTC/TDF ausgeprägt ist.

Übelkeit

Weder zu Woche 48 (OR [95%-KI]: 1,03 [0,71;1,49], $p=0,8952$) noch zu Woche 96 (OR [95%-KI]: 0,99 [0,69;1,42], $p=0,9574$) besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen hinsichtlich der Vermeidung von **Übelkeit**.

Zu Woche 48 besteht ein Hinweis auf Effektmodifikation durch Alter ($p=0,0541$; $I^2=73,1\%$). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Ein weiterer Hinweis auf Effektmodifikation besteht durch CD4-Zellzahl zu Baseline ($p=0,14$; $I^2=53,3\%$). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL beträgt das OR [95%-KI] 1,33 [0,79;2,25] mit $p=0,2791$ und in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL ergibt sich das OR [95%-KI] zu 0,76 [0,44;1,30] mit $p=0,3207$. Zu Woche 96 besteht ein Beleg für Effektmodifikation durch Alter ($p=0,04$; $I^2=76,8\%$) sowie ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,16$; $I^2=48,5\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)

Zu Woche 48 besteht zwischen den Behandlungen hinsichtlich **allgemeiner Erkrankungen und Beschwerden am Verabreichungsort (SOC)** kein statistisch signifikanter Unterschied (OR [95%-KI]: 1,04 [0,74;1,45], $p=0,8297$). Und auch zu Woche 96 kann kein statistisch signifikanter Unterschied zwischen den Behandlungen abgeleitet werden (OR [95%-KI]: 0,81 [0,59;1,11], $p=0,1844$).

Es bestehen keine Hinweise oder Belege auf Effektmodifikation.

Fatigue

Die Ereignisrate von **Fatigue** war zu Woche 48 in beiden Behandlungsarmen ähnlich, woraus sich kein statistisch signifikanter Unterschied zwischen den Behandlungen ableiten lässt (OR [95%-KI]: 1,14 [0,74;1,75], $p=0,5615$). Zu Woche 96 kann ebenso kein statistisch signifikanter Unterschied zwischen den Behandlungen festgestellt werden (OR [95%-KI]: 0,92 [0,61;1,39], $p=0,7020$).

Zu Woche 48 besteht jeweils ein Hinweis auf Effektmodifikation durch Alter ($p=0,14$; $I^2=55,2\%$) und CD4-Zellzahl zu Baseline ($p=0,14$; $I^2=53,7\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

Erkrankungen des Nervensystems (SOC)

Zu Woche 48 besteht hinsichtlich **Erkrankungen des Nervensystems (SOC)** kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,98 [0,70;1,37], $p=0,8991$). Auch zu Woche 96 ist der Unterschied zwischen den Behandlungen nicht statistisch signifikant (OR [95%-KI]: 0,91 [0,66;1,25], $p=0,5565$).

Es besteht zu Woche 48 ein Hinweis auf Effektmodifikation durch Alter ($p=0,11$; $I^2=61,8\%$). Außerdem besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Rasse ($p=0,18$; $I^2=43,5\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

Kopfschmerzen

Hinsichtlich des Auftretens von **Kopfschmerzen** besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,25 [0,81;1,92], $p=0,3114$). Auch zu Woche 96 ist der Unterschied zwischen den Behandlungen statistisch nicht signifikant (OR [95%-KI]: 1,14 [0,76;1,71], $p=0,5156$).

Es besteht zu Woche 48 ein Hinweis auf Effektmodifikation durch Rasse ($p=0,11$; $I^2=60,3\%$), und zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,18$; $I^2=45,2\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

Psychiatrische Erkrankungen (SOC)

Zu Woche 48 besteht hinsichtlich des Auftretens **psychiatrischer Erkrankungen (SOC)** kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,91 [0,64;1,30], $p=0,6143$). Und auch zu Woche 96 kann kein statistisch signifikanter Unterschied zwischen den Behandlungen abgeleitet werden (OR [95%-KI]: 0,90 [0,66;1,25], $p=0,5423$).

Zu Woche 48 gibt es einen Beleg für Effektmodifikation durch Geschlecht ($p=0,004$; $I^2=87,8\%$). In der Gruppe der Männer beträgt das OR [95%-KI] 0,76 [0,53;1,11] mit $p=0,1597$ und in der Gruppe der weiblichen Patienten ergibt sich das OR [95%-KI] zu 5,35 [1,49;19,19] mit $p=0,0101$. In der Gruppe der weiblichen Patienten besteht somit ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF. Allerdings ist diese Subgruppenanalyse, aufgrund der sehr geringen Anzahl an Frauen in der Studie, mit Vorsicht zu interpretieren. Außerdem besteht ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,12$; $I^2=57,5\%$). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Des Weiteren besteht ein Beleg für Effektmodifikation durch Therapieadhärenz ($p=0,01$; $I^2=84,2\%$). In der Gruppe der Patienten, die zu $<95\%$ adhären waren, beträgt das OR [95%-KI] 0,42 [0,21;0,84] mit $p=0,0147$ und in der Gruppe der zu $\geq 95\%$ adhären Patienten ergibt sich das OR [95%-KI] zu 1,19 [0,78;1,82] mit $p=0,4096$. In der Gruppe der Patienten, die weniger adhären waren, besteht also ein statistisch signifikanter Vorteil von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF. Zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch Alter ($p=0,17$; $I^2=47,1\%$). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus. Des Weiteren besteht ein Beleg für Effektmodifikation durch Geschlecht ($p=0,009$; $I^2=85,4\%$). In der Gruppe der männlichen Patienten beträgt das OR [95%-KI] 0,78 [0,56;1,10] mit $p=0,1581$ und für die weiblichen Patienten ergibt sich ein OR [95%-KI] von 4,16 [1,25;13,82] mit $p=0,0202$. In der Gruppe der Frauen besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF. Allerdings ist diese Subgruppenanalyse aufgrund der sehr geringen Anzahl an Frauen in der Studie mit Vorsicht zu interpretieren.

Abnorme Träume

Es besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen den Behandlungen hinsichtlich des Auftretens **abnormer Träume** (OR [95%-KI]: 0,86 [0,39;1,88], $p=0,7005$). Zu Woche 96 traten abnorme Träume nicht bei mindestens 10% der Patienten in einem der Behandlungsarme auf.

Für abnorme Träume wurden keine Subgruppenanalysen durchgeführt, da sie in Studie S103 nicht bei $\geq 10\%$ der Patienten in einem der Behandlungsarme auftraten.

Schlaflosigkeit

Zu Woche 48 besteht hinsichtlich des Auftretens von **Schlaflosigkeit** kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,68 [0,91;3,08], $p=0,0958$). Zu Woche 96 trat Schlaflosigkeit nicht bei mindestens 10% der Patienten in einem der Behandlungsarme auf.

Für Schlaflosigkeit wurden keine Subgruppenanalysen durchgeführt, da sie in Studie S103 nicht bei $\geq 10\%$ der Patienten in einem der Behandlungsarme auftrat.

Depression

Zu Woche 48 ist der Unterschied zwischen den Behandlungen hinsichtlich des Auftretens von **Depressionen** statistisch nicht signifikant (OR [95%-KI]: 0,87 [0,47;1,61], $p=0,6508$). Und auch zu Woche 96 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,82 [0,50;1,32], $p=0,4077$).

Zu Woche 96 bestehen Hinweise auf Effektmodifikation durch Alter ($p=0,19$; $I^2=42,1\%$), Geschlecht ($p=0,13$; $I^2=55,4\%$) und Rasse ($p=0,13$; $I^2=55,8\%$). Da jedoch in keiner der in diesen Subgruppenanalysen betrachteten Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

Erkrankungen der Haut und des Unterhautzellgewebes (SOC)

Zu Woche 48 gibt es hinsichtlich des Auftretens von **Erkrankungen der Haut und des Unterhautzellgewebes (SOC)** keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,95 [0,69;1,32], $p=0,7743$). Mit einem OR [95%-KI] von 0,82 [0,60;1,11] und $p=0,2008$ besteht auch zu Woche 96 kein statistisch signifikanter Unterschied zwischen den Behandlungen.

Zu Woche 48 besteht ein Hinweis auf Effektmodifikation durch Alter ($p=0,10$; $I^2=63,6\%$). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Hautausschlag

Zu Woche 48 besteht hinsichtlich des Auftretens von **Hautausschlag** statistisch kein signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,15 [0,64;2,09], $p=0,6365$). Auch zu Woche 96 ist der Unterschied zwischen den Behandlungen statistisch nicht signifikant (OR [95%-KI]: 0,94 [0,55;1,59], $p=0,8049$).

Für Hautausschlag wurden keine Subgruppenanalysen durchgeführt, da er in Studie S103 nicht bei $\geq 10\%$ der Patienten in einem der Behandlungsarme auftrat.

Augenerkrankungen (SOC)

Zu Woche 48 besteht hinsichtlich des Auftretens von **Augenerkrankungen (SOC)** ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,19 [0,11;0,33], $p<0,00001$), der zugunsten von EVG/COBI/FTC/TDF ausgeprägt ist. Dieser statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF besteht auch noch nach 96 Wochen Behandlung (OR [95%-KI]: 0,27 [0,17;0,43], $p<0,00001$).

Zu Woche 48 bestehen Hinweise auf Effektmodifikation durch Rasse ($p=0,15$; $I^2=51,5\%$), Ausgangsviruslast ($p=0,16$; $I^2=50,1\%$) und Therapieadhärenz ($p=0,18$; $I^2=43,7\%$). Außerdem besteht ein Beleg für Effektmodifikation durch Alter ($p=0,02$; $I^2=81,9\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] mit 0,11 [0,05;0,23] mit $p<0,00001$ und unter den älteren Patienten ergibt sich das OR [95%-KI] zu 0,42 [0,18;0,93] mit $p=0,0334$. In

der Gruppe der kaukasischen Patienten ist das OR [95%-KI] 0,24 [0,13;0,44] mit $p < 0,00001$ und für die nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,09 [0,02;0,31] mit $p = 0,0001$. In der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL ergibt sich das OR [95%-KI] zu 0,13 [0,06;0,29] mit $p < 0,00001$ und in der Gruppe der Patienten mit Ausgangsviruslast > 100.000 Kopien/mL ist das OR [95%-KI] 0,28 [0,14;0,58] mit $p = 0,0007$. Über alle Subgruppen hinweg besteht ein statistisch signifikanter Vorteil von EVG/COBI/FTC/TDF, welcher sich in manchen Subgruppen lediglich quantitativ von der jeweils anderen Subgruppe unterscheidet.

Zu Woche 96 bestehen Hinweise auf Effektmodifikation durch Alter ($p = 0,15$; $I^2 = 51,6\%$) und Therapieadhärenz ($p = 0,14$; $I^2 = 53,1\%$). Außerdem besteht ein Beleg für Effektmodifikation durch Rasse ($p = 0,04$; $I^2 = 76,7\%$). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,21 [0,11;0,37] mit $p < 0,00001$ und in der Gruppe der mindestens 40-jährigen Patienten ergibt sich das OR [95%-KI] zu 0,41 [0,20;0,83] mit $p = 0,0138$. In der Gruppe der kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,36 [0,22;0,60] mit $p = 0,0001$ und unter den nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,09 [0,02;0,31] mit $p = 0,0001$. Für Patienten, die zu $< 95\%$ adhärenz waren, beträgt das OR [95%-KI] 0,13 [0,04;0,40] mit $p = 0,0003$ und für Patienten, die eine Therapieadhärenz $\geq 95\%$ aufwiesen ergibt sich das OR [95%-KI] zu 0,33 [0,20;0,55] mit $p < 0,0001$. Auch hier kann konsistent in allen Subgruppen ein statistisch signifikanter Vorteil von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF beobachtet werden.

Skleraler Ikterus

Hinsichtlich des Auftretens von skleralem Ikterus besteht nach 48 Wochen ein statistisch signifikanter Unterschied zwischen den Behandlungen, der zugunsten von EVG/COBI/FTC/TDF ausgeprägt ist (OR [95%-KI]: 0,03 [0,01;0,14], $p < 0,00001$). Dieser statistisch signifikante Unterschied besteht ebenfalls noch nach 96 Wochen Behandlung (OR [95%-KI]: 0,03 [0,01;0,14], $p < 0,00001$).

Es bestehen keine Hinweise oder Belege auf Effektmodifikation.

Generell sind Subgruppenanalysen, stratifiziert nach Therapieadhärenz, kritisch zu betrachten, da naturgemäß nicht zu Baseline stratifiziert werden konnte, sondern sich die Zuteilung zu den Subgruppen erst durch die erhobenen Daten ergeben hat. Da die Behandlung und möglicherweise auch andere Faktoren einen Einfluss auf die Subgruppenbildung hatten, ist die Aussagekraft dieser Subgruppenanalyse als niedrig einzuschätzen. Die gefundenen Ergebnisse zeigen zumeist nur, dass eine höhere Therapieadhärenz zu mehr Nebenwirkungen führt, womit gegebenenfalls vorliegende Unterschiede im Nebenwirkungsprofil signifikant werden können. Da die Therapieadhärenz vor dem Hintergrund einer möglichen Resistenzentwicklung in jedem Fall unabhängig vom antiretroviralen Regime bei idealerweise 100% liegen muss, sind die gefundenen Daten klinisch von minderer Bedeutung. Darüber hinaus sind die in der vorliegenden Nutzenbewertung betrachteten Daten auf die klinische Praxis und den tatsächlichen Einsatz von EVG/COBI/FTC/TDF nicht übertragbar, da die

Patienten aufgrund des double-dummy Designs der hier eingeschlossenen Studien täglich jeweils eine Behandlung mit aktivem Wirkstoff und eine Behandlung mit Placebo einnehmen mussten. Alle Patienten nahmen somit mehrere Tabletten täglich ein, wodurch eine Übertragbarkeit der Adhärenz-Daten auf die Praxis, in der EVG/COBI/FTC/TDF-Patienten nur eine Tablette täglich und ATV/r+FTC/TDF-Patienten drei Tabletten täglich (ATV, RTV und FTC/TDF) einnehmen müssen, nicht gegeben. Aus diesem Grund wurden die Subgruppenanalysen zwar vollständig dargestellt, finden jedoch keine weitere Berücksichtigung in der Bewertung und Nutzenableitung.

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.

Da die Informationsbeschaffung zu vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, ergeben hat, dass es keine Studien des pharmazeutischen Unternehmers gibt, die sich für einen indirekten Vergleich geeignet hätten und aufgrund der Übertragbarkeit der Daten der nicht-vorbehandelten HIV-infizierten Patienten auf vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, wird im Weiteren nur die Informationsbeschaffung zu den nicht-vorbehandelten HIV-infizierten Patienten dargestellt.

Es wird ein indirekter Vergleich durchgeführt:

- Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil (EVG/COBI/FTC/TDF) vs. Raltegravir + Emtricitabin/ Tenofoviridisoproxil (RAL+FTC/TDF)

Der indirekte Vergleich wird über einen geeigneten Brückenkomparator durchgeführt. Die Ergebnisse der Informationsbeschaffung, wie auch die Ergebnisse des indirekten Vergleichs selbst sind gemeinsam in den nachfolgenden Abschnitten dargestellt.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

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

- Studien des pharmazeutischen Unternehmers

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

4.3.2.1.1 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.

Es werden nur Studien des pharmazeutischen Unternehmers mit dem zu bewertenden Arzneimittel berücksichtigt, die entweder mit dem Single-Tablet-Regimen (EVG/COBI/FTC/TDF) oder mit den einzelnen, aber allen vier Komponenten in Kombination (EVG+COBI+FTC+TDF) in einem Studienarm durchgeführt wurden [59]. Studien, die nur mit einzelnen, aber nicht allen vier Komponenten in Kombination durchgeführt wurden, werden hier nicht gelistet, unabhängig davon, ob sie mit der Marketing Authorisation eingereicht wurden. Zudem werden in der folgenden Tabelle nur die Studien im entsprechenden Anwendungsgebiet aufgenommen (Patienten, die nicht antiretroviral vorbehandelt sind), aus welchem die Daten auf die vorbehandelten Patienten zu übertragen sind.

Tabelle 4-124: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel für indirekte Vergleiche

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
GS-US-236-0102	ja	laufend 48-Wochen-Interim-Bericht verfügbar 96-Wochen-Interim-Bericht verfügbar	192 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo Efavirenz 600 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo
GS-US-236-0103	ja	laufend 48-Wochen-Interim-Bericht verfügbar 96-Wochen-Interim-Bericht verfügbar	192 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo Ritonavir 100 mg + Atazanavir 300 mg + Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo
GS-US-236-0104	ja	abgeschlossen 96-Wochen-Interim-Bericht verfügbar	60 Wochen, im Anschluss daran Rollover Open-label Extension	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo Efavirenz 600 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo
GS-US-236-0128	nein	laufend	48 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo Ritonavir 100 mg + Atazanavir 300 mg + Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg plus Placebo
GS-US-292-0102	nein	laufend	mindestens 48 Wochen, im Anschluss daran Rollover Open-label Extension	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviralafenamid 10 mg plus Placebo Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
				245 mg plus Placebo
GS-US-292-0111	nein	laufend	mindestens 48 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviralfenamid 10 mg plus Placebo Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofovirdisoproxil 245 mg plus Placebo
GS-US-292-0104	nein	laufend	96 Wochen	Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviralfenamid 10 mg plus Placebo Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofovirdisoproxil 245 mg plus Placebo

Geben Sie an, welchen Stand die Information in Tabelle 4-124 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.

Tabelle 4-124 bildet den Studienstatus vom 30. April 2013 ab.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-124 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-125: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – Studien für indirekte Vergleiche

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
GS-US-292-0102	Es liegen lediglich 24-Wochen-Daten vor. Außerdem wurde die Studie nicht mit einem für einen indirekten Vergleich geeigneten Brückenkomparator als Vergleichstherapie (EFV/FTC/TDF oder ATV/r+FTC/TDF) durchgeführt.
GS-US-292-0111	Es liegen keine Studienergebnisse vor
GS-US-236-0128	Es liegen keine Studienergebnisse vor.
GS-US-292-0104	Es liegen keine Studienergebnisse vor.

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

4.3.2.1.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

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

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

Da keine Studien vorlagen, die einen direkten Vergleich zwischen dem zu bewertenden Arzneimittel und der zweckmäßigen Vergleichstherapie, Raltegravir in Kombination mit Emtricitabin und Tenofoviridisoproxil (RAL+FTC/TDF), ermöglichen, wurde ein indirekter Vergleich von Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil (EVG/COBI/FTC/TDF) mit Raltegravir in Kombination mit Emtricitabin und Tenofoviridisoproxil (RAL+FTC/TDF) durchgeführt.

Wie in Abschnitt 4.3.2.1.1.1 beschrieben, kamen grundsätzlich drei RCT mit dem zu bewertenden Arzneimittel für einen indirekten Vergleich in Frage (S102, S103 und S104). In diesen Studien wurde EVG/COBI/FTC/TDF gegen Efavirenz (EFV) bzw. Ritonavir-geboostertem Atazanavir (ATV/r), jeweils in Kombination mit Emtricitabin und Tenofoviridisoproxil (FTC/TDF) verglichen. Deshalb wurde nach RCT mit der zweckmäßigen Vergleichstherapie, Raltegravir in Kombination mit FTC/TDF, gesucht, die mit EFV oder ATV/r, jeweils in Kombination mit FTC/TDF, kontrolliert waren.

Die umfassende bibliographische Literaturrecherche vom 21.04.2013 nach RCT mit der zweckmäßigen Vergleichstherapie Raltegravir erzielte insgesamt 41 Treffer (ohne Duplikate), die zunächst auf Grundlage von Titel und Abstract nach den in Abschnitt 4.2.2 definierten Kriterien gescreent wurden. Dabei wurden 36 Publikationen als nicht relevant eingestuft und

ausgeschlossen. Für die verbleibenden 6 Publikationen wurden im Anschluss die Volltexte besorgt und hinsichtlich dieser Kriterien überprüft. Im Ergebnis entsprachen 3 Publikationen nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C).

Es wurden Publikationen zu zwei RCT gefunden, in denen Raltegravir mit Efavirenz verglichen wurde:

- STARTMRK: Raltegravir gegen Efavirenz, jeweils in Kombination mit Emtricitabin und Tenofoviridisoproxil; vier Publikationen, von denen eine ausgeschlossen wurde, da keine relevanten Zusatzinformationen berichtet wurden
 - STARTMRK ist eine doppelblinde, randomisierte, aktiv-kontrollierte Phase III-Studie zum Vergleich von RAL 400 mg BID und EFV 600 mg QHS, jeweils in Kombination mit FTC/TDF
- Studie 004: Raltegravir gegen Efavirenz, jeweils in Kombination mit Lamivudin und Tenofoviridisoproxil; ausgeschlossen:
 - Protocol 004 ist eine doppelblinde, randomisierte, Dosisfindungsstudie. Teil I umfasste eine Monotherapie mit RAL über 10 Tage (n=35), Teil II untersuchte die Sicherheit, Verträglichkeit und Wirksamkeit von RAL in Dosierungen von 100, 200, 400, oder 600 mg BID versus EFV 600 mg QD, jeweils in Kombination mit Tenofoviridisoproxil 245 mg und Lamivudin (3TC) 300 mg (n=30 aus Teil I + n=171 randomisierte Patienten) über 48 Wochen - damit entspricht die Vergleichstherapie nicht der als Brückenkomparator festgelegten Kombination aus EFV, FTC und TDF. Eine Vergleichbarkeit ist trotz der strukturellen Ähnlichkeit zwischen Lamivudin (3TC) und Emtricitabin (FTC) klar nicht gegeben: Klinische Daten zeigen, dass es deutliche pharmakokinetische und klinische Unterschiede gibt [13, 14]:
 - Lamivudin ist weniger wirksam als Emtricitabin [105, 106] und führt entsprechend zu einer deutlich höheren Resistenzentwicklung [107, 108]
 - Hinzu kommt, dass zwischen TDF und FTC synergistische Effekte gezeigt werden konnten, die zu jeweils höheren Konzentrationen der eigentlichen Wirksubstanzen (Tri-Phosphate) führen [109]; der Effekt zwischen TDF und 3TC hingegen ist lediglich additiv [110]
 - 3TC hat eine deutlich geringere Halbwertszeit als FTC, die zudem eine hohe interindividuelle Variabilität aufweist [13, 14]
 - Dies bestätigen auch die Leitlinien: Die Kombination von Tenofovir und Lamivudin wird in Leitlinien nicht oder nur als Alternative empfohlen [2].

Es wurden keine Publikationen zu RCT mit Raltegravir im Vergleich zu Ritonavir-geboostertem Atazanavir gefunden.

Somit ergab die bibliographische Literaturrecherche drei relevante Publikationen zu insgesamt einer Studie. Bei den identifizierten Publikationen handelt es sich um Lennox (2009 und 2010) und Rockstroh [111-113]. Für das zu bewertende Arzneimittel wurde keine bibliografische Literaturrecherche durchgeführt.

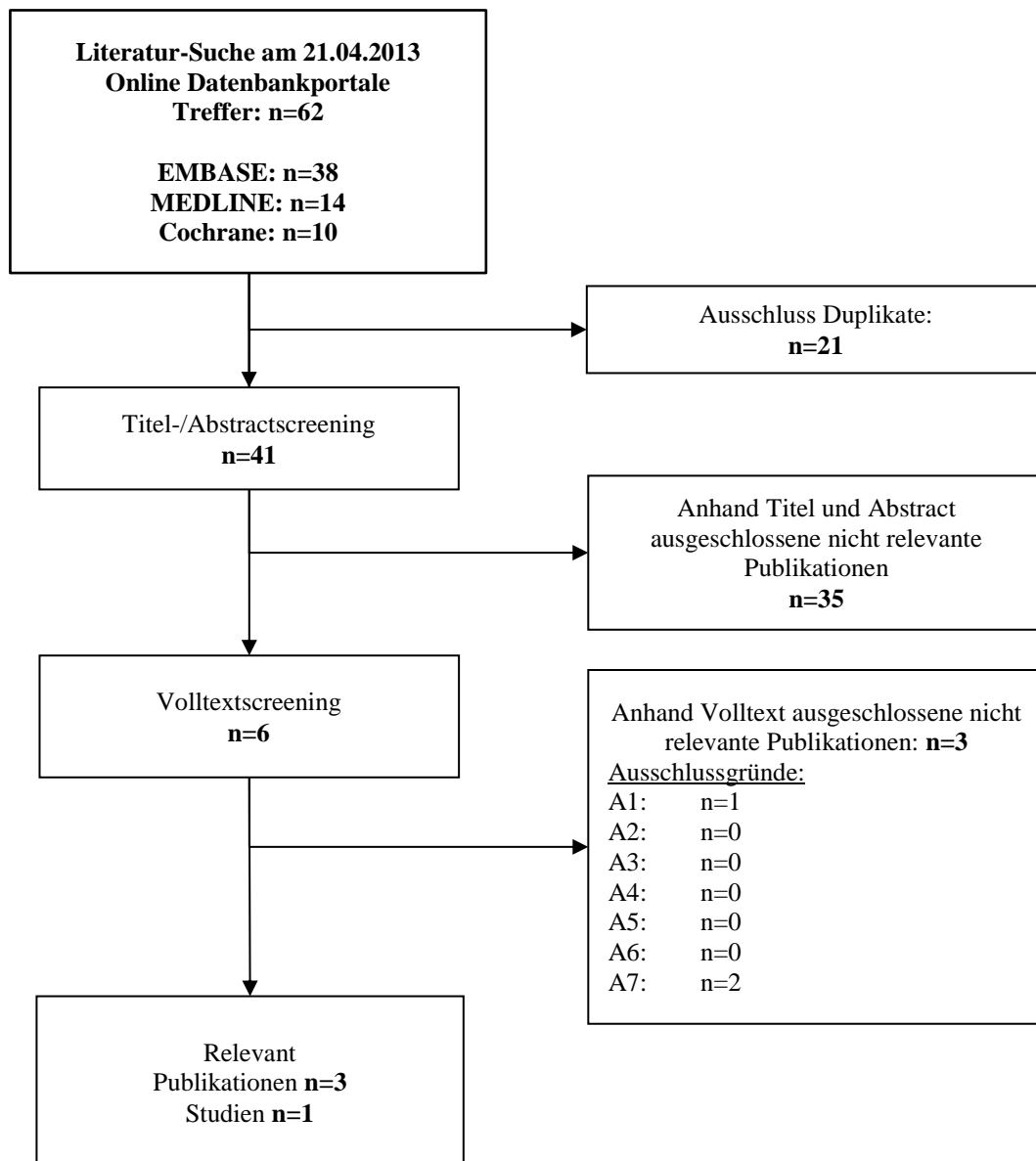


Abbildung 11: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit der zweckmäßigen Vergleichstherapie

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 4-3) 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 relevanten Studien aus der Suche in Studienregistern sind in Tabelle 4-126 (zu bewertendes Arzneimittel) und Tabelle 4-127 (zweckmäßige Vergleichstherapie) aufgeführt.

Tabelle 4-126: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel für indirekte Vergleiche

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)
GS-US-236-0102	clinicaltrials.gov [NCT01095796] [84], ICTRP [NCT01095796] [85]	ja	n.d.
GS-US-236-0103	clinicaltrials.gov [NCT01106586] [86], ICTRP Search Portal [NCT01106586] [87]	ja	n.d.
GS-US-236-0104	clinicaltrials.gov [NCT00869557] [88], ICTRP [NCT00869557] [89]	ja	n.d.
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.			

Tabelle 4-127: Relevante Studien aus der Suche in Studienregistern – RCT mit der zweckmäßigen Vergleichstherapie für indirekte Vergleiche

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)
STARTMRK	clinicaltrials.gov [NCT00369941] [114], ICTRP [NCT00369941, EUCTR2006- 003109-23-DE] [115, 116]	nein	ja
<p>a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.</p>			

Um jede verfügbare Information für den indirekten Vergleich zu nutzen, wurden zusätzlich zu der bibliographischen Literaturrecherche die auf der Studienregisterseite der Studie STARTMRK eingetragenen acht Publikationen betrachtet (Stand: 09.04.2013)

- Lennox JL, DeJesus E, Lazzarin A, Pollard RB, Madruga JV, Berger DS, Zhao J, Xu X, Williams-Diaz A, Rodgers AJ, Barnard RJ, Miller MD, DiNubile MJ, Nguyen BY, Leavitt R, Sklar P; STARTMRK investigators. Safety and efficacy of raltegravir-based versus efavirenz-based combination therapy in treatment-naïve patients with HIV-1 infection: a multicentre, double-blind randomised controlled trial. *Lancet*. 2009 Sep 5;374(9692):796-806. Epub 2009 Aug 3.
- DeJesus E, Rockstroh JK, Lennox JL, Saag MS, Lazzarin A, Zhao J, Wan H, Rodgers AJ, Walker ML, Miller M, DiNubile MJ, Nguyen BY, Teppler H, Leavitt R, Sklar P; STARTMRK Investigators. Efficacy of raltegravir versus efavirenz when combined with tenofovir/emtricitabine in treatment-naïve HIV-1-infected patients: week-192 overall and subgroup analyses from STARTMRK. *HIV Clin Trials*. 2012 Jul-Aug;13(4):228-32.
- Rockstroh JK, Lennox JL, Dejesus E, Saag MS, Lazzarin A, Wan H, Walker ML, Xu X, Zhao J, Teppler H, Dinubile MJ, Rodgers AJ, Nguyen BY, Leavitt R, Sklar P; STARTMRK Investigators. Long-term treatment with raltegravir or efavirenz combined with tenofovir/emtricitabine for treatment-naïve human immunodeficiency virus-1-infected patients: 156-week results from STARTMRK. *Clin Infect Dis*. 2011 Oct;53(8):807-16.
- Rockstroh J, Teppler H, Zhao J, Sklar P, Harvey C, Strohmaier K, Leavitt R, Nguyen BY. Safety and efficacy of raltegravir in patients with HIV-1 and hepatitis B and/or C virus coinfection. *HIV Med*. 2012 Feb;13(2):127-31. doi: 10.1111/j.1468-1293.2011.00933.x. Epub 2011 May 22.

- Rockstroh JK, Teppler H, Zhao J, Sklar P, Miller MD, Harvey CM, Strohmaier KM, Leavitt RY, Nguyen BY. Clinical efficacy of raltegravir against B and non-B subtype HIV-1 in phase III clinical studies. *AIDS*. 2011 Jul 17;25(11):1365-9. doi: 10.1097/QAD.0b013e328348065a.
- Nguyen BY, Isaacs RD, Teppler H, Leavitt RY, Sklar P, Iwamoto M, Wenning LA, Miller MD, Chen J, Kemp R, Xu W, Fromtling RA, Vacca JP, Young SD, Rowley M, Lower MW, Gottesdiener KM, Hazuda DJ. Raltegravir: the first HIV-1 integrase strand transfer inhibitor in the HIV armamentarium. *Ann N Y Acad Sci*. 2011 Mar;1222:83-9. doi: 10.1111/j.1749-6632.2011.05972.x. Review.
- Teppler H, Brown DD, Leavitt RY, Sklar P, Wan H, Xu X, Lievano F, Lehman HP, Mast TC, Nguyen BY. Long-term safety from the raltegravir clinical development program. *Curr HIV Res*. 2011 Jan;9(1):40-53.
- Lennox JL, DeJesus E, Berger DS, Lazzarin A, Pollard RB, Ramalho Madruga JV, Zhao J, Wan H, Gilbert CL, Teppler H, Rodgers AJ, Barnard RJ, Miller MD, Dinubile MJ, Nguyen BY, Leavitt R, Sklar P; STARTMRK Investigators. Raltegravir versus Efavirenz regimens in treatment-naive HIV-1-infected patients: 96-week efficacy, durability, subgroup, safety, and metabolic analyses. *J Acquir Immune Defic Syndr*. 2010 Sep;55(1):39-48.

Drei dieser Publikationen entsprechen den bei der Literaturrecherche identifizierten und eingeschlossenen Publikationen [111-113]. DeJesus et al. 2012 enthält keine zusätzlichen Daten und wurde deswegen nicht weiter berücksichtigt (Ausschluss in der bibliographischen Recherche). Rockstroh et al. 2011 und 2012 enthalten ebenfalls keine zusätzlichen Informationen und wurden bereits in der bibliographischen Recherche ausgeschlossen. Auch Nguyen et al. 2011 enthält keine zusätzlichen Daten. Teppler et al. 2011 [117] hingegen stellt ergänzende Daten zu unerwünschten Ereignissen auf SOC-Ebene in der STARTMRK dar und wurde (abweichend vom primären Ausschluss in der Literaturrecherche) für den Vergleich zusätzlich herangezogen. Eine Liste mit Ausschlussgründen der Publikationen, die nicht für die Nutzenbewertung herangezogen werden, findet sich in Anhang 4-C.

Somit werden insgesamt vier relevante Publikationen [111-113, 117] zur STARTMRK für die Nutzenbewertung herangezogen.

4.3.2.1.1.4 Resultierender Studienpool aus den einzelnen Suchschritten

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.2.1.1.1, 4.3.2.1.1.2 und 4.3.2.1.1.3) resultierenden Pool relevanter Studien für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle für die Nutzenbewertung berücksichtigten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilen-

überschriften 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 4-128: Studienpool – Studien für indirekte Vergleiche (zu bewertendes Arzneimittel und zweckmäßige Vergleichstherapie)

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						
aktivkontrolliert, zu bewertendes Arzneimittel						
GS-US-236-0102	ja	ja	ja	ja [90, 91]	ja [NCT01095796] [84, 85]	ja [92]
GS-US-236-0104	ja	ja	ja	ja [93]	ja [NCT00869557] [88, 89]	ja [94]
aktivkontrolliert, zweckmäßige Vergleichstherapie						
START MRK	nein	nein	ja	nein	ja [NCT00369941] [114], [NCT00369941, EUCTR2006-003109-23-DE] [115, 116]	ja [111-113, 117]
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.						

Die beiden Studien GS-US-236-0102 und GS-US-236-0104 wurden beide mit der fixen Kombination EFV/FTC/TDF als Vergleichstherapie durchgeführt. Untersuchte Patientenpopulation waren jeweils antiretroviral nicht vorbehandelte erwachsene Patienten mit HIV-1 Infektion. Auch wenn EFV/FTC/TDF - EFV/FTC/TDF ist für vorbehandelte, virologisch supprimierte Patienten zugelassen - hier rein formal außerhalb seiner in Deutschland gültigen Zulassung gegeben wurde, sind diese beiden Studien in der vorliegenden Nutzenbewertung dennoch einzuschließen und zur Ableitung eines Nutzens resp. Zusatznutzens geeignet. Zum Einen hält das IQWiG den Einschluss von Studien, in welchen Patienten nicht gemäß Zulassung behandelt wurden, für gerechtfertigt, sofern die Behandlungseffekte in einer nicht zulassungsgemäßen Behandlung mit einer Behandlung innerhalb der Zulassung hinreichend vergleichbar sind [60]. Des Weiteren sieht der G-BA die

beiden Studien (GS-US-236-0102 und GS-US-236-0104) als geeignet, um valide Schlüsse zum Nutzen und Zusatznutzen von EVG/COBI/FTC/TDF zu ziehen [59]. Aus diesem Grund wird davon ausgegangen, dass die beiden Studien GS-US-236-0102 und GS-US-236-0104 geeignet sind, um EVG/COBI/FTC/TDF im Vergleich zur zweckmäßigen Vergleichstherapie RAL+FTC/TDF in der Indikation der HIV-1 Infektion bei antiretroviral nicht vorbehandelten erwachsenen Patienten, zu bewerten.

Im Folgenden werden die Studien GS-US-236-0102 und GS-US-236-0104 aus Übersichtlichkeitsgründen mit S102 und S104 abgekürzt.

Da keine relevante Studie mit der zweckmäßigen Vergleichstherapie Raltegravir in Kombination mit Emtricitabin und Tenofoviridisoproxil (RAL+FTC/TDF) versus Ritonavir-geboostertes Atazanavir in Kombination mit Emtricitabin und Tenofoviridisoproxil (ATV/r+FTC/TDF) identifiziert wurde, wurde Studie S103 nicht in den Studienpool aufgenommen. Studie GS-US-236-0103 und ihre Ergebnisse werden in dem in Modul 5 hinterlegten Modul 4 zum Vergleich von EVG/COBI/FTC/TDF versus ATV/r+FTC/TDF dargestellt.

Die Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF im Vergleich zu RAL+FTC/TDF basiert im Folgenden auf zwei aktiv-kontrollierten randomisierten Studien des pharmazeutischen Unternehmers, die das zu bewertende Arzneimittel EVG/COBI/FTC/TDF mit EFV/FTC/TDF direkt vergleichen (GS-US-236-0104 und GS-US-236-0102), und außerdem einer randomisierten aktiv-kontrollierten Studie die RAL+FTC/TDF mit EFV+FTC/TDF direkt vergleicht. Es wird also ein adjustierter indirekter Vergleich über EFV+FTC/TDF als Brückenkompator durchgeführt. Es werden Daten zu Woche 48 und 96 präsentiert.

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.

Im vorliegenden indirekten Vergleich von Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil (EVG/COBI/FTC/TDF) versus Raltegravir in Kombination mit Emtricitabin und Tenofoviridisoproxil (RAL+FTC/TDF) werden nur Ergebnisse der Endpunkte dargestellt, für die theoretisch ein indirekter Vergleich durchgeführt werden kann. Vollständige Daten aus den Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (GS-US-236-0102 und GS-US-236-0104), sind Modul 4A zu entnehmen.

Tabelle 4-129: Charakterisierung der eingeschlossenen Studien – Studien für den indirekten Vergleich

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
S102	RCT, doppelblind, parallel	Antiretroviral nicht- vorbehandelte HIV- 1-infizierte Erwachsene	EVG/COBI/FTC/TDF: n= 353 EFV/FTC/TDF: n= 354	Screening: 5 Wochen Behandlung: 96 Wochen + Zeit bis Entblindung dann entweder: Open-label- Behandlung: bis Produkt kommerziell verfügbar oder Gilead das Forschungsprogram m dazu stoppt oder Follow-Up: 30 Tage	USA und Puerto Rico Woche 48: 03/2010 – 08/2011 Woche 96: 03/2010 – 07/2012	<u>Primärer Endpunkt:</u> Virologisches Ansprechen zu Woche 48 <u>Sekundäre Endpunkte:</u> Zeit bis Verlust des virologischen Ansprechens Veränderung der CD4-Zellzahl Mortalität Unerwünschte Ereignisse
S104	RCT, doppelblind, parallel	Antiretroviral nicht- vorbehandelte HIV- 1-infizierte Erwachsene	EVG/COBI/FTC/TDF: n= 48 EFV/FTC/TDF: n= 23	Screening: 4 Wochen Behandlung: 48 Wochen + Zeit bis Entblindung (Woche 60) dann entweder: Open-label- Behandlung: bis Produkt kommerziell verfügbar oder Gilead das Forschungsprogram m dazu stoppt	USA 03/2009 – 05/2011	<u>Primärer Endpunkt:</u> Virologisches Ansprechen zu Woche 24 <u>Sekundäre Endpunkte:</u> Virologisches Ansprechen zu Woche 48 Veränderung der CD4-Zellzahl Mortalität

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

				oder Follow-Up: 30 Tage	Unerwünschte Ereignisse	
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF						
START MRK	RCT, doppelblind, parallel	Antiretroviral nicht- vorbehandelte HIV- 1-infizierte Erwachsene	RAL+FTC/TDF: n= 282 EFV+FTC/TDF: n=284	Screening: k.A. Behandlung: 96 Wochen Follow-up: k.A.	International: Australien, Brasilien, Chile, Deutschland, Frankreich, Indien, Italien, Kanada, Kolumbien, Mexiko, Peru, Spanien, Thailand und USA 09/2006 - 06/2008	<u>Primärer Endpunkt:</u> Virologisches Ansprechen zu Woche 48 <u>Patientenrelevante sekundäre Endpunkte:</u> Veränderung der CD4-Zellzahl Mortalität Unerwünschte Ereignisse

Tabelle 4-130: Charakterisierung der Interventionen – Studien für indirekte Vergleiche

Studie	Intervention	Vergleichstherapie	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
S102	150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin, 245 mg Tenofovir- disoproxil einmal täglich (EVG/COBI/FTC/TDF)	600 mg Efavirenz, 200 mg Emtricitabin, 245 mg Tenofovir- disoproxil einmal täglich (EFV/FTC/TDF)	Screening: 5 Wochen Behandlung: 96 Wochen + Zeit bis zur Entblindung (Patienten kamen alle 12 Wochen zur Visite) Anschließend: Rollover Open-label Studie bis zur kommerziellen Verfügbarkeit des Medikaments oder bis Gilead das Forschungsprogramm dazu stoppt oder Follow-Up: 30 Tage
S104	150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin, 245 mg Tenofovir- disoproxil einmal täglich (EVG/COBI/FTC/TDF)	600 mg Efavirenz, 200 mg Emtricitabin, 245 mg Tenofovir- disoproxil einmal täglich (EFV/FTC/TDF)	Screening: 4 Wochen Behandlung: 48 Wochen + Zeit bis zur Entblindung (Woche 60) Anschließend: Rollover Open-label Studie bis zur kommerziellen Verfügbarkeit des Medikaments oder bis Gilead das Forschungsprogramm dazu stoppt oder Follow-Up: 30 Tage
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF			
STARTMRK	400 mg Raltegravir zweimal täglich + 200 mg Emtricitabin, 245 mg Tenofovir- disoproxil einmal täglich	600 mg Efavirenz + 200 mg Emtricitabin, 245 mg Tenofovir- disoproxil einmal täglich	Behandlungsdauer: 96 Wochen

Tabelle 4-131: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.)

Studie Gruppe	N	Alter (Jahre)	Geschlecht	Baseline BMI (kg/m ²)
		Mittelwert (SD)	w/m (%)	Mittelwert (SD)
S102				
EVG/COBI/FTC/TDF	348	38 (10,4)	11,8/88,2	26,7 (5,9)
EFV/FTC/TDF	352	38 (10,6)	10,2/89,8	26,1 (5,2)
S104				
EVG/COBI/FTC/TDF	48	36 (8,9)	8,3/91,7	25,9 (3,6)
EFV/FTC/TDF	23	35 (9,6)	8,7/91,3	27,8 (4,2)
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF				
STARTMRK				
RAL+FTC/TDF	281	37,6 (9,0)	19,0/81,0	24 (5)
EFV+FTC/TDF	282	36,9 (10,0)	18,0/82,0	24 (5)
BMI = Body Mass Index, SD = Standardabweichung				

Tabelle 4-132: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.)

Studie Gruppe	Rasse n (%)					
	Kaukasisch	Asiatisch	Schwarz (oder afrikanischer Abstammung) ^a	Indianer/ Ureinwohner Alaskas	Hawaiianer/ Pazifik-insulaner	Andere
S102						
EVG/COBI/FTC/TDF	214 (61,5)	6 (1,7)	106 (30,5)	2 (0,6)	4 (1,1)	16 (4,6)
EFV/FTC/TDF	227 (64,5)	10 (2,8)	91 (25,9)	4 (1,1)	1 (0,3)	19 (5,4)
S104						
EVG/COBI/FTC/TDF	33 (68,8)	1 (2,1)	12 (25,0)	1 (2,1)	k.A.	1 (2,1)
EFV/FTC/TDF	18 (78,3)	0 (0,0)	5 (21,7)	0 (0,0)	k.A.	0 (0,0)
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF						
STARTMRK						
RAL+FTC/TDF	116 (41,0)	36 (13,0)	33 (12,0)	1 (<1,0) ^b	k.A.	^c
EFV+FTC/TDF	123 (44,0)	32 (11,0)	23 (8,0)	1 (<1,0) ^b	k.A.	^c
^a Angabe in Studie S102 ^b Diese Patienten waren in STARTMRK-Studie Indianer oder Ureinwohner Amerikas. ^c Außerdem waren in STARTMRK-Studie im RAL+FTC/TDF-Arm 35 (12,0%) Patienten gemischtrassig und im EFV+FTC/TDF-Arm waren dies 36 (13,0%). In dieser Studie wurde hispanische Abstammung, im Gegensatz zu Studien S102 und S104 als Rasse betrachtet. Deswegen wird dieser Anteil Patienten hier nicht mir dargestellt, weswegen sich die Patientenanteile nicht zu 100% aufsummieren. k.A. = keine Angaben, n: Anzahl Patienten in der Kategorie						

Tabelle 4-133: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.)

Studie Gruppe	Ethnische Abstammung n (%)	
	Hispanisch/Latino	Nicht hispanisch/Latino
S102		
EVG/COBI/FTC/TDF	82 (23,6)	266 (76,4)
EFV/FTC/TDF	85 (24,1)	267 (75,9)
S104		
EVG/COBI/FTC/TDF	4 (8,3)	44 (91,7)
EFV/FTC/TDF	3 (13,0)	20 (87,0)
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF		
STARTMRK		
RAL+FTC/TDF	a	a
EFV+FTC/TDF		
^a In dieser Studie wurde hispanische Abstammung als Rasse betrachtet. 60 (21,0%) Patienten im RAL+FTC/TDF-Arm und 67 (24,0%) der EFV+FTC/TDF-Patienten waren hispanischer Abstammung. n: Anzahl der Patienten in der Kategorie		

Tabelle 4-134: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.)

Studie Gruppe	Baseline HIV-1-RNA (log ₁₀ Kopien/mL) Mittelwert (SD)	Baseline HIV-1-RNA Kopien/mL n (%)	
		≤100.000	>100.000
S102			
EVG/COBI/FTC/TDF	4,73 (0,6)	230 (66,1)	118 (33,9)
EFV/FTC/TDF	4,78 (0,6)	236 (67,0)	116 (33,0)
S104			
EVG/COBI/FTC/TDF	4,59 (0,6)	37 (77,1)	11 (22,9)
EFV/FTC/TDF	4,58 (0,6)	18 (78,3)	5 (21,7)
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF			
STARTMRK			
RAL+FTC/TDF	5,0 (0,6)	127 (45,0)	154 (55,0)
EFV+FTC/TDF	5,0 (0,6)	139 (49,0)	143 (51,0)
n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung			

Tabelle 4-135: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche (S102 und S104: Safety Analyse-Set, STARTMRK: k.A.)

Studie Gruppe	CD4-Zellzahl zu Baseline (/mm ³) Mittelwert (SD)	CD4-Zellzahl zu Baseline (/mm ³) n (%)				
		≤50	51-≤200	201-≤350	351-≤500	>500
S102						
EVG/COBI/FTC/TDF	391 (188,6)	7 (2,0)	36 (10,3)	112 (32,2)	113 (32,5)	80 (23,0)
EFV/FTC/TDF	382 (170,2)	6 (1,7)	45 (12,8)	96 (27,3)	136 (38,6)	69 (19,6)
S104						
EVG/COBI/FTC/TDF	378 (179,8)	0 (0,0)	7 (14,6)	17 (35,4)	14 (29,2)	10 (20,8)
EFV/FTC/TDF	454 (137,6)	0 (0,0)	0 (0,0)	8 (34,8)	5 (21,7)	10 (43,5)
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF						
STARTMRK						
RAL+FTC/TDF	218,9 (124,2)	27 (10,0)	104 (37,0)		150 (53,0)	
EFV+FTC/TDF	217,4 (133,6)	31 (11,0)	105 (37,0)		145 (51,0)	

n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung

Tabelle 4-136: Charakterisierung der Studienpopulationen (Patientendisposition) – Studien für indirekte Vergleiche

Studie Disposition	Intervention	Vergleichstherapie
	n (%)	n (%)
S102	EVG/COBI/FTC/TDF	EFV/FTC/TDF
Randomisiert	353	354
Behandlung erhalten	348 (100,0)	352 (100,0)
Woche 48 bis zum Daten Cut-Off in Behandlung abgeschlossen	311 (89,4)	306 (86,9)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	12 (3,4)	18 (5,1)
Tod	1 (0,3)	1 (0,3)
Schwangerschaft	1 (0,3)	0 (0,0)
Fehlende Wirksamkeit	5 (1,4)	4 (1,1)
Entscheidung des Prüfarztes	1 (0,3)	0 (0,0)
Einverständnis zurückgezogen	3 (0,9)	5 (1,4)
Lost-to-follow-up	10 (2,9)	12 (3,4)
Non-Compliance	3 (0,9)	6 (1,7)
Protokollverletzung	1 (0,3)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Woche 48 bis zum Daten Cut-Off in Studie abgeschlossen	319 (91,7)	316 (89,8)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	6 (1,7)	12 (3,4)
Tod	1 (0,3)	1 (0,3)
Schwangerschaft	0 (0,0)	0 (0,0)
Fehlende Wirksamkeit	3 (0,9)	0 (0,0)

Studie	Intervention	Vergleichstherapie
Disposition	n (%)	n (%)
Entscheidung des Prüfarztes	1 (0,3)	0 (0,0)
Einverständnis zurückgezogen	5 (1,4)	5 (1,4)
Lost-to-follow-up	10 (2,9)	12 (3,4)
Non-Compliance	2 (0,6)	6 (1,7)
Protokollverletzung	1 (0,3)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Woche 96 bis zum Daten Cut-Off in Behandlung abgeschlossen	295 (84,8)	291 (82,7)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	16 (4,6)	24 (6,8)
Tod	1 (0,3)	1 (0,3)
Schwangerschaft	2 (0,6)	0 (0,0)
Fehlende Wirksamkeit	6 (1,7)	5 (1,4)
Entscheidung des Prüfarztes	2 (0,6)	0 (0,0)
Einverständnis zurückgezogen	3 (0,9)	7 (2,0)
Lost-to-follow-up	17 (4,9)	17 (4,8)
Non-Compliance	5 (1,4)	7 (2,0)
Protokollverletzung	1 (0,3)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Woche 96 bis zum Daten Cut-Off in Studie abgeschlossen	304 (87,4)	299 (84,9)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	8 (2,3)	17 (4,8)
Tod	1 (0,3)	1 (0,3)
Schwangerschaft	1 (0,3)	0 (0,0)
Fehlende Wirksamkeit	4 (1,1)	1 (0,3)
Entscheidung des Prüfarztes	2 (0,6)	0 (0,0)
Einverständnis zurückgezogen	7 (2,0)	7 (2,0)
Lost-to-follow-up	17 (4,9)	20 (5,7)
Non-Compliance	3 (0,9)	7 (2,0)
Protokollverletzung	1 (0,3)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
S104	EVG/COBI/FTC/TDF	EFV/FTC/TDF
Randomisiert	48	23
Behandlung erhalten	48 (100,0)	23 (100,0)
Randomisierte Phase ^a bis zum Daten Cut-Off in Behandlung beendet	45 (93,8)	20 (87,0)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	0 (0,0)	1 (4,3)
Tod	0 (0,0)	0 (0,0)
Schwangerschaft	0 (0,0)	0 (0,0)
Fehlende Wirksamkeit	0 (0,0)	0 (0,0)
Entscheidung des Prüfarztes	1 (2,1)	0 (0,0)
Einverständnis zurückgezogen	0 (0,0)	1 (4,3)
Lost-to-follow-up	2 (4,2)	1 (4,3)

Studie	Intervention	Vergleichstherapie
Disposition	n (%)	n (%)
Non-Compliance	0 (0,0)	0 (0,0)
Protokollverletzung	0 (0,0)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Randomisierte Phase ^a bis zum Daten Cut-Off in Studie beendet	45 (93,8)	21 (91,3)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	0 (0,0)	0 (0,0)
Tod	0 (0,0)	0 (0,0)
Schwangerschaft	0 (0,0)	0 (0,0)
Fehlende Wirksamkeit	0 (0,0)	0 (0,0)
Entscheidung des Prüfarztes	1 (2,1)	0 (0,0)
Einverständnis zurückgezogen	0 (0,0)	1 (4,3)
Lost-to-follow-up	2 (4,2)	1 (4,3)
Non-Compliance	0 (0,0)	0 (0,0)
Protokollverletzung	0 (0,0)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF		
STARTMRK	RAL+FTC/TDF	EFV+FTC/TDF
Randomisiert	282	284
Behandlung erhalten	281 (100,0)	282 (100,0)
Woche 48 bis zum Daten Cut-Off in Studie beendet	257 (91,5)	247 (87,6)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	8 (2,8)	17 (6,0)
Tod	0 (0,0)	0 (0,0)
Schwangerschaft	0 (0,0)	0 (0,0)
Fehlende Wirksamkeit	4 (1,4)	2 (0,7)
Entscheidung des Prüfarztes	0 (0,0)	0 (0,0)
Einverständnis zurückgezogen	0 (0,0)	0 (0,0)
Lost-to-follow-up	3 (1,1)	7 (2,5)
Non-Compliance	0 (0,0)	0 (0,0)
Protokollverletzung	0 (0,0)	0 (0,0)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Andere Gründe	9 (3,2)	9 (3,2)
Woche 96 bis zum Daten Cut-Off in Studie beendet	245 (87,2)	232 (82,3)
Gründe für vorzeitigen Abbruch		
Unerwünschtes Ereignis	10 (3,6)	19 (6,7)
Tod	0 (0,0)	0 (0,0)
Schwangerschaft	0 (0,0)	0 (0,0)
Fehlende Wirksamkeit	5 (1,8)	5 (1,8)
Entscheidung des Prüfarztes	0 (0,0)	0 (0,0)
Einverständnis zurückgezogen	0 (0,0)	0 (0,0)
Lost-to-follow-up	4 (1,4)	12 (4,3)
Non-Compliance	0 (0,0)	0 (0,0)
Protokollverletzung	0 (0,0)	0 (0,0)

Studie	Intervention	Vergleichstherapie
Disposition	n (%)	n (%)
Studie durch Sponsor abgebrochen	0 (0,0)	0 (0,0)
Andere Gründe	17 (6,0)	14 (5,0)

^a Randomisierte Phase dauerte bis Woche 60
n: Anzahl der Patienten in der Kategorie

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

Design

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Die in die Nutzenbewertung von EVG/COBI/FTC/TDF gegenüber der Vergleichstherapie EFV/FTC/TDF eingeschlossenen Studien, S102 und S104, wurden beide in einem randomisierten aktiv-kontrollierten Parallelgruppendesign durchgeführt. Die Patienten wurden jeweils entweder mit EVG/COBI/FTC/TDF (150 mg Elvitegravir/150 mg Cobicistat/200 mg Emtricitabin/245 mg Tenofoviridisoproxil) oder mit EFV/FTC/TDF (600 mg Efavirenz/200 mg Emtricitabin/245 mg Tenofoviridisoproxil) behandelt.

Nach einer 5-wöchigen Screening-Phase in Studie S102 wurden die Patienten 96 Wochen lang in einem doppelblinden Design behandelt. Anschließend wurden die Patienten weiter bis zur Entblindung verblindet behandelt, wobei sie alle 12 Wochen zur Visite erscheinen mussten. Danach hatten die Patienten die Möglichkeit eines Rollovers in eine Open-Label-Studie, in der sie mit EVG/COBI/FTC/TDF behandelt wurden, bis dieses Arzneimittel kommerziell verfügbar war oder Gilead das Forschungsprogramm dazu stoppt. Nutzten die Patienten diese Option nicht, traten sie in die 30-tägige Follow-up Phase ein.

In Studie S104 dauerte die Screening-Phase 4 Wochen, wonach die Patienten in eine 48-wöchige randomisierte Behandlungsphase eintraten. Anschließend wurden sie bis zur Entblindung zu Woche 60 weiter randomisiert behandelt. Danach hatten die Patienten die Möglichkeit eines Rollovers in eine Open-Label-Studie, in der sie mit EVG/COBI/FTC/TDF behandelt wurden, bis das Arzneimittel kommerziell verfügbar war oder Gilead das Forschungsprogramm dazu stoppt. Nutzten die Patienten diese Option nicht, traten sie gleich in die 30-tägige Follow-up Phase ein.

Studie S102 wurde in den USA und Puerto Rico durchgeführt und Studie S104 wurde in den USA durchgeführt.

Die Studien wurden jeweils in einem doppel-blinden double-dummy Design durchgeführt, um die Verblindung zu gewährleisten, d.h. weder den Prüfarzten noch den Patienten war die Gruppenzugehörigkeit der Patienten bekannt. Dazu wies in beiden Studien das interactive voice response system (IVRS)/ interactive web response system (IWRS) den Studienmedikations-Flaschen verblindete Flaschennummern zu.

In Studie S102 wurden die Patienten durch das IVRS/IWRS in einem Verhältnis von 1:1 auf die beiden Behandlungsarme randomisiert (Randomisierung durchgeführt durch Bracket). Dabei wurde nach Screening HIV-1-RNA Level (≤ 100.000 Kopien/mL oder > 100.000 Kopien/mL) stratifiziert und es wurde eine Blocklänge von 4 verwendet. In Studie S104 wurden die Patienten durch das IVRS/IWRS in einem Verhältnis von 2:1 auf die beiden Behandlungsarme randomisiert. Dabei wurde nach Screening HIV-1-RNA Level (≤ 100.000 Kopien/mL oder > 100.000 Kopien/mL) stratifiziert und es wurde eine Blocklänge von 6 verwendet.

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

Die STARTMRK-Studie wurde in einem randomisierten, aktivkontrollierten Parallelgruppendesign durchgeführt. Die Patienten wurden entweder mit 400 mg Raltegravir (zweimal täglich) oder mit 600 mg Efavirenz (einmal täglich), jeweils in Kombination mit 245 mg Tenofoviridisoproxil und 200 mg Emtricitabin einmal täglich, behandelt.

Die Patienten wurden für insgesamt 96 Wochen doppelblind behandelt.

Die Studie wurde international durchgeführt und die Zentren erstreckten sich über fünf Kontinente (Europa, Asien, Nordamerika, Südamerika und Ozeanien).

Die Studie wurde in einem double-dummy Design durchgeführt, um die Verblindung zu gewährleisten, d.h. weder den Prüfern noch den Patienten war die Gruppenzugehörigkeit der Patienten bekannt. Die Patienten wurden durch ein IVRS mittels eines Computer-generierten randomisierten Zuteilungsplan im Verhältnis 1:1 auf die Behandlungen randomisiert. Dabei wurde nach Ausgangsviruslast (> 50.000 vs. ≤ 50.000 Kopien/mL) und Hepatitis-Koinfektionsstatus stratifiziert.

Studienpopulation

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

In beiden Studien S102 und S104 wurden jeweils erwachsene Patienten eingeschlossen, die eine Infektion mit HIV-1 aufwiesen und antiretroviral nicht-vorbehandelt waren. Die Patienten mussten außerdem unter anderem eine Viruslast ≥ 5.000 Kopien/mL aufweisen, eine Lebenserwartung von ≥ 1 Jahr haben und außerdem eine adäquate Nierenfunktion, definiert durch eGFR ≥ 70 mL/min (Kreatinin-Clearance) in Studie S102 bzw. ≥ 80 mL/min in Studie S104 gemäß Cockcroft-Gault, aufweisen. Des Weiteren durften in beiden Studien die Patienten keinerlei antiretrovirale Vorbehandlung bekommen haben; die Sensitivität gegenüber den Wirkstoffen FTC, TDF und EFV musste zusätzlich anhand eines genotypischen Resistenztests dokumentiert worden sein. Da INI-assoziierte Primärresistenzen nicht beschrieben sind (siehe Module 3A und 3B), wurden durch die im Studiendesign vorgegebene fehlende Vorbehandlung per se nur Patienten eingeschlossen, die keine Resistenz-assoziierten Mutationen gegen EVG aufweisen; Resistenz-assoziierte Mutationen gegen FTC, TDF und EFV wurden zusätzlich direkt ausgeschlossen (genotypischer Resistenztest). Somit entsprechen die Studienpopulationen der Population von HIV-1 infizierten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit

Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind. Wichtige Ausschlusskriterien waren unter anderen Schwangerschaft, schwere Infektionen oder die Einnahme von Medikamenten, die nicht zusammen mit den in der Studie verabreichten Behandlungen eingenommen werden sollte.

Die Patienten in Studie S102 waren in beiden Behandlungsarmen im Mittel 38 Jahre alt. In Studie S104 waren die Patienten im EVG/COBI/FTC/TDF-Arm im Mittel 36 Jahre alt und im EFV/FTC/TDF-Arm lag das durchschnittliche Alter bei 35 Jahren. Es wurden in beiden Studien deutlich mehr Männer eingeschlossen (EVG/COBI/FTC/TDF vs. EFV/FTC/TDF S102: 88,2% vs. 89,8%, S104: 91,7% vs. 91,3%), was die höhere Prävalenz der HIV-1-Infektion unter Männern in Europa und den USA widerspiegelt. Der Baseline-BMI lag in der EVG/COBI/FTC/TDF-Gruppe der Studie S102 im Mittel bei 26,7 kg/m² und in der EFV/FTC/TDF-Gruppe bei durchschnittlich 26,1 kg/m². In Studie S104 lag der durchschnittliche Baseline-BMI in der EVG/COBI/FTC/TDF-Gruppe bei 25,9 kg/m² und in der EFV/FTC/TDF-Gruppe bei 27,8 kg/m² (siehe Tabelle 4-131). Hier waren sowohl die Behandlungsarme in den beiden Studien als auch die beiden Studien untereinander balanziert.

Der Großteil der Patienten in den beiden Studien war kaukasischer Herkunft (S102: EVG/COBI/FTC/TDF: 61,5%, EFV/FTC/TDF: 64,5%; S104: EVG/COBI/FTC/TDF: 68,8%, EFV/FTC/TDF: 78,3%), wobei dieser Anteil im EFV/FTC/TDF-Arm jeweils etwas höher war als im EVG/COBI/FTC/TDF-Arm. Mit 30,5% der Patienten im EVG/COBI/FTC/TDF-Arm und 25,9% der Patienten im EFV/FTC/TDF-Arm bildeten schwarze Patienten bzw. Patienten afrikanischer Abstammung in Studie S102 die zweitgrößte Gruppe. Ein ähnliches Bild ergibt sich für Studie S104, dort waren 25,0% der Patienten im EVG/COBI/FTC/TDF-Arm und 21,7% der Patienten im EFV/FTC/TDF-Arm schwarz oder afrikanischer Abstammung. Die übrigen Patienten waren asiatischer (EVG/COBI/FTC/TDF vs. EFV/FTC/TDF: S102: 1,7% vs. 2,8%; S104: 2,1% vs. 0,0%) oder hawaiianischer (EVG/COBI/FTC/TDF vs. EFV/FTC/TDF: S102: 1,1% vs. 0,3%; S104: keine Angaben) Herkunft, und ein kleiner Teil der Patienten waren Indianer oder Ureinwohner Alaskas (EVG/COBI/FTC/TDF vs. EFV/FTC/TDF: S102: 0,6% vs. 1,1%; S104: 2,1% vs. 0,0%). Außerdem gehörten im EVG/COBI/FTC/TDF-Arm in Studie S102 4,6% und im EFV/FTC/TDF-Arm 5,4% der Patienten und in Studie S104 2,1% der Patienten im EVG/COBI/FTC/TDF-Arm und 0,0% der Patienten im EFV/FTC/TDF-Arm einer anderen Rasse an (siehe Tabelle 4-132). Der Großteil der Patienten war nicht hispanischer/ Latino-Abstammung (EVG/COBI/FTC/TDF vs. EFV/FTC/TDF: S102: 76,4% vs. 75,9%, S104: 91,7% vs. 87,0%; siehe Tabelle 4-133). Hier konnten somit keine relevanten Unterschiede zwischen den Behandlungsarmen der beiden Studien oder zwischen den beiden Studien gefunden werden.

Die Ausgangsviruslast zu Baseline in Studie S102 lag bei durchschnittlich 4,73 log₁₀ HIV-1-Kopien/mL im EVG/COBI/FTC/TDF-Arm und 4,78 log₁₀ Kopien/mL im EFV/FTC/TDF-Arm. In Studie S104 lag die Ausgangsviruslast im EVG/COBI/FTC/TDF-Arm bei 4,59 log₁₀ HIV-1-Kopien/mL und 4,58 log₁₀ HIV-1-Kopien/mL im EFV/FTC/TDF-Arm (siehe Tabelle 4-134). Auch hinsichtlich der kategorialen Viruslast waren die Patienten zwischen den beiden Behandlungsarmen in beiden Studien balanziert

(≤ 100.000 / >100.000 Kopien/mL EVG/COBI/FTC/TDF: S102: 66,1%/33,9%; S104: 77,1%/22,9%; im EFV/FTC/TDF-Arm: S102: 77,1%/22,9%; S104: 78,3%/21,7%). Die mittlere CD4-Zellzahl lag zu Baseline von Studie S102 bei 391 Zellen/ μ L im EVG/COBI/FTC/TDF-Arm und 382 Zellen/ μ L im EFV/FTC/TDF-Arm. In Studie S104 lag die mittlere CD4-Zellzahl zu Baseline im EVG/COBI/FTC/TDF-Arm bei 378 Zellen/ μ L und im EFV/FTC/TDF-Arm bei 454 Zellen/ μ L. Im EFV/FTC/TDF-Arm der Studie S104 wiesen die Patienten zu Baseline also eine höhere CD4-Zellzahl auf als im EVG/COBI/FTC/TDF-Arm. Bei dem Großteil der Patienten in beiden Studien lag die CD4-Zellzahl zu Baseline zwischen 201 Zellen/ μ L und 500 Zellen/ μ L. In Studie S102 waren dies im EVG/COBI/FTC/TDF-Arm in der Gruppe 201- \leq 350 Zellen/ μ L 32,2% und in der Gruppe 351- \leq 500 Zellen/ μ L 32,5% der Patienten. Im EFV/FTC/TDF-Arm lag dieser Anteil bei 27,3% in der 201- \leq 350 Zellen/ μ L-Gruppe und 38,6% in der 351- \leq 500 Zellen/ μ L-Gruppe (siehe Tabelle 4-179). In Studie S104 wiesen 35,4% der Patienten im EVG/COBI/FTC/TDF-Arm und 34,8% der Patienten im EFV/FTC/TDF-Arm eine CD4-Zellzahl zu Baseline zwischen 201 bis \leq 350 auf. In der Gruppe 351- \leq 500 lagen die Anteile bei 29,2% im EVG/COBI/FTC/TDF-Arm und 21,7% im EFV/FTC/TDF-Arm. Insgesamt sind die Behandlungsarme recht ausgeglichen. Kleinere Unbalanziertheiten, vor allem hinsichtlich der CD4-Zellzahl, indizieren hier, dass im EVG/COBI/FTC/TDF-Arm eher die etwas schwerer erkrankten HIV-1-infizierten Patienten waren.

In Studie S102 erhielten 348 Patienten EVG/COBI/FTC/TDF und 352 Patienten EFV/FTC/TDF. Davon beendeten 89,4% der Patienten im EVG/COBI/FTC/TDF-Arm und 86,9% der Patienten im EFV/FTC/TDF-Arm die ersten 48 Wochen der Behandlung in Behandlung (siehe Tabelle 4-136). Hauptgründe für den vorzeitigen Abbruch der Behandlung waren unerwünschte Ereignisse (EVG/COBI/FTC/TDF: 3,4%; EFV/FTC/TDF: 5,1%) und lost-to-follow-up (EVG/COBI/FTC/TDF: 2,9%; EFV/FTC/TDF: 3,4%). 96 Wochen in Behandlung beendeten 84,8% der EVG/COBI/FTC/TDF-Patienten und 82,7% der EFV/FTC/TDF-Patienten. Hauptgrund für vorzeitiges Beenden der Behandlung war lost-to-follow-up (EVG/COBI/FTC/TDF: 4,9%; EFV/FTC/TDF: 4,8%). In Studie S104 erhielten 48 Patienten EVG/COBI/FTC/TDF und 23 Patienten EFV/FTC/TDF. 93,8% der EVG/COBI/FTC/TDF-Patienten beendeten 60 Wochen der randomisierten Behandlung in Behandlung und in der EFV/FTC/TDF-Gruppe waren dies 87,0%. Häufigster Grund für einen vorzeitigen Abbruch waren lost-to-follow-up (EVG/COBI/FTC/TDF: 4,2%; EFV/FTC/TDF: 4,3%).

Insgesamt sind die Behandlungsarme untereinander balanziert.

Haupt-Analyse-Set für Wirksamkeitsendpunkte war in beiden Studien das ITT-Set. In Studie S102 umfasste das ITT-Set 348 Patienten im EVG/COBI/FTC/TDF-Arm und 352 Patienten im EFV/FTC/TDF-Arm. In Studie S104 umfasste das ITT-Set 48 Patienten im EVG/COBI/FTC/TDF-Arm und 23 Patienten im EFV/FTC/TDF-Arm. Das Safety Analyse-Set (Haupt-Analyse-Set für Sicherheitsendpunkte) schloss in beiden Studien jeweils die gleiche Anzahl an Patienten wie im ITT-Set ein. In das Per Protokoll-Set wurden in Studie S102 für die Woche 48-Analyse 312 Patienten in den EVG/COBI/FTC/TDF-Arm und

300 Patienten in den EFV/FTC/TDF-Arm eingeschlossen und für die Woche 96-Analyse waren dies 293 Patienten bzw. 294 Patienten. In Studie S104 schloss das Per Protokoll-Set 45 Patienten im EVG/COBI/FTC/TDF-Arm und 20 Patienten im EFV/FTC/TDF-Arm ein.

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

In die Studien wurden erwachsene Patienten eingeschlossen, die eine Infektion mit HIV-1 aufweisen und antiretroviral nicht-vorbehandelt waren. Die Patienten mussten außerdem unter anderem eine Viruslast ≥ 5000 Kopien/mL aufweisen. Des Weiteren durften die Patienten keinerlei antiretrovirale Vorbehandlung bekommen haben; die Sensitivität gegenüber den Wirkstoffen FTC, TDF und EFV musste zusätzlich anhand eines genotypischen Resistenztests dokumentiert worden sein. Resistenz-assoziierte Mutationen gegen FTC, TDF und EFV wurden zusätzlich direkt ausgeschlossen (genotypischer Resistenztest). Patienten mit akuter oder dekompensierter chronischer Hepatitis, Niereninsuffizienz (Bedarf einer Dialyse, Serum Kreatinin-Konzentration größer als die 2fache Obergrenze des Normbereichs oder geschätzte Kreatinin-Clearance beim Screening ≤ 30 mL/min nach Cockcroft-Gault) oder einer medizinischen Erkrankung, die die Durchführung oder Interpretation der Studie beeinflusst hätte, wurden ausgeschlossen.

Die Patienten im RAL+FTC/TDF-Arm waren im Mittel 37,6 Jahre alt und im EFV+FTC/TDF-Arm lag das durchschnittliche Alter bei 36,9 Jahren. 81,0% der Patienten im RAL+FTC/TDF-Arm und 82,0% der Patienten im EFV+FTC/TDF-Arm waren männlich. Der Baseline BMI lag in beiden Behandlungsarmen bei 24 kg/m^2 (siehe Tabelle 4-131).

41,0% der RAL+FTC/TDF-Patienten und 44,0% der EFV+FTC/TDF-Patienten waren kaukasischer Abstammung, während 13,0% bzw. 11,0% der Patienten asiatischer Abstammung waren. Außerdem waren 12,0% der RAL+FTC/TDF-Patienten und 8,0% der EFV-Patienten schwarz und es gab je einen indianischen Patienten pro Behandlungsarm. Außerdem waren 12,0% der RAL+FTC/TDF-Patienten und 13,0% der EFV+FTC/TDF-Patienten gemischtrassig. In der Studie wurden hispanische Patienten unter dem Merkmal „Rasse“ aufgeführt. Hispanisch waren 21,0% der RAL+FTC/TDF-Patienten bzw. 24,0% der EFV+FTC/TDF-Patienten (siehe Tabelle 4-132 und Tabelle 4-133).

Zu Baseline lag die durchschnittliche Viruslast in beiden Armen bei $5,0 \log_{10}$ Kopien/mL; 45,0% bzw. 49,0% der RAL+FTC/TDF- bzw. EFV+FTC/TDF-Patienten wiesen eine Ausgangsviruslast ≤ 100.000 Kopien/mL auf (siehe Tabelle 4-134).

Die CD4-Zellzahl zu Baseline lag im RAL+FTC/TDF-Arm bei $218,9 \text{ Zellen}/\mu\text{L}$ und im EFV+FTC/TDF-Arm bei $217,4 \text{ Zellen}/\mu\text{L}$. 53,0% der RAL+FTC/TDF-Patienten bzw. 51,0% der EFV+FTC/TDF-Patienten wiesen eine CD4-Zellzahl zu Baseline $>200 \text{ Zellen}/\mu\text{L}$ auf, während jeweils 37,0% der Patienten eine CD4-Zellzahl zu Baseline zwischen 51 und $200 \text{ Zellen}/\mu\text{L}$ aufwiesen und 10,0% der RAL+FTC/TDF-Patienten und 11,0% der EFV+FTC/TDF-Patienten hatten eine CD4-Zellzahl zu Baseline $\leq 50 \text{ Zellen}/\mu\text{L}$ (siehe Tabelle 4-179).

Insgesamt beendeten 91,5% der Patienten im RAL+FTC/TDF-Arm die 48 Wochen in der Studie und im EFV+FTC/TDF-Arm waren dies 87,6%. Hauptgrund für einen vorzeitigen Studienabbruch waren unerwünschte Ereignisse (RAL+FTC/TDF: 2,8% vs. EFV+FTC/TDF: 6,0%). Außerdem gaben jeweils 3,2% der Patienten andere Gründe für den vorzeitigen Studienabbruch an. 87,2% der Patienten im RAL+FTC/TDF-Arm beendeten 96 Wochen der Studie und im EFV+FTC/TDF-Arm waren dies 82,3%. Hauptgründe für einen vorzeitigen Studienabbruch waren lost-to-follow-up (RAL+FTC/TDF: 1,4% vs. EFV+FTC/TDF: 4,3%) und andere Gründe (RAL+FTC/TDF: 6,0% vs. EFV+FTC/TDF: 5,0%).

Alle randomisierten Patienten, die mit der Studienmedikation behandelt wurden, wurden in die Wirksamkeits- und Sicherheitsanalysen zu Woche 48 und 96 eingeschlossen.

In der STARTMRK-Studie war je ein Patient pro Behandlungsarm zu Woche 48 nicht verfügbar, zu Woche 96 allerdings schon. Aus diesem Grund unterscheiden sich die Patientenzahlangaben zu Woche 48 und 96 bei den Endpunkten.

Insgesamt sind die Behandlungsarme untereinander balanziert.

In der STARTMRK-Studie wurden etwas mehr Frauen eingeschlossen als in den EVG/COBI/FTC/TDF-Studien und der Anteil der kaukasischen Patienten ist etwas geringer. Hinsichtlich der Ausgangsviruslast gibt es keine Unterschiede zwischen den Studien. Allerdings weisen mehr Patienten in der STARTMRK-Studie eine Ausgangsviruslast >100.000 Kopien/mL auf als in den EVG/COBI/FTC/TDF-Studien. Dies spiegelt sich entsprechend in der CD4-Zellzahl zu Baseline wider. In der STARTMRK-Studie liegt die durchschnittliche CD4-Zellzahl zu Baseline unter der in den EVG/COBI/FTC/TDF Studien.

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 4-137: 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			
S102	ja	ja	ja	ja	nein	nein	niedrig
S104	ja	ja	ja	ja	nein	nein	niedrig
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF							
STARTMRK	ja	ja	ja	ja	unklar*	nein	niedrig
* Da kein Studienprotokoll oder statistischer Analyseplan vorliegt können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden.							

Begründen Sie für jede Studie die abschließende Einschätzung.

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

In beiden Studien S102 und S104 werden keine Aussagen zur Erzeugung der Randomisierungsliste oder zur Verdeckung der Gruppenzuteilung gemacht. Da jedoch in beiden Studien ein IVRS/IWRS zur Randomisierung verwendet wurde, wird davon ausgegangen dass die Erzeugung der Randomisierungsliste adäquat durchgeführt wurde und dass die Gruppenzuteilung verdeckt war.

In beiden Studien wurde eine nach HIV-1-Viruslast zum Zeitpunkt des Screenings (≤ 100.000 oder >100.000 Kopien/mL) stratifizierte Randomisierung durchgeführt. Dabei wurde in Studie S102 eine Blocklänge von 4 für die Randomisierung verwendet und in Studie S104 betrug die Blocklänge 6. Das Zuteilungsverhältnis war in Studie S102 1:1 und in Studie S104 2:1. Es waren in beiden Studien sowohl die Patienten als auch die Prüfarzte gegenüber der Gruppenzugehörigkeit der Patienten verblindet.

In Studie S104 wurde die Imputationsmethode Missing/ ART Switch=Failure zur Berechnung des Anteils der virologischen Ansprecher erst im statistischen Analyseplan zu Woche 96 definiert. Da der statistische Analyseplan zu Woche 96 erst nach dem Daten Cut-Off für die Woche 48-Analyse angefertigt wurde, kann eine ergebnisgesteuerte Berichterstattung nicht ausgeschlossen werden. Außerdem wurden unerwünschte Ereignisse von Interesse weder im Protokoll noch im Analyseplan präspezifiziert. Auch hier kann eine ergebnisgesteuerte Berichterstattung nicht ausgeschlossen werden. Sonstige das Verzerrungspotenzial beeinflussende Faktoren gibt es nicht.

Es konnten in Studie S102 keine Hinweise auf ergebnisgesteuerte Berichterstattung gefunden werden und es gab auch keine sonstigen das Verzerrungspotenzial beeinflussenden Punkte. Insgesamt wird das Verzerrungspotenzial auf Studienebene in beiden Studien als niedrig eingestuft.

Für die unterschiedlichen Analysezeitpunkte die im vorliegenden Nutzendossier dargestellt werden, wurden jeweils die zulassungsrelevanten Daten verwendet. D.h. für Studie S102 gilt: Für Endpunkte zu Woche 48 wurde der Daten Cut-Off zu Woche 48 verwendet und für Endpunkte zu Woche 96 wurde der Daten Cut-Off zu Woche 96 verwendet. Für Studie S104 wurde der Woche 96-Daten Cut-Off verwendet.

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

In der STARTMRK-Studie wurden die Patienten gemäß einem Computer-generierten randomisierten Zuteilungsplan auf die Behandlungen randomisiert. Dabei wurde ein zentrales IVRS verwendet. Die Randomisierungssequenz wurde somit adäquat erzeugt. Die Gruppenteilung war für die Prüfarzte, das Studienpersonal, die Patienten, die Monitoren und das Zentrallaborpersonal verdeckt.

Da die Studie in einem doppelblinden Design durchgeführt wurde, war die Gruppenteilung der Patienten zu jeder Zeit sowohl den Prüfarzten bzw. Behandlern als auch den Patienten unbekannt.

Da kein Studienprotokoll oder statistischer Analyseplan vorliegt, können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden. Es können keine sonstigen das Verzerrungspotenzial beeinflussenden Punkte gefunden werden.

Das Verzerrungspotenzial wird für diese Studie als niedrig eingestuft.

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 4-138: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Virologisches Ansprechen	Zeit bis zum Verlust des virologischen Ansprechens	Veränderung der CD4-Zellzahl	Mortalität ^a
S102	ja	ja	ja	ja
S104	ja	ja	ja	ja
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF				
STARTMRK	ja	nein	ja	ja
^a Die Endpunkte Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse und häufigste unerwünschte Ereignisse (SOC) werden im Folgenden in einem gemeinsamen Abschnitt "Unerwünschte Ereignisse" dargestellt. Das Verzerrungspotenzial wird für diese Endpunkte pro Studie gemeinsam bewertet und unter Abschnitt 4.3.2.1.3.1.3 detailliert diskutiert. Es ist möglich diese Endpunkte pro Studie zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und pro Studie die gleiche Anzahl an Patienten in die Analyse einging.				

Tabelle 4-139: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Gesamtrate unerwünschter Ereignisse ^a	Schwerwiegende unerwünschte Ereignisse ^a	Unerwünschte Ereignisse Grad 3-4	Therapieabbrüche aufgrund unerwünschter Ereignisse ^a
S102	ja	ja	ja	ja
S104	ja	ja	ja	ja
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF				
STARTMRK	ja	ja	nein	ja
^a Die Endpunkte Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse und häufigste unerwünschte Ereignisse (SOC) werden im Folgenden in einem gemeinsamen Abschnitt "Unerwünschte Ereignisse" dargestellt. Das Verzerrungspotenzial wird für diese Endpunkte pro Studie gemeinsam bewertet und unter Abschnitt 4.3.2.1.3.1.3 detailliert diskutiert. Es ist möglich diese Endpunkte pro Studie zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und pro Studie die gleiche Anzahl an Patienten in die Analyse einging.				

Tabelle 4-140: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Wichtige unerwünschte Ereignisse ^{a,b}	Unerwünschte Ereignisse von Interesse	Häufigste unerwünschte Ereignisse (SOC) ^a
S102	ja	ja	ja
S104	ja	ja	ja
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF			
STARTMRK	ja*	nein	ja

^a Die Endpunkte Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse und häufigste unerwünschte Ereignisse (SOC) werden im Folgenden in einem gemeinsamen Abschnitt "Unerwünschte Ereignisse" dargestellt. Das Verzerrungspotenzial wird für diese Endpunkte pro Studie gemeinsam bewertet und unter Abschnitt 4.3.2.1.3.1.3 detailliert diskutiert. Es ist möglich diese Endpunkte pro Studie zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und pro Studie die gleiche Anzahl an Patienten in die Analyse einging.

^b Neurologische & psychiatrische Ereignisse wurden in STARTMRK-Studie unter einer anderen Gruppierung als neuropsychiatrische Ereignisse erfasst

4.3.2.1.3.1 Endpunkte – indirekte Vergleiche aus RCT

4.3.2.1.3.1.1 Virologisches Ansprechen – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

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

Tabelle 4-141: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Referenzen Studien	EVG/COBI/FTC/TDF	EFV+FTC/TDF	RAL+FTC/TDF
2	S102 S104	• •	• •	
1	STARTMRK		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

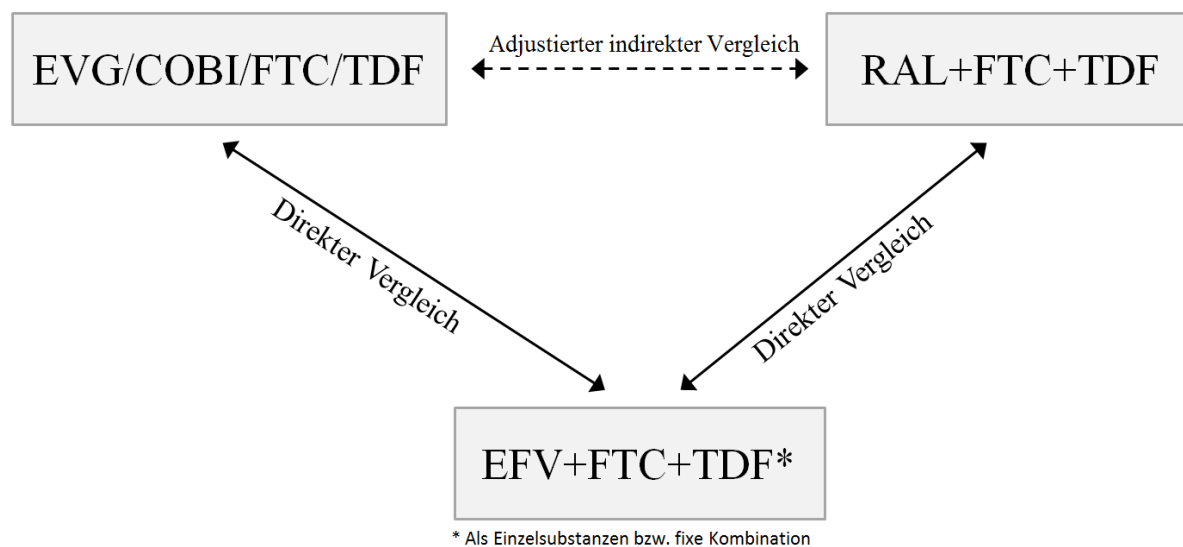


Abbildung 12: Netzwerkstruktur des indirekten Vergleichs zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF

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

Tabelle 4-142: Operationalisierung von virologisches Ansprechen

Studie	Operationalisierung
S102	Anteil an Patienten mit Viruslast <50 Kopien/mL zu Woche 48 bzw. 96. Zur Messung der HIV-1-RNA Plasma-Konzentrationen wurde der COBAS Amplicor HIV-1 Monitor-Test (Version 1.5) verwendet. Aus Vergleichbarkeitsgründen wird hier der Missing=Failure-Ansatz betrachtet.
S104	Anteil an Patienten mit Viruslast <50 Kopien/mL zu Woche 48. Zur Messung der HIV-1-RNA Plasma-Konzentrationen wurde der COBAS Amplicor HIV-1 Monitor-Test (Version 1.5) verwendet. Aus Vergleichbarkeitsgründen wird hier der Missing=Failure-Ansatz betrachtet.
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	
STARTMRK	Anteil an Patienten mit Viruslast <50 Kopien/mL zu Woche 48 bzw. 96. Zur Messung der HIV-1-RNA Plasma-Konzentration wurde der COBAS Amplicor HIV-1 Monitor Assay (Version 1.5; Roche Diagnostics, Branchburg, NJ, USA) verwendet. Das Quantifizierungslevel liegt bei 400 Kopien/mL. Die Anzahl an virologischen Ansprechern berechnet durch den Missing=Failure-Ansatz. Hierbei wurden fehlende HIV-1-RNA Messungen als Versagen gewertet, unabhängig von den Gründen des Fehlens. Wenn der letzte und erste Wert vor und nach der fehlenden Messung eine Viruslast <50 Kopien/mL zeigt, wird der fehlende Wert als fehlend gewertet. Bis Woche 16 wurde der Ultrasensitive Amplicor HIV-1 Monitor Assay (Version 1.5; Roche Diagnostics) mit niedrigem Quantifizierungslevel von 50 Kopien für alle Proben die laut Standard-Test unter 400 Kopien/mL lagen, verwendet. Ab Woche 16 wurde der ultrasensitive Test für alle Proben, die laut Standard-Test >50 Kopien/mL lagen, 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 4-143: Bewertung des Verzerrungspotenzials für virologisches Ansprechen 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
S102	niedrig	ja	ja	nein	nein	niedrig
S104	niedrig	ja	ja	nein	nein	niedrig
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF						
STARTMRK	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Studie zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Der Endpunkt wurde in beiden Studien, S102 und S104, definiert als Absinken der HIV-1-RNA Viruslast unter die Nachweisgrenze von 50 Kopien/mL bis Woche 48 und in Studie S102 auch bis Woche 96.

In Studie S102 wurde zur Messung der Plasma Viruslast der COBAS Amplicor HIV-1 Monitor-Test (Version 1,5) verwendet. In Studie S104 wurde anfangs der COBAS AmpliPrep/COBAS Taqman HIV-1 Test (Version 2.0) verwendet, der allerdings im ersten Amendment des Protokolls (29. Juli 2009) durch den COBAS Amplicor HIV-1 Monitor-Test (Version 1.5) ersetzt wurde.

Zur Vergleichbarkeit der Studienergebnisse wird für die beiden Studien S102 und S104 hier der Missing=Failure-Algorithmus verwendet. Der Missing=Failure-Algorithmus wurde in beiden Studien so operationalisiert, dass alle fehlenden HIV-1-RNA-Werte als Versagen gewertet werden.

Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung.

Das Verzerrungspotenzial wird für beide Studien als niedrig eingestuft.

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

Virologisches Ansprechen war in Studie STARTMRK definiert als Abfall der Viruslast unter die Nachweisgrenze von 50 Kopien/mL bis Woche 48 und 96.

In STARTMRK-Studie wurde der COBAS Amplicor HIV-1 Monitor Test (Version 1.5; Roche Diagnostics, Branchburg, NJ, USA) verwendet. Bis Woche 16 wurde der Ultrasensitive Amplicor HIV-1 Monitor Assay (Version 1.5; Roche Diagnostics) mit niedrigem Quantifizierungslevel von 50 Kopien für alle Proben die laut Standard-Test unter 400 Kopien/mL lagen, verwendet. Ab Woche 16 wurde der ultrasensitive Test für alle Proben, die laut Standard-Test >50 Kopien/mL lagen, verwendet.

Der Missing=Failure-Algorithmus wurde STARTMRK-Studie so operationalisiert, dass alle fehlenden HIV-1-RNA-Werte als Versagen gewertet werden. Die Analyse zum virologischen Versagen wurde wie im Protokoll geplant durchgeführt, und wurde konzipiert um Nichtunterlegenheit von RAL+FTC/TDF gegenüber EFV+FTC/TDF zur 12%-Grenze zu zeigen. Das Konfidenzintervall wurde mit der Methode von [118] errechnet.

In STARTMRK-Studie waren die Endpunkterheber verblindet. Das ITT-Prinzip wurde in allen drei Studien adäquat umgesetzt und es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung. Sonstige das Verzerrungspotenzial beeinflussende Punkte können nicht gefunden werden.

Die Operationalisierung und das Verzerrungspotenzial sind an beiden Zeitpunkten gleich.

Das Verzerrungspotenzial wird für die Studie als niedrig eingestuft.

Entsprechend der Auslegung des IQWiG in der Nutzenbewertung zu Rilpivirin (A12-04, Version 1.0, Stand 12.04.2012) wird in der vorliegenden Nutzenbewertung nur der Anteil an virologischen Ansprechern betrachtet, da davon ausgegangen wird, dass der Endpunkt virologisches Versagen (Wirksamkeit bzw. Resistenzen) hinreichend durch den Endpunkt virologisches Ansprechen abgedeckt ist [62].

Für Informationen zu Operationalisierung und Ergebnissen zu virologischem Versagen mittels Snapshot- bzw. TLOVR-Algorithmus siehe Modul 4A. Da dieser Endpunkt als durch virologisches Ansprechen abgedeckt angesehen wird, wird kein indirekter Vergleich zu diesem Endpunkt durchgeführt.

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 4-144: Ergebnisse für virologisches Ansprechen (M=F) aus RCT für indirekte Vergleiche (ITT Analyse-Set)

Zeitpunkt Studie	Intervention EVG/COBI/FTC/ TDF		Brücken- komparator EFV+FTC/TDF		Zweckmäßige Vergleichs- therapie RAL+FTC/TDF		Behandlungseffekt	
	N	n (%)	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert	RD [95%-KI], p-Wert
	Woche 48							
S102	348	309 (88,8)	352	301 (85,5)			1,34 [0,86;2,10], p=0,19 ^a	3,3 [-1,6;8,3]
S104	48	43 (89,6)	23	20 (87,0)			1,29 [0,28;5,94], p=0,72 ^a	2,8 [-15,1;20,8]
Woche 96								
S102	348	300 (86,2)	352	293 (83,2)			1,26 [0,83;1,90], p=0,27 ^a	3,0 [-2,4;8,3]
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF:								
Woche 48								
START MRK			281	230 (81,9)	280	241 (86,1)	1,37 [0,87;2,16], p=0,1743 ^b	4,2 [-1,9;10,3]
Woche 96								
START MRK			282	223 ^c (79,1)	281	228 ^d (81,1)	1,14 [0,75;1,72], p=0,5404 ^b	2,0 [-4,0;9,0]
^a aus Chochrane Mantel Heanzel Test, stratifiziert nach Baseline HIV-1-RNA Kategorie (≤ 100.000 vs. > 100.000 Kopien/mL) ^b zusätzlich berechnet ^c wurde aus einer gerundeten Prozentzahlangebe in der Publikation nachberechnet und von 222,78 aufgerundet. ^d wurde aus einer gerundeten Prozentzahlangebe in der Publikation nachberechnet und von 227,61 aufgerundet.								
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz								

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Es besteht zu Woche 48 in Studie S102 kein statistisch signifikanter Unterschied zwischen den Behandlungen (EVG/COBI/FTC/TDF vs. EFV/FTC/TDF: 88,8% vs. 85,5%). Das OR [95%-KI] beträgt hier 1,34 [0,86;2,10] mit p=0,19 (RR [95%-KI]: 1,04 [0,98;1,10], p=0,1949). Auch in Studie S104 (89,6% vs. 87,0%) kann kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (OR [95%-KI]: 1,29 [0,28;5,94], p=0,72; RR [95%-KI]: 1,03 [0,86;1,24], p=0,7530).

Zu Woche 96 besteht ebenso kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,26 [0,83;1,90], $p=0,27$; RR [95%-KI]: 1,04 [0,97;1,10], $p=0,2754$).

Allerdings kann in Studie S102 zu jedem Zeitpunkt die Nichtunterlegenheit zur 12% Grenze gezeigt werden. Lediglich in Studie S104 kann keine Nichtunterlegenheit zur 12% Grenze gezeigt werden.

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

In der STARTMRK-Studie besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,37 [0,87;2,16], $p=0,1743$; RR [95%-KI]: 1,05 [0,98;1,13], $p=0,1738$).

Zu Woche 96 ist der Unterschied zwischen den Behandlungen ebenfalls nicht statistisch signifikant (OR [95%-KI]: 1,14 [0,75;1,72], $p=0,5404$; RR [95%-KI]: 1,03 [0,94;1,11], $p=0,5404$).

Es kann jedoch zu beiden Zeitpunkten die Nichtunterlegenheit von RAL+FTC/TDF gegenüber EFV+FTC/TDF zur 12% Grenze gezeigt werden.

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.*

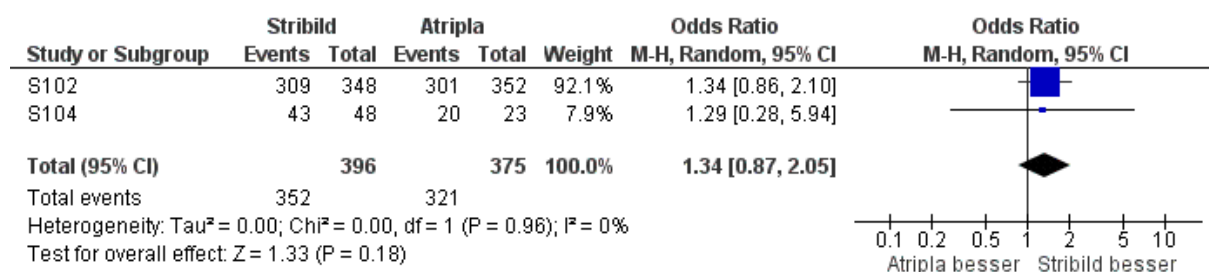


Abbildung 13: Meta-Analyse für virologisches Ansprechen (M=F) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, ITT-Set)

Der OR-Gesamtschätzer [95%-KI] liegt bei 1,34 [0,87;2,05] mit $p=0,18$, was keinen statistisch signifikanten Unterschied anzeigt (RR [95%-KI]: 1,04 [0,98;1,10], $p=0,18$). Die Datenlage ist dabei homogen ($p=0,96$; $I^2=0\%$). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse (OR [95%-KI]: 1,34 [0,87;2,05], $p=0,18$).

Da keine direkte und indirekte Evidenz vorlag, werden keine Konsistenzprüfungen durchgeführt.

Tabelle 4-145: Ergebnisse für virologisches Ansprechen (M=F) aus RCT für indirekte Vergleiche

Virologisches Ansprechen (M=F)	OR [95%-KI]	
	Woche 48	Woche 96
Direkter Vergleich EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	1,34 [0,87;2,05]	1,26 [0,83;1,90]
Direkter Vergleich RAL+FTC/TDF vs. EFV+FTC/TDF	1,37 [0,87;2,16]	1,14 [0,75;1,72]
Indirekter Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	0,98 [0,52;1,83]	1,11 [0,61;1,99]

KI = Konfidenzintervall, M=F: Missing=Failure, OR = Odds Ratio

Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF, zeigt zu Woche 48 keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,98 [0,52;1,83]; RR [95%-KI]: 0,99 [0,90;1,09]). Und auch zu Woche 96 zeigt sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,11 [0,61;1,99]; RR [95%-KI]: 1,01 [0,91;1,12]).

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.1.2 Veränderung der CD4-Zellzahl – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

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

Tabelle 4-146: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Referenzen Studien	EVG/COBI/FTC/TDF	EFV+FTC/TDF	RAL+FTC/TDF
2	S102 S104	• •	• •	
1	STARTM RK		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

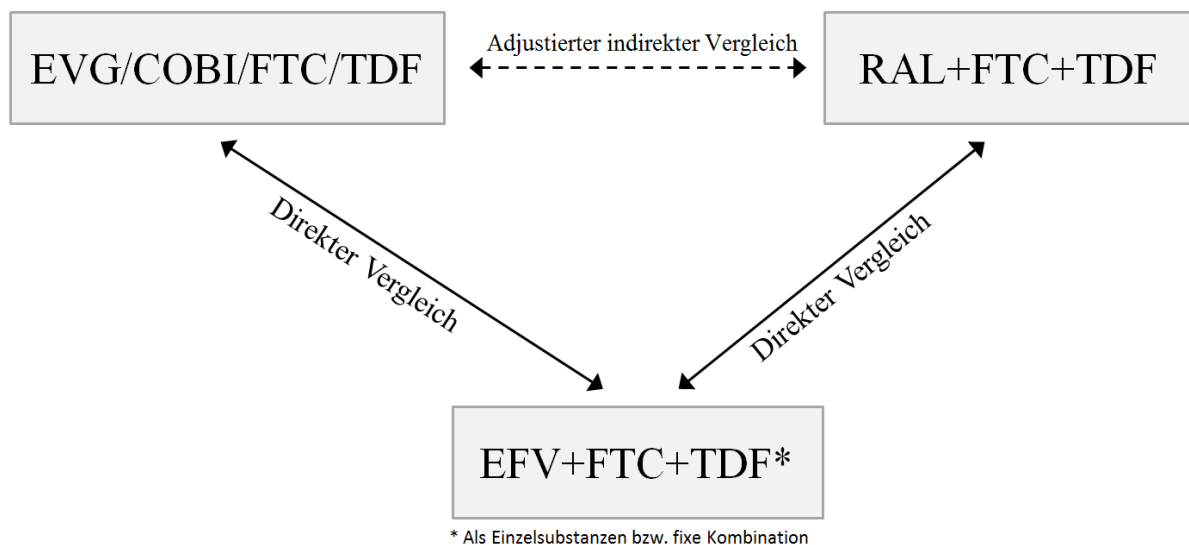


Abbildung 14: Netzwerkstruktur des indirekten Vergleichs zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF

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

Tabelle 4-147: Operationalisierung von Veränderung der CD4-Zellzahl

Studie	Operationalisierung
S102	Mittlere Veränderung der CD4-Zellzahl ab Baseline bis Woche 48 bzw. 96. Die CD4-Zellzahl wurde mittels Durchflusszytometrie erfasst.
S104	Mittlere Veränderung der CD4-Zellzahl ab Baseline bis Woche 48.
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	
STARTMRK	Mittlere Veränderung der CD4-Zellzahl ab Baseline bis Woche 48 und 96.

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 4-148: Bewertung des Verzerrungspotenzials für Veränderung der CD4-Zellzahl 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
S102	niedrig	unklar	ja	nein	nein	niedrig
S104	niedrig	unklar	ja	nein	nein	niedrig
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF						
STARTMRK	niedrig	ja	ja	unklar ^a	nein	niedrig
^a Da kein Studienprotokoll oder statistischer Analyseplan vorliegt können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden.						

Begründen Sie für jede Studie die abschließende Einschätzung.

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Der Endpunkt Veränderung der CD4-Zellzahl ist operationalisiert als Veränderung der mittleren Anzahl der CD4-Zellzahl zwischen Baseline und Woche 48 und Woche 96 (nur Studie S102). In Studie S102 wurde die CD4-Zellzahl mittels Durchflusszytometrie gemessen. Der Nichtberücksichtigungsanteil in Studie S102 lag zu Woche 48 bei ca. 6,6% im EVG/COBI/FTC/TDF-Arm und 10,5% im EFV/FTC/TDF-Arm und zu Woche 96 bei 11,8% bzw. 14,2%. In Studie S104 wurden 6,3% der Patienten im EVG/COBI/FTC/TDF-Arm und 8,7% im EFV/FTC/TDF-Arm in der Analyse zu Woche 48 nicht berücksichtigt. Allerdings

sind die Studienabbrecher detailliert beschrieben (siehe Tabelle 4-136), so dass eine Abschätzung des Einflusses möglich ist. In Studie S102 werden die Observed-Values betrachtet und in Studie S104 werden darüber keine Aussagen gemacht.

Es wurden keine Angaben gemacht, ob das Labor, welches die Proben zur Messung der CD4-Zellzahl untersuchte, gegenüber der Behandlung verblindet war. Da aber die Beeinflussbarkeit der Labortestergebnisse durch Kenntnis der Gruppenzugehörigkeit als gering eingeschätzt wird, wirkt sich dies nicht auf die Bewertung des Verzerrungspotenzials aus. Es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Punkte. Das Verzerrungspotenzial wird somit für beide Studien als niedrig eingestuft.

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

Der Endpunkt Veränderung der CD4-Zellzahl ist operationalisiert als Veränderung der mittleren Anzahl der CD4-Zellzahl zwischen Baseline und Woche 48 und Woche 96.

Für Patienten, die aufgrund fehlender Wirksamkeit die Studie abbrechen, wurden die Baseline-Werte fortgeschrieben (Observed-Failure-Methode). Patienten, die aus anderen Gründen die Studie abbrechen, wurden im Weiteren nicht mehr in die Analyse aufgenommen.

In STARTMRK-Studie waren die Endpunkterheber verblindet. Der Nichtberücksichtigungsanteil lag zu Woche 48 bei 8,2% im RAL+FTC/TDF-Arm und bei 10,7% im EFV+FTC/TDF-Arm bzw. zu Woche 96 bei 11,0% im RAL+FTC/TDF- und 13,8% im EFV+FTC/TDF-Arm. Hier wurde der Patientenfluss detailliert beschrieben. Das ITT-Prinzip wurde in der Studie somit adäquat umgesetzt. Da kein Studienprotokoll oder statistischer Analyseplan vorliegt können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden. Sonstige das Verzerrungspotenzial beeinflussende Punkte können nicht gefunden werden.

Das Verzerrungspotenzial wird für die Studie als niedrig eingestuft.

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 4-149: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT für indirekte Vergleiche

STARTMRK Endpunkt	Baseline		Woche x ^a		Veränderung bis Woche x		Behandlungseffekt MWD [95%-KI], p- Wert
	N	MW(SD)	N	MW(SD)	N	MW(SD)	
Woche 48							
S102							
EVG/COBI/FTC /TDF	348	391 (188,6)	325	631 (249,1)	325	239 (167,2)	Adj. ^b MWD: 33 [8;58], p=0,009 ^a
EFV/FTC/TDF	352	382 (170,2)	315	586 (213,9)	315	206 (153,4)	
S104							
EVG/COBI/FTC /TDF	48	378 (179,8)	45	619 (269,5)	45	240 (172,8)	Adj. ^b MWD: 74 [-15;164], p=0,10 ^a
EFV/FTC/TDF	23	454 (137,6)	21	628 (201,7)	21	166 (158,5)	
Woche 96							
S102:							
EVG/COBI/FTC /TDF	348	391 (188,6)	307	688 (263,6)	307	295 (213,3)	Adj. ^b MWD: 22 [-10;54], p=0,19 ^a
EFV/FTC/TDF	352	382 (170,2)	302	655 (237,1)	302	273 (189,7)	
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF							
Woche 48							
STARTMRK							
RAL+FTC/TDF	281	218,9 (124,2)	258	k.A.	258	189 (122,9 ^c)	26 [4;47], p=0,0184
EFV+FTC/TDF	282	217,4 (133,6)	251	k.A.	251	163 (121,2 ^c)	
Woche 96							
STARTMRK							
RAL+FTC/TDF	281	218,9 (124,2)	250	k.A.	250	240 (153,3 ^c)	15 [-13;42], p=0,2745 ^d
EFV+FTC/TDF	282	217,4 (133,6)	243	k.A.	243	225 (151,1 ^c)	
^a für Studie S102 und S104: Quelle: [119-121]							
^b S102: ANOVA adjustiert für Baseline HIV-1-RNA Level; S104: aus ANOVA adjustiert für Baseline HIV-1-RNA Kategorie (≤100.000 und >100.000 Kopien/mL)							
^c wurde aus dem angegebenen Konfidenzintervall berechnet (mittels Normalapproximation)							
^d zusätzlich berechnet							
Woche x steht für Woche 48 oder Woche 96							
k.A. = keine Angaben, KI = Konfidenzintervall, MWD = Mittelwertdifferenz, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis							

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

In Studie S102 hatten die Patienten im EVG/COBI/FTC/TDF-Arm und unter Verwendung des ITT Analyse-Sets zu Baseline eine durchschnittliche CD4-Zellzahl von 391 Zellen/ μ L. Diese konnte bis Woche 48 um durchschnittlich 239 Zellen/ μ L (SD: 167,2) auf 631 Zellen/ μ L (SD: 249,1) gesteigert werden. Im EFV/FTC/TDF-Arm wiesen die Patienten zu Baseline durchschnittlich eine CD4-Zellzahl von 382 Zellen/ μ L (SD: 170,2) auf; diese stieg bis Woche 48 um durchschnittlich 206 Zellen/ μ L (SD: 153,4) auf 586 Zellen/ μ L (SD: 213,9) an. Daraus ergibt sich eine adj. MWD [95%-KI] von 33 [8;58] mit $p=0,009$, womit ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF besteht. In Studie S104 wiesen die Patienten zu Baseline eine durchschnittliche CD4-Zellzahl von 378 Zellen/ μ L (SD: 179,8) im EVG/COBI/FTC/TDF-Arm bzw. 454 Zellen/ μ L (SD: 137,6) im EFV/FTC/TDF-Arm auf. Im EVG/COBI/FTC/TDF-Arm erhöhte sich die CD4-Zellzahl um durchschnittlich 240 Zellen/ μ L (SD: 172,8) auf 619 Zellen/ μ L (SD: 269,5), wohingegen die CD4-Zellzahl der EFV/FTC/TDF-Patienten lediglich um 166 Zellen/ μ L (SD: 158,5) auf 628 Zellen/ μ L (SD: 201,7) anstieg. Der Unterschied zwischen den beiden Behandlungen ist allerdings nicht signifikant (adj. MWD [95%-KI]: 74 [-15;164], $p=0,10$).

Nach 96 Wochen Behandlung erhöhte sich die CD4-Zellzahl der EVG/COBI/FTC/TDF-Patienten durchschnittlich um 295 Zellen/ μ L (SD: 213,3) auf 688 Zellen/ μ L (SD: 263,6) und im EFV/FTC/TDF-Arm erhöhte sich die CD4-Zellzahl der Patienten um durchschnittlich 273 Zellen/ μ L (SD: 189,7) auf 655 Zellen/ μ L (SD: 237,1).

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

In der STARTMRK-Studie besteht zu Woche 48 hinsichtlich der Veränderung der CD4-Zellzahl ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF (MWD [95%-KI]: 26 [4;47], $p=0,0184$). Dieser Unterschied besteht zu Woche 96 nicht; die MWD [95%-KI] beträgt zu Woche 96 15 [-13;42].

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.*

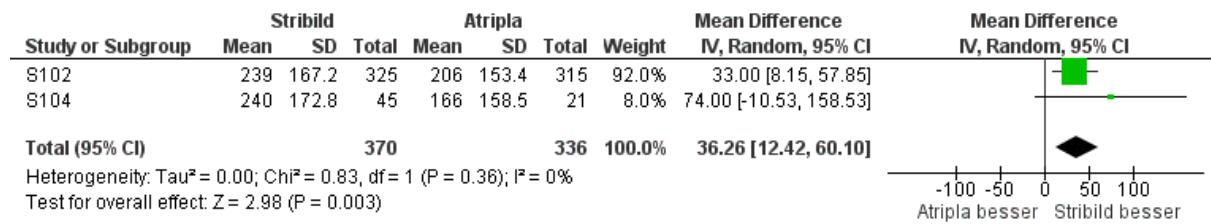


Abbildung 15: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, MWD, ITT)

Die Meta-Analyse zeigt – bei homogener Datenlage (Heterogenitätstest: p=0,36, I²=0%) – einen statistisch signifikant höheren Anstieg der CD4-Zellzahl im EVG/COBI/FTC/TDF-Arm gegenüber dem EFV/FTC/TDF-Arm. Der OR-Gesamtschätzer [95%-KI] beträgt 36,26 [12,42;60,10] mit p=0,003. Eine Sensitivitätsanalyse, unter Verwendung eines FEM, zeigt robuste Ergebnisse (OR [95%-KI]: 36,26 [12,42;60,10], p=0,003).

In der folgenden Abbildung ist die Meta-Analyse mittels Hedges' g dargestellt.

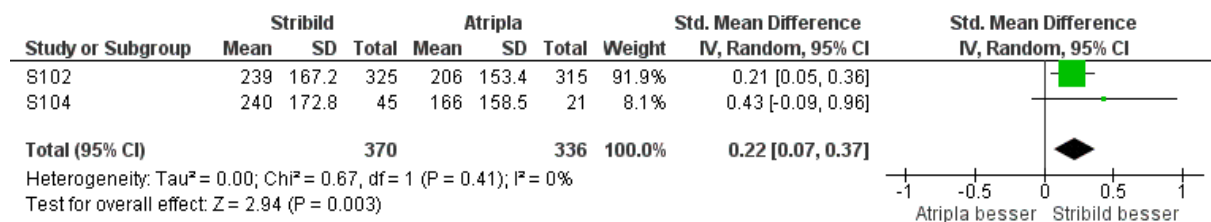


Abbildung 16: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, Hedges' g, ITT)

Die Meta-Analyse mittels Hedges' g zeigt einen statistisch signifikanten, aber, gemäß IQWiG Methoden 4,0 [76], klinisch nicht relevanten Effekt (Hedges' g [95%-KI]: 0,22 [0,07;0,37]), da die Untergrenze des Konfidenzintervalls die Grenze von 0,2 nicht übersteigt.

Da keine direkte und indirekte Evidenz vorlag, werden keine Konsistenzprüfungen durchgeführt.

Tabelle 4-150: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT für indirekte Vergleiche (Woche 48)

Veränderung der CD4-Zellzahl	MWD [95%-KI]	
	Woche 48	Woche 96

Veränderung der CD4-Zellzahl	MWD [95%-KI]	
	Woche 48	Woche 96
Direkter Vergleich EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	36,26 [12,42;60,10]	22 [-10;54]
Direkter Vergleich RAL+FTC/TDF vs. EFV+FTC/TDF	26 [4;47]	15 [-13;42]
Indirekter Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	10,26 [-21,84;42,36]	7 [-35,19;49,19]
KI = Konfidenzintervall, OR = Odds Ratio		

Der indirekte Vergleich zeigt, dass hinsichtlich der Veränderung der CD4-Zellzahl zu Woche 48 zwischen den beiden Behandlungen EVG/COBI/FTC/TDF und RAL+FTC/TDF kein statistisch signifikanter Unterschied besteht (MWD [95%-KI]: 10,26 [-21,84;42,36]). Und auch zu Woche 96 ist der Unterschied zwischen den beiden Behandlungen statistisch nicht signifikant (MWD [95%-KI]: 7 [-35,19;49,19]).

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.

4.3.2.1.3.1.3 Unerwünschte Ereignisse – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

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

Tabelle 4-151: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Referenzen Studien	EVG/COBI/FTC/TDF	EFV+FTC/TDF	RAL+FTC/TDF
2	S102 S104	• •	• •	
1	STARTMRK		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

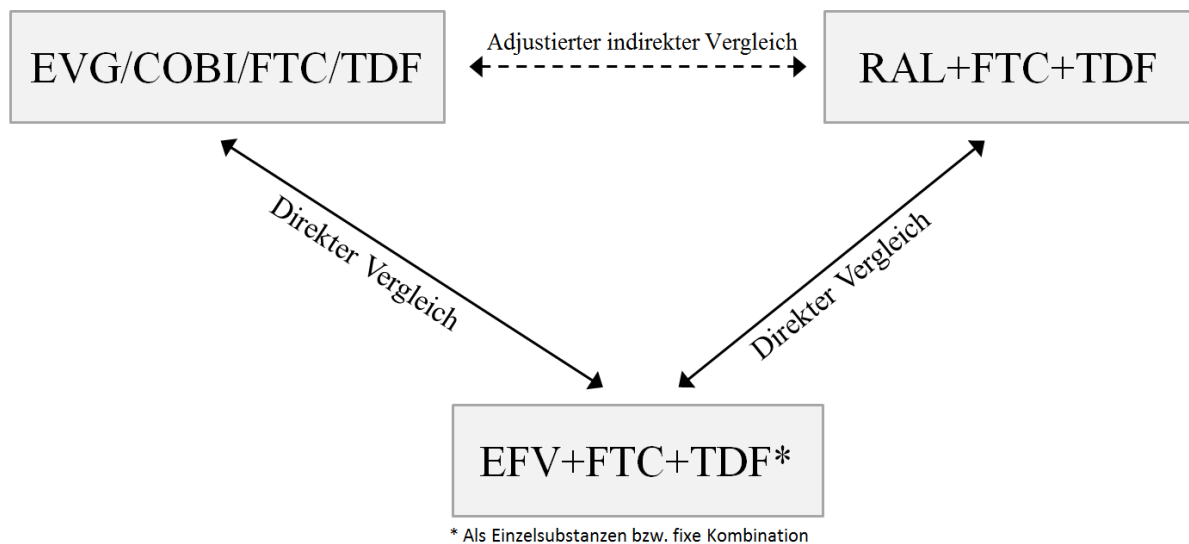


Abbildung 17: Netzwerkstruktur des indirekten Vergleichs zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF

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

Tabelle 4-152: Operationalisierung von unerwünschten Ereignissen

Studie	Operationalisierung
S102	<p><u>Mortalität</u> Anteil verstorbener Patienten im Behandlungszeitraum bis Woche 48 und 96 (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage)</p> <p><u>Gesamtrate unerwünschter Ereignisse</u> Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96</p> <p><u>Schwerwiegende unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96</p> <p><u>Therapieabbrüche aufgrund unerwünschter Ereignisse</u> Anteil an Patienten, die die Studie aufgrund eines unerwünschten Ereignisses abgebrochen haben (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96</p> <p><u>Neurologische & psychiatrische Ereignisse</u> Anteil an Patienten mit mindestens einem neurologischen & psychiatrischen Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96.</p> <p><u>Häufigste unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem häufigsten unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 96. Da die Operationalisierung der häufigsten unerwünschten Ereignisse (per SOC) in der STARTMRK-Studie zu Woche 48 von der in den beiden EVG/COBI/FTC/TDF-Studien abwich, wird kein indirekter Vergleich für die häufigsten unerwünschten Ereignisse (per SOC) zu Woche 48 durchgeführt. Ergebnisse für häufigste unerwünschte Ereignisse zu Woche 48 werden für die Studie S102 nicht dargestellt. Für Details zu den Ergebnissen dieser Endpunkte sei auf das Modul 4A verwiesen. Zu Woche 96 werden die häufigsten unerwünschten Ereignisse per SOC dargestellt. Darunter werden die PTs dargestellt, die bei mindestens 10% der Patienten in einem der Behandlungsarme auftraten.</p>
S104	<p><u>Mortalität</u> Anteil verstorbener Patienten im Behandlungszeitraum bis Woche 60 (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage)</p> <p><u>Gesamtrate unerwünschter Ereignisse</u> Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48</p> <p><u>Schwerwiegende unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48</p> <p><u>Therapieabbrüche aufgrund unerwünschter Ereignisse</u> Anteil an Patienten, die die Studie aufgrund eines unerwünschten Ereignisses abgebrochen haben (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48</p> <p><u>Wichtige unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem wichtigen unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48. Wichtige unerwünschte Ereignisse, die hier betrachtet werden, waren neurologische & psychiatrische Ereignisse.</p> <p><u>Häufigste unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem häufigsten unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48. Da die Operationalisierung der häufigsten unerwünschten Ereignisse (per SOC) in der</p>

Studie	Operationalisierung
	STARTMRK-Studie zu Woche 48 von der in den beiden EVG/COBI/FTC/TDF-Studien abwich, wird kein indirekter Vergleich für die häufigsten unerwünschten Ereignisse (per SOC) zu Woche 48 durchgeführt. Ergebnisse für häufigste unerwünschte Ereignisse zu Woche 48 werden für die Studie S104 nicht dargestellt.
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	
STARTMRK	<p><u>Mortalität</u> Anteil verstorbener Patienten im Behandlungszeitraum bis Woche 48 und 96 (Zeitraum zwischen erster und letzter Behandlungsdosis + 14 Tage)</p> <p><u>Gesamtrate unerwünschter Ereignisse</u> Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art bis Woche 48 und 96 (Zeitraum zwischen erster und letzter Behandlungsdosis + 14 Tage)</p> <p><u>Schwerwiegende unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis bis Woche 48 und 96 (Zeitraum zwischen erster und letzter Behandlungsdosis + 14 Tage)</p> <p><u>Therapieabbrüche aufgrund unerwünschter Ereignisse</u> Anteil an Patienten, die die Studie aufgrund eines unerwünschten Ereignisses abgebrochen haben bis Woche 48 und 96 (Zeitraum zwischen erster und letzter Behandlungsdosis + 14 Tage)</p> <p><u>Neuropsychiatrische Ereignisse</u> Anteil an Patienten mit mindestens einem neuropsychiatrischen Ereignis bis Woche 48 und 96 (Zeitraum zwischen erster und letzter Behandlungsdosis + 14 Tage).</p> <p><u>Häufigste unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem schwerwiegenden oder Therapie-assoziierten mittel- bis schwerwiegenden häufigsten unerwünschten Ereignis bis Woche 48, bzw. Anteil an Patienten mit mindestens einem häufigsten unerwünschten Ereignis bis Woche 96 (Zeitraum zwischen erster und letzter Behandlungsdosis + 14 Tage). Da die Operationalisierung der häufigsten unerwünschten Ereignisse (per SOC) zu Woche 48 in der STARTMRK-Studie von der in den beiden EVG/COBI/FTC/TDF-Studien abwich, wird kein indirekter Vergleich für die häufigsten unerwünschten Ereignisse (per SOC) zu Woche 48 durchgeführt. Zu Woche 96 werden die häufigsten unerwünschten Ereignisse per SOC dargestellt. Darunter werden die PTs dargestellt, die bei mindestens 10% der Patienten in einem der Behandlungsarme auftraten.</p>

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

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

Tabelle 4-153: 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
S102	niedrig	ja	ja	nein	nein	niedrig
S104	niedrig	ja	ja	unklar ^a	nein	niedrig
Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF						
STARTMRK	niedrig	ja	ja	unklar ^b	nein	niedrig
^a Wichtige unerwünschte Ereignisse (neurologische & psychiatrische Ereignisse) wurden in Studie S104 basierend auf verblindeten Daten derselben Studie und Studie S102 zu der Analyse hinzugefügt. Angaben zum Zeitpunkt der Aufnahme der Endpunkte wurden dabei nicht gemacht. Unerwünschte Ereignisse von Interesse wurden weder im Protokoll noch im statistischen Analyseplan präspezifiziert. Eine ergebnisgesteuerte Berichterstattung kann somit nicht ausgeschlossen werden. Das Verzerrungspotenzial wird bei den entsprechenden Endpunkten separat berücksichtigt. ^b Da kein Studienprotokoll oder statistischer Analyseplan vorliegt können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden.						

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

In den beiden Studien S102 und S104 waren Patienten und Endpunkterheber zu jeder Zeit gegenüber der Gruppenzugehörigkeit verblindet. Das ITT-Prinzip wurde adäquat umgesetzt. Wichtige unerwünschte Ereignisse (neurologische & psychiatrische Ereignisse) wurden in Studie S104 basierend auf verblindeten Daten derselben Studie und Studie S102 zu der Analyse hinzugefügt. Angaben zum Zeitpunkt der Aufnahme der Endpunkte wurden dabei nicht gemacht. Unerwünschte Ereignisse von Interesse wurden weder im Protokoll noch im statistischen Analyseplan präspezifiziert. Eine ergebnisgesteuerte Berichterstattung kann somit nicht ausgeschlossen werden. Es konnten keine sonstigen das Verzerrungspotenzial beeinflussende Punkte gefunden werden. Das Verzerrungspotenzial auf Endpunktebene wird hier für beide Studien als niedrig bewertet.

Unter unerwünschten Ereignissen wurden verschiedene sicherheitsbezogene Ereignisse betrachtet. Hier werden folgende unerwünschte Ereignisse betrachtet:

- Anteil verstorbener Patienten (Mortalität)
- Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art (Gesamtrate unerwünschter Ereignisse)
- Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Schwerwiegende unerwünschte Ereignisse)
- Anteil an Patienten die die Therapie aufgrund unerwünschter Ereignisse vorzeitig abbrechen (Therapieabbrüche aufgrund unerwünschter Ereignisse)

Zusätzlich dazu wurden wichtige unerwünschte Ereignisse betrachtet. Dies sind unerwünschte Ereignisse die besonders in Zusammenhang mit Efavirenz auftraten und in der Fachinformation von Efavirenz als übliche Nebenwirkungen dieser Therapie aufgelistet sind. In den beiden Studien wurden folgende unerwünschte Ereignisse als wichtig erachtet und werden hier betrachtet:

- Anteil an Patienten mit unerwünschten Ereignissen aus dem neurologischen bzw. psychiatrischen Bereich (Neurologische & psychiatrische Ereignisse):

Diese präspezifizierte Gruppe von PTs schloss die folgenden Terme ein: **abnorme Träume, akute Psychose**, affektive Erkrankung, Agitation, **Amnesie, Angst, Angststörung, bipolare Störung, vollendeter Suizid, Verwirrheitszustände, depressive Stimmung, Depression, Desorientierung**, Aufmerksamkeitsstörungen, **Schwindel**, euphorische Stimmung, Halluzination, **Kopfschmerzen, Schlaflosigkeit**, Reizbarkeit, **verringerte Libido, starke Depressionen, veränderte Stimmung, Stimmungsschwankungen, Nervosität, Alpträume, Panikattacke, Paranoia**, Unruhe, **Schlafstörungen, Somnolenz, Suizidgedanken, Suizidversuch und Gleichgewichtsstörung.**

Bei neurologischen & psychiatrischen Ereignissen wurden in Studie S104 einige PTs weniger in dieser Gruppe betrachtet. Die PTs die in beiden Studien eingeschlossen waren, sind oben fett gedruckt.

Da die Operationalisierung der häufigsten unerwünschten Ereignisse (per SOC) zu Woche 48 in der STARTMRK-Studie von der in den beiden EVG/COBI/FTC/TDF-Studien abwich, wird kein indirekter Vergleich für die häufigsten unerwünschten Ereignisse (per SOC) zu Woche 48 durchgeführt. Zu Woche 96 werden die häufigsten unerwünschten Ereignisse per SOC dargestellt. Darunter werden die PTs dargestellt, die bei mindestens 10% der Patienten in einem der Behandlungsarme auftraten. Für Details zu den ausführlichen Ergebnissen dieser Endpunkte zu Woche 48 und Woche 96 sei auf das Modul 4A verwiesen. Hier werden die folgenden SOC betrachtet:

- Erkrankungen des Gastrointestinaltrakts
- Allgemeine Erkrankungen und Beschwerden am Verabreichungsort
- Erkrankungen des Nervensystems
- Psychiatrische Erkrankungen
- Erkrankungen der Haut und des Unterhautzellgewebes

Die darunter am häufigsten auftretenden PTs (bei mindestens 10% der Patienten in einem der Behandlungsarme) werden dargestellt.

Die Endpunkterheber waren in beiden Studien verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung. Auch gibt es keine sonstigen, das Verzerrungspotenzial beeinflussenden Punkte. Somit ist das Verzerrungspotenzial in beiden Studien als niedrig einzustufen.

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

In der STARTMRK-Studie waren Patienten und Endpunkterheber zu jeder Zeit gegenüber der Gruppenzugehörigkeit verblindet. Das ITT-Prinzip wurde adäquat umgesetzt. Es können keine sonstigen das Verzerrungspotenzial beeinflussende Punkte gefunden werden. Das Verzerrungspotenzial auf Endpunktebene wird hier als niedrig bewertet.

Unter unerwünschten Ereignissen wurden verschiedene sicherheitsbezogene Ereignisse betrachtet. Diese waren die folgenden:

- Anteil verstorbener Patienten (Mortalität)
- Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art (Gesamtrate unerwünschter Ereignisse)
- Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Schwerwiegende unerwünschte Ereignisse)
- Anteil an Patienten die die Therapie aufgrund unerwünschter Ereignisse vorzeitig abbrechen (Therapieabbrüche aufgrund unerwünschter Ereignisse)

Daten zu diesen Endpunkten werden einer Publikation entnommen, welche ausschließlich Resultate zu der STARTMRK-Studie berichtet [111].

Zusätzlich dazu wurden neuropsychiatrische Ereignisse betrachtet. Dies sind unerwünschte Ereignisse, die besonders in Zusammenhang mit Efavirenz auftraten und in der Fachinformation von Efavirenz als übliche Nebenwirkungen dieser Therapie aufgelistet sind. Folgende Ereignisse wurden unter neuropsychiatrischen Ereignissen als Gruppe von PT betrachtet:

Depression, Alpträume, Verwirrheitszustände, Suizidgedanken, Nervensystemstörungen, psychotische Störung, abnorme Träume, Suizidversuch, akute Psychose, Delirium, eingeschränktes Bewusstsein, Halluzinationen, akustische Halluzination, vollendeter Suizid, und starke Depressionen.

Nach der Entblindung wurden weitere, für klinisch relevant befundene PT in die Gruppe der neuropsychiatrischen Ereignisse aufgenommen: Dies waren Schwindel, Schlaflosigkeit, Somnolenz und Konzentrationsstörungen. Es wurden Sensitivitätsanalysen unter Einschluss der zusätzlichen PT durchgeführt.

Die PT wurden aus der MedDRA Version 11.0 entnommen.

In der STARTMRK-Studie werden zu Woche 48 nur häufige unerwünschte Ereignisse dargestellt, die entweder in Verbindung mit der Behandlung gebracht wurden (möglicherweise, wahrscheinlich oder definitiv mit der Therapie assoziiert), schwerwiegend waren oder die in Verbindung mit der Behandlung gebracht wurden und von mittlerer bis moderater Schwere waren. Da sich diese Operationalisierungen nicht für einen indirekten Vergleich eignen, werden die Ergebnisse zu Woche 48 hier nur informativ dargestellt.

Daten zu den häufigsten unerwünschten Ereignissen zu Woche 96 werden einer Publikation [117] entnommen, in welcher die Daten zu Mortalität, der Gesamtrate unerwünschter Ereignisse, schwerwiegenden unerwünschten Ereignissen und Therapieabbrüchen aufgrund unerwünschter Ereignisse von den Daten in Lennox [111] abweichen. Diese Abweichung könnte zum Beispiel daran liegen, dass in der Publikation von Teppler [117] eine andere MedDRA Version verwendet wurde (Lennox: MedDRA Version 11.0, Teppler: MedDRA 12.0), als in der Publikation von Lennox [111]. Die hier interessierenden SOC umfassen

- Erkrankungen des Gastrointestinaltrakts
- Allgemeine Erkrankungen und Beschwerden am Verabreichungsort
- Erkrankungen des Nervensystems
- Psychiatrische Erkrankungen
- Erkrankungen der Haut und des Unterhautzellgewebes

Darunter werden je die am häufigsten auftretenden PT beobachtet. Dies sind die Erkrankungen, die bei mindestens 10% der Patienten in jeglichem Behandlungsarm auftraten. Ereignisraten für die gesamte SOC werden nicht angegeben.

Die Endpunkterheber waren in beiden Studien verblindet, das ITT-Prinzip wurde adäquat umgesetzt. Da kein Studienprotokoll oder statistischer Analyseplan vorliegt können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden. Es gibt keine sonstigen, das Verzerrungspotenzial beeinflussenden Punkte. Somit ist das Verzerrungspotenzial in beiden Studien als niedrig einzustufen.

Begründen Sie für jede Studie die abschließende Einschätzung.

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 4-154: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 48, Safety Set)

Endpunkt Studie	Intervention EVG/COBI/FTC/ TDF		Brücken- komparator EFV+FTC/TDF		Zweckmäßige Vergleichs- therapie RAL+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert ^a
	N	n (%)	N	n (%)	N	n (%)	
Mortalität							
S102	348	1 (0,3)	352	2 (0,6)			0,50 [0,05;5,59], p=0,5769
S104	48	0 (0,0)	23	0 (0,0)			n.b.
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	0 (0,0)	281	2 (0,7)	5,05 [0,24;105,74], p=0,2964

Endpunkt Studie	Intervention EVG/COBI/FTC/ TDF		Brücken- komparator EFV+FTC/TDF		Zweckmäßige Vergleichs- therapie RAL+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert ^a
	N	n (%)	N	n (%)	N	n (%)	
Gesamtrate unerwünschter Ereignisse							
S102	348	327 (94,0)	352	334 (94,9)			0,84 [0,44;1,60], p=0,5958
S104	48	44 (91,7)	23	20 (87,0)			1,65 [0,34;8,07], p=0,5364
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	272 (96,5)	281	253 (90,0)	0,33 [0,16;0,70], p=0,002 ^b
Gesamtrate schwerwiegender unerwünschter Ereignisse							
S102	348	41 (11,8)	352	24 (6,8)			1,83 [1,08;3,09], p=0,0253
S104	48	2 (4,2)	23	1 (4,3)			0,96 [0,08;11,12], p=0,9717
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	27 (9,6)	281	28 (10,0)	1,05 [0,60;1,82], p=0,888 ^b
Therapieabbrüche aufgrund unerwünschter Ereignisse							
S102	348	13 (3,7)	352	18 (5,1)			0,72 [0,35;1,49], p=0,3775
S104	48	0 (0,0)	23	1 (4,3)			0,15 [0,01;3,95], p=0,2587
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	17 (6,0)	281	9 (3,2)	0,52 [0,23;1,18], p=0,159 ^b
^a zusätzlich berechnet							
^b Fishers exakter Test							
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio							

Tabelle 4-155: Ergebnisse für wichtige unerwünschte Ereignisse (S102: neurologische & psychiatrische Ereignisse; STARTMRK: neuropsychiatrische Ereignisse) aus RCT für indirekte Vergleiche (Woche 48, Safety Set)

Studie	Intervention EVG/COBI/FTC/ TDF		Brücken- komparator EFV+FTC/TDF		Zweckmäßige Vergleichs- therapie RAL+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert ^b
	N	n (%)	N	n (%)	N	n (%)	
S102	348	149 (42,8)	352	220 (62,5)			0,45 [0,33;0,61], p<0,001
S104	48	20 (41,7)	23	13 (56,5)			0,55 [0,20;1,50], p=0,31
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK							
Prä- spezifizierter Endpunkt			282	65 (23,0)	281	38 (13,5)	0,52 [0,34;0,81], p=0,0044
Sensitivitäts-			282	165 (58,5)	281	73 (26,0)	0,25 [0,17;0,36], p<0,0001

Studie	Intervention		Brückenkomparator		Zweckmäßige Vergleichstherapie		Behandlungseffekt
	EVG/COBI/FTC/TDF		EFV+FTC/TDF		RAL+FTC/TDF		
	N	n (%)	N	n (%)	N	n (%)	OR ^a [95%-KI] ^a , p-Wert ^b
analyse							
^a zusätzlich berechnet							
^b Fisher's exakter Test							
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio							

Tabelle 4-156: Ergebnisse für die häufigsten schwerwiegenden unerwünschten Ereignisse (SOC) aus RCT für indirekte Vergleiche (Woche 48, Safety Set)

STARTMRK Endpunkt (SOC) Preferred term	RAL+FTC/TDF		EFV+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert ^a
	N	n (%)	N	n (%)	
Erkrankungen des Gastrointestinaltrakts (SOC)	281	3 (1,1)	282	3 (1,1)	1,00 [0,20;5,02], p=0,9965
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)	281	1 (0,4)	282	1 (0,4)	1,00 [0,06;16,12], p=0,9980
<i>Brustschmerzen</i>	281	1 (0,4)	282	1 (0,4)	1,00 [0,06;16,12], p=0,9980
Erkrankungen des Nervensystems (SOC)	281	2 (0,7)	282	4 (1,4)	0,50 [0,09;2,74], p=0,4233
Psychiatrische Erkrankungen (SOC)	281	2 (0,7)	282	4 (1,4)	0,50 [0,09;2,74], p=0,4233
<i>Psychische Störung</i>	281	1 (0,4)	282	1 (0,4)	1,00 [0,06;16,12], p=0,9980
<i>Depression</i>	281	1 (0,4)	282	1 (0,4)	1,00 [0,06;16,12], p=0,9980
^a zusätzlich berechnet					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Tabelle 4-157: Ergebnisse für die häufigsten mittel- bis schwerwiegenden Therapieassoziierten unerwünschten Ereignisse (PTs) aus RCT für indirekte Vergleiche (Woche 48, Safety Set)

STARTMRK Preferred term	RAL+FTC/TDF		EFV+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert ^a
	N	n (%)	N	n (%)	
<i>Diarrhoe</i>	281	3 (1,1)	282	8 (2,8)	0,37 [0,10;1,41], p=0,1446
<i>Übelkeit</i>	281	8 (2,8)	282	10 (3,5)	0,80 [0,31;2,05], p=0,6380
<i>Fatigue</i>	281	4 (1,4)	282	8 (2,8)	0,49 [0,15;1,66], p=0,2548
<i>Schwindel</i>	281	4 (1,4)	282	18 (6,4)	0,21 [0,07;0,63], p=0,0055
<i>Kopfschmerzen</i>	281	11 (3,9)	282	13 (1,8)	0,84 [0,37;1,92], p=0,6833
<i>Schlaflosigkeit</i>	281	10 (3,6)	282	9 (3,2)	1,12 [0,45;2,80], p=0,8094
^a zusätzlich berechnet					

STARTMRK	RAL+FTC/TDF		EFV+FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	
Preferred term					OR ^a [95%-KI] ^a , p-Wert ^a

KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio

Unerwünschte Ereignisse

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Bis Woche 48 verstarben in Studie S102 ein Patient (0,3%) im EVG/COBI/FTC/TDF-Arm und zwei Patienten (0,6%) im EFV/FTC/TDF-Arm. Der Unterschied zwischen den beiden Behandlungen ist statistisch nicht signifikant (OR [95%-KI]: 0,50 [0,05;5,59] p=0,5769, RR [95%-KI]: 0,51 [0,05;5,55], p=0,5771). In Studie S104 verstarb in den 48 Wochen kein Patient.

94,0% der Patienten im EVG/COBI/FTC/TDF-Arm und 94,9% der Patienten im EFV/FTC/TDF-Arm der Studie S102 hatten mindestens ein unerwünschtes Ereignis. Das OR [95%-KI] beträgt 0,84 [0,44;1,60] mit p=0,5958 (RR [95%-KI]: 0,99 [0,96;1,03], p=0,5956), was keinen statistisch signifikanten Unterschied anzeigt. Auch in Studie S104 besteht, bei 91,7% der EVG/COBI/FTC/TDF-Patienten und 87,0% der EFV/FTC/TDF-Patienten mit unerwünschten Ereignissen, kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,65 [0,34;8,07], p=0,5364; RR [95%-KI]: 1,05 [0,88;1,26], p=0,5653).

In Studie S102 hatten 11,8% der Patienten im EVG/COBI/FTC/TDF-Arm und 6,8% im EFV/FTC/TDF-Arm ein schwerwiegendes unerwünschtes Ereignis. Der Unterschied zwischen den beiden Behandlungen ist statistisch signifikant (OR [95%-KI]: 1,83 [1,08;3,09], p=0,0253; RR [95%-KI]: 1,73 [1,07;2,80], p=0,0260). In Studie S104 hatten 4,2% der EVG/COBI/FTC/TDF- und 4,3% der EFV/FTC/TDF-Patienten ein schwerwiegendes unerwünschtes Ereignis. Der Unterschied zwischen den Behandlungen ist nicht statistisch signifikant (OR [95%-KI]: 0,96 [0,08;11,12], p=0,9717; RR [95%-KI]: 0,96 [0,09;10,03], p=0,9717).

In Studie S102 brachen 3,7% der EVG/COBI/FTC/TDF-Patienten und 5,1% der EFV/FTC/TDF-Patienten die Behandlung aufgrund unerwünschter Ereignisse ab. Der Unterschied zwischen den Behandlungen ist jedoch nicht statistisch signifikant (OR [95%-KI]: 0,72 [0,35;1,49], p=0,3775; RR [95%-KI]: 0,73 [0,36;1,47], p=0,3778). Kein Patient im EVG/COBI/FTC/TDF-Arm, aber 4,3% der Patienten im EFV/FTC/TDF-Arm der Studie S104 brachen die Behandlung aufgrund unerwünschter Ereignisse ab. Der Unterschied zwischen den Behandlungen ist jedoch nicht statistisch signifikant (OR [95%-KI]: 0,15 [0,01;3,95], p=0,2587; RR [95%-KI]: 0,16 [0,01;3,86], p=0,2614).

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

In STARTMRK starben zwei Patienten im RAL+FTC/TDF-Arm. Das OR [95%-KI] beträgt 5,05 [0,24;105,74] mit p=0,2964 (RR [95%-KI]: 5,02 [0,24;104,05], p=0,2971), was keinen

statistisch signifikanten Unterschied anzeigt. Hinsichtlich der Gesamtrate unerwünschter Ereignisse ergibt sich ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF. Das OR [95%-KI] beträgt dabei 0,33 [0,16;0,70] mit $p=0,002$ (RR [95%-KI]: 0,93 [0,89;0,98], $p=0,0026$). In STARTMRK gibt es hinsichtlich schwerwiegender unerwünschter Ereignisse keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,05 [0,60;1,82], $p=0,888$; RR [95%-KI]: 1,04 [0,63;1,72], $p=0,8762$). Es gibt hinsichtlich Therapieabbrüche aufgrund unerwünschter Ereignisse keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,52 [0,23;1,18], $p=0,159$; RR [95%-KI]: 0,53 [0,24;1,17], $p=0,1170$).

Neurologische & psychiatrische bzw. neuropsychiatrische Ereignisse

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

In Studie S102 traten neurologische & psychiatrische Ereignisse bei 42,8% der EVG/COBI/FTC/TDF-Patienten und bei 62,5% der EFV/FTC/TDF-Patienten auf. Somit besteht hier ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Das OR [95%-KI] beträgt 0,45 [0,33;0,61] mit $p<0,001$ (RR [95%-KI]: 0,69 [0,59;0,79], $p<0,00001$). Auch in Studie S104 traten im EVG/COBI/FTC/TDF-Arm (41,7%) bei numerisch weniger Patienten neurologische & psychiatrische Ereignisse auf als im EFV/FTC/TDF-Arm (56,5%), allerdings ist dieser Unterschied nicht statistisch signifikant (OR [95%-KI]: 0,55 [0,20;1,50], $p=0,2426$; RR [95%-KI]: 0,74 [0,45;1,20], $p=0,2230$).

Da für den Endpunkt neurologische & psychiatrische Ereignisse das Verzerrungspotenzial in Studie S104 nicht sicher ausgeschlossen werden kann, ist die Aussagekraft dieser Endpunkte für Studie S104 niedriger einzustufen.

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

In STARTMRK besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF. Das OR [95%-KI] beträgt dabei 0,52 [0,34;0,81] mit $p=0,0044$ (RR [95%-KI]: 0,59 [0,41;0,84], $p=0,0041$). In der Sensitivitätsanalyse, unter Einschluss der zusätzlichen PTs Schwindel, Schlaflosigkeit, Somnolenz und Konzentrationsstörungen, ergibt sich ebenso ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF (OR [95%-KI]: 0,25 [0,17;0,36], $p<0,0001$; RR [95%-KI]: 0,44 [0,36;0,55], $p<0,00001$).

Häufigste unerwünschte Ereignisse

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

In STARTMRK wurden nur häufige schwerwiegende bzw. häufige mittel- bis schwerwiegende Therapie-assoziierte unerwünschte Ereignisse berichtet. Schwerwiegende *Erkrankungen des Gastrointestinaltrakts (SOC)* traten bei jeweils drei Patienten pro Behandlungsarm auf (OR [95%-KI]: 1,00 [0,20;5,02], $p=0,9965$; RR [95%-KI]: 1,00 [0,20;4,93], $p=0,9965$), was keinen statistisch signifikanten Unterschied anzeigt. Je ein Patient pro Behandlungsarm wies eine schwerwiegende *allgemeine Erkrankung oder Beschwerden*

am Verabreichungsort (SOC) auf (OR [95%-KI]: 1,00 [0,06;16,12], p=0,9980; RR [95%-KI]: 1,00 [0,06;15,97], p=0,9980); dies waren hier Brustschmerzen. Hinsichtlich schwerwiegender *Erkrankungen des Nervensystems (SOC)* besteht ebenso kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,50 [0,09;2,74], p=0,4233; RR [95%-KI]: 0,50 [0,09;2,72], p=0,4237). Schwerwiegende *psychiatrische Erkrankungen (SOC)* traten statistisch nicht signifikant unterschiedlich häufig in den Behandlungsarmen auf (OR [95%-KI]: 0,50 [0,09;2,74], p=0,4233; RR [95%-KI]: 0,50 [0,09;2,72], p=0,4237). Schwerwiegende Depression (OR [95%-KI]: 1,00 [0,06;16,12], p=0,9980; RR [95%-KI]: 1,00 [0,06;15,97], p=0,9980) trat ebenfalls gleichhäufig in den Behandlungsarmen auf. Für schwerwiegende psychische Störungen besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,00 [0,06;16,12], p=0,9980; RR [95%-KI]: 1,00 [0,06;15,97], p=0,9980).

Hinsichtlich des Therapie-assoziierten Auftretens von mittel- bis schwerwiegender Diarrhoe (OR [95%-KI]: 0,37 [0,10;1,41], p=0,1446; RR [95%-KI]: 0,38 [0,10;1,40], p=0,1457) und mittel- bis schwerwiegender Übelkeit (OR [95%-KI]: 0,80 [0,31;2,05], p=0,6380; RR [95%-KI]: 0,80 [0,32;2,00], p=0,6381), besteht ebenso kein statistisch signifikanter Unterschied zwischen den Behandlungen. Hinsichtlich Therapie-assoziiertes mittel- bis schwerwiegender Fatigue besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,49 [0,15;1,66], p=0,2548; RR [95%-KI]: 0,50 [0,15;1,65], p=0,2556). Therapie-assoziierte mittel- bis schwerwiegende Kopfschmerzen (OR [95%-KI]: 0,84 [0,37;1,92], p=0,6833; RR [95%-KI]: 0,85 [0,39;1,86], p=0,6834) traten statistisch nicht signifikant unterschiedlich häufig in den Behandlungen auf. Hinsichtlich Therapie-assoziiertem mittel- bis schwerwiegendem Schwindel besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF (OR [95%-KI]: 0,21 [0,07;0,63], p=0,0055; RR [95%-KI]: 0,22 [0,08;0,65], p=0,0060). Therapie-assoziiertem mittel- bis schwerwiegende Schlaflosigkeit (OR [95%-KI]: 1,12 [0,45;2,80], p=0,8094; RR [95%-KI]: 1,12 [0,46;2,70], p=0,8095) trat in beiden Behandlungsarmen gleich häufig auf.

Tabelle 4-158: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 96, Safety Set)

Endpunkt Studie	Intervention EVG/COBI/FTC/ TDF		Brücken- komparator EFV+FTC/TDF		Zweckmäßige Vergleichstherapie RAL+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert
	N	n (%)	N	n (%)	N	n (%)	
Mortalität							
S102	348	1 (0,3)	352	2 (0,6)			0,50 [0,05;5,59], p=0,5769 ^a
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	0 (0,0)	281	3 (1,1)	7,10 [0,37;138,10], p=0,1955 ^a
Gesamtrate unerwünschter Ereignisse							
S102	348	337 (96,8)	352	342 (97,2)			0,90 [0,38;2,14], p=0,8041 ^a
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	275 (97,5)	281	266 (94,7)	0,45 [0,18;1,12], p=0,086 ^b
Gesamtrate schwerwiegender unerwünschter Ereignisse							
S102	348	56 (16,1)	352	33 (9,4)			1,85 [1,17;2,93], p=0,0083 ^a
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	34 (12,1)	281	40 (14,2)	1,21 [0,74;1,98], p=0,457 ^b
Therapieabbrüche aufgrund unerwünschter Ereignisse							
S102	348	17 (4,9)	352	24 (6,8)			0,70 [0,37;1,33], p=0,2782 ^a
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	17 (6,0)	281	11 (3,9)	0,64 [0,29;1,38], p=0,333 ^b
^a zusätzlich berechnet ^b Fishers exakter Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio							

Tabelle 4-159: Ergebnisse für wichtige unerwünschte Ereignisse (S102: neurologische & psychiatrische Ereignisse; STARTMRK: neuropsychiatrische Ereignisse) aus RCT für indirekte Vergleiche (Woche 96, Safety Set)

Studie	Intervention EVG/COBI/FTC/ TDF		Brücken- komparator EFV+FTC/TDF		Zweckmäßige Vergleichs- therapie RAL+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert
	N	n (%)	N	n (%)	N	n (%)	
S102	348	162 (46,6)	352	232 (65,9)			0,45 [0,33;0,61], p<0,001 ^b
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	71 ^c (25,2)	281	46 ^d (16,4)	0,58 [0,38;0,88], p=0,0105 ^a
^a zusätzlich berechnet ^b Fisher's exakter Test ^c wurde aus einer gerundeten Prozentzahlerhöhung (gegenüber Woche 48) in der Publikation nachberechnet und von 5,64 aufgerundet. ^d wurde aus einer gerundeten Prozentzahlerhöhung (gegenüber Woche 48) in der Publikation nachberechnet und von 8,43 abgerundet. KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio							

Tabelle 4-160: Ergebnisse für häufigste unerwünschte Ereignisse (SOC, Grenze: ≥10%) aus RCT für indirekte Vergleiche (Woche 96, Safety Set)

Endpunkt Studie	Intervention EVG/COBI/ FTC/TDF		Brücken- komparator EFV+FTC/TDF		Zweckmäßige Vergleichs- therapie RAL+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert ^a
	N	n (%)	N	n (%)	N	n (%)	
Erkrankungen des Gastrointestinaltrakts (SOC)							
<i>Diarrhoe</i>							
S102	348	87 (25,0)	352	83 (23,6)			1,08 [0,76;1,53], p=0,6613
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	69 (24,5)	281	48 (17,1)	0,64 [0,42;0,96], p=0,0315
<i>Übelkeit</i>							
S102	348	76 (21,8)	352	53 (15,1)			1,58 [1,07;2,32], p=0,0213
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	36 (12,8)	281	40 (14,2)	1,13 [0,70;1,84], p=0,6102
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)							
<i>Fatigue</i>							
S102	348	45 (12,9)	352	52 (14,8)			0,86 [0,56;1,32], p=0,4810
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	33 (11,7)	281	19 (6,8)	0,55 [0,30;0,99], p=0,0454
Erkrankungen des Nervensystems (SOC)							

Endpunkt Studie	Intervention EVG/COBI/ FTC/TDF		Brücken- komparator EFV+FTC/TDF		Zweckmäßige Vergleichs- therapie RAL+FTC/TDF		Behandlungseffekt OR ^a [95%-KI] ^a , p-Wert ^a
	N	n (%)	N	n (%)	N	n (%)	
Schwindel							
S102	348	25 (7,2)	352	91 (25,9)			0,22 [0,14;0,36], p<0,00001
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	104 (36,9)	281	23 (8,2)	0,15 [0,09;0,25], p<0,00001
Kopfschmerzen							
S102	348	56 (16,1)	352	40 (11,4)			1,50 [0,97;2,31], p=0,0702
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	69 (24,5)	281	64 (22,8)	0,91 [0,62;1,34], p=0,6365
Psychiatrische Erkrankungen (SOC)							
Abnorme Träume							
S102	348	53 (15,2)	352	97 (27,6)			0,47 [0,32;0,69], p=0,0001
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	37 (13,1)	281	21 (7,5)	0,53 [0,30;0,94], p=0,0294
Schlaflosigkeit							
S102	348	38 (10,9)	352	57 (16,2)			0,63 [0,41;0,99], p=0,0428
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	31 (11,0)	281	34 (12,1)	1,11 [0,66;1,87], p=0,6813
Depression^b							
S102	348	42 (12,1)	352	49 (13,9)			0,85 [0,55;1,32], p=0,4668
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	25 (8,9)	281	21 (7,5)	0,83 [0,45;1,52], p=0,5470
Erkrankungen der Haut und des Unterhautzellgewebes (SOC)							
Hautausschlag							
S102	348	26 (7,5)	352	49 (13,9)			0,50 [0,30;0,82], p=0,0066
<i>Studie für den indirekten Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF</i>							
STARTMRK			282	34 (12,1)	281	17 (6,0)	0,47 [0,26;0,86], p=0,0148
^a zusätzlich berechnet							
^b Der PT Depression zu Woche 96 wird auch in der Publikation von Lennox berichtet (Ereignisrate in beiden Publikationen gleich)							
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio							

Unerwünschte Ereignisse

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Im EVG/COBI/FTC/TDF-Arm verstarb bis Woche 96 ein Patient (0,3%) und im EFV/FTC/TDF-Arm verstarben 2 Patienten (0,6%). Es besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,50 [0,05;5,59], p=0,5769; RR [95%-KI]: 0,51 [0,05;5,55], p=0,5771). Insgesamt hatten 337 (96,8%) der Patienten im EVG/COBI/FTC/TDF-Arm und 342 (97,2%) der Patienten im EFV/FTC/TDF-Arm ein unerwünschtes Ereignis jeglicher Art. Auch hier besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,90 [0,38;2,14], p=0,8041; RR [95%-KI]: 1,00 [0,97;1,02], p=0,8041). Hinsichtlich der schwerwiegenden unerwünschten Ereignisse wird ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (EVG/COBI/FTC/TDF: 16,1%; EFV/FTC/TDF: 9,4%), der sich zugunsten der EFV/FTC/TDF-Gruppe ausprägt, gefunden (OR [95%-KI]: 1,85 [1,17;2,93], p=0,0083; RR [95%-KI]: 1,72 [1,15;2,57], p=0,0087). 4,9% der Patienten im EVG/COBI/FTC/TDF-Arm und 6,8% der Patienten im EFV/FTC/TDF-Arm brachen die Therapie aufgrund unerwünschter Ereignisse ab (OR [95%-KI]: 0,70 [0,37;1,33], p=0,2782; RR [95%-KI]: 0,72 [0,39;1,31], p=0,2788).

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

Zu Woche 96 besteht hinsichtlich der Mortalitätsrate (RAL+FTC/TDF vs. EFV+FTC/TDF: 1,1% vs. 0,0%) kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 7,10 [0,37;138,10], p=0,1955; RR [95%-KI]: 7,02 [0,36;135,37], p=0,1966). Hinsichtlich der Gesamtrate unerwünschter Ereignisse (RAL+FTC/TDF vs. EFV+FTC/TDF: 94,7% vs. 97,5%) kann zu Woche 96 ebenso kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (OR [95%-KI]: 0,45 [0,18;1,12], p=0,086; RR [95%-KI]: 0,97 [0,94;1,00], p=0,0814). Insgesamt traten unter RAL+FTC/TDF bei 14,2% der Patienten schwerwiegende unerwünschte Ereignisse auf, und im EFV+FTC/TDF-Arm waren dies 12,1%. Der Unterschied ist nicht statistisch signifikant (OR [95%-KI]: 1,21 [0,74;1,98], p=0,457; RR [95%-KI]: 1,18 [0,77;1,81], p=0,4451). Es besteht hinsichtlich der Therapieabbrüche aufgrund unerwünschter Ereignisse (RAL+FTC/TDF vs. EFV+FTC/TDF: 3,9% vs. 6,0%) kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,64 [0,29;1,38], p=0,333; RR [95%-KI]: 0,65 [0,31;1,36], p=0,2529).

Neurologische & psychiatrische bzw. neuropsychiatrische Ereignisse

Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Neurologische & psychiatrische Ereignisse wurden bei 46,6% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 65,9% der Patienten im EFV/FTC/TDF-Arm der Studie S102 beobachtet. Es traten somit statistisch signifikant weniger neurologische & psychiatrische Ereignisse im EVG/COBI/FTC/TDF-Arm als im EFV/FTC/TDF-Arm auf (OR [95%-KI]: 0,45 [0,33;0,61], p<0,001; RR [95%-KI]: 0,71 [0,62;0,81], p<0,00001).

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

Neuropsychiatrische Ereignisse wurden zu Woche 96 16,4% bzw. 25,2% der Patienten im RAL+FTC/TDF-Arm bzw. EFV+FTC/TDF-Arm beobachtet. Der Unterschied zwischen den Behandlungen ist statistisch signifikant zugunsten von RAL+FTC/TDF (OR [95%-KI]: 0,58 [0,38;0,88], $p=0,0105$; RR [95%-KI]: 0,65 [0,47;0,91], $p=0,0111$). Es werden keine Angaben darüber gemacht, ob die präspezifizierte oder die im Nachhinein ergänzte Gruppe von PTs verwendet wird.

Häufigste unerwünschte Ereignisse**Studien zu EVG/COBI/FTC/TDF vs. EFV/FTC/TDF**

Die unter *Erkrankungen des Gastrointestinaltrakts (SOC)* am häufigsten beobachteten Ereignisse waren Diarrhoe (EVG/COBI/FTC/TDF: 25,0%; EFV/FTC/TDF: 23,6%; OR [95%-KI]: 1,08 [0,76;1,53], $p=0,6613$; RR [95%-KI]: 1,06 [0,82;1,38], $p=0,6613$) und Übelkeit (EVG/COBI/FTC/TDF: 21,8%; EFV/FTC/TDF: 15,1%). Lediglich bei Übelkeit besteht ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen zugunsten von EFV/FTC/TDF (OR [95%-KI]: 1,58 [1,07;2,32], $p=0,0213$; RR [95%-KI]: 1,45 [1,06;1,99], $p=0,0219$).

Am häufigsten beobachtet wurde im Zusammenhang mit *allgemeinen Erkrankungen und Beschwerden am Verabreichungsort (SOC)* Fatigue (EVG/COBI/FTC/TDF: 12,9%; EFV/FTC/TDF: 14,8%; OR [95%-KI]: 0,86 [0,56;1,32], $p=0,4810$; RR [95%-KI]: 0,88 [0,60;1,27], $p=0,4812$).

Die häufigsten *Erkrankungen des Nervensystems (SOC)* waren Schwindel (EVG/COBI/FTC/TDF: 7,2%; EFV/FTC/TDF: 25,9%) und Kopfschmerzen (EVG/COBI/FTC/TDF: 16,1%; EFV/FTC/TDF: 11,4%; OR [95%-KI]: 1,50 [0,97;2,31], $p=0,0702$; RR [95%-KI]: 1,42 [0,97;2,07], $p=0,0710$). Im EVG/COBI/FTC/TDF-Arm wurde statistisch signifikant weniger Schwindel beobachtet als im EFV/FTC/TDF-Arm (OR [95%-KI]: 0,22 [0,14;0,36], $p<0,00001$; RR [95%-KI]: 0,28 [0,18;0,42], $p<0,00001$).

Die häufigsten *psychiatrischen Erkrankungen (SOC)* waren abnorme Träume (EVG/COBI/FTC/TDF: 15,2%; EFV/FTC/TDF: 27,6%), Schlaflosigkeit (EVG/COBI/FTC/TDF: 10,9%; EFV/FTC/TDF: 16,2%) und Depression (EVG/COBI/FTC/TDF: 12,1%; EFV/FTC/TDF: 13,9%; OR [95%-KI]: 0,85 [0,55;1,32], $p=0,4668$; RR [95%-KI]: 0,87 [0,59;1,27], $p=0,4670$). Hinsichtlich abnormer Träume (OR [95%-KI]: 0,47 [0,32;0,69], $p=0,0001$; RR [95%-KI]: 0,55 [0,41;0,75], $p=0,0001$) und Schlaflosigkeit (OR [95%-KI]: 0,63 [0,41;0,99], $p=0,0428$; RR [95%-KI]: 0,67 [0,46;0,99], $p=0,0436$) besteht dabei jeweils ein statistisch signifikanter Vorteil der EVG/COBI/FTC/TDF-Behandlung gegenüber der EFV/FTC/TDF-Behandlung.

Hautausschlag war das unter den *Erkrankungen der Haut und des Unterhautzellgewebes (SOC)* am häufigsten beobachtete Ereignis (EVG/COBI/FTC/TDF: 7,5%; EFV/FTC/TDF: 13,9%), welches statistisch signifikant seltener im EVG/COBI/FTC/TDF-Arm als im

EFV/FTC/TDF-Arm beobachtet wurde (OR [95%-KI]: 0,50 [0,30;0,82], $p=0,0066$; RR [95%-KI]: 0,54 [0,34;0,84], $p=0,0070$).

Studie zu RAL+FTC/TDF vs. EFV+FTC/TDF

Unter *Erkrankungen des Gastrointestinaltrakts (SOC)* traten am häufigsten Diarrhoe (RAL+FTC/TDF vs. EFV+FTC/TDF: 17,1% vs. 24,5%; OR [95%-KI]: 0,64 [0,42;0,96], $p=0,0315$; RR [95%-KI]: 0,70 [0,50;0,97], $p=0,0324$) und Übelkeit auf (RAL+FTC/TDF vs. EFV+FTC/TDF: 14,2% vs. 12,8%; OR [95%-KI]: 1,13 [0,70;1,84], $p=0,6102$; RR [95%-KI]: 1,12 [0,73;1,70], $p=0,6103$). Der Unterschied zwischen den Behandlungen hinsichtlich des Auftretens von Diarrhoe zeigte statistische Signifikanz zugunsten von RAL+FTC/TDF.

Unter *allgemeinen Erkrankungen und Beschwerden am Verabreichungsort (SOC)* besteht hinsichtlich Fatigue ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF (RAL+FTC/TDF vs. EFV+FTC/TDF: 6,8% vs. 11,7%; OR [95%-KI]: 0,55 [0,30;0,99], $p=0,0454$; RR [95%-KI]: 0,58 [0,34;0,99], $p=0,0464$).

Unter *Erkrankungen des Nervensystems (SOC)* traten Schwindel (RAL+FTC/TDF vs. EFV+FTC/TDF: 8,2% vs. 36,9%; OR [95%-KI]: 0,15 [0,09;0,25], $p<0,00001$; RR [95%-KI]: 0,22 [0,15;0,34], $p<0,00001$) und Kopfschmerzen (RAL+FTC/TDF vs. EFV+FTC/TDF: 22,8% vs. 24,5%; OR [95%-KI]: 0,91 [0,62;1,34], $p=0,6365$; RR [95%-KI]: 0,93 [0,69;1,25], $p=0,6366$) am häufigsten auf. Dabei besteht hinsichtlich des Auftretens von Schwindel ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF.

Unter *psychiatrischen Erkrankungen (SOC)* besteht für abnorme Träume ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,53 [0,30;0,94], $p=0,0294$; RR [95%-KI]: 0,57 [0,34;0,95], $p=0,0303$). Dabei traten abnorme Träume bei 7,5% der RAL+FTC/TDF-Patienten und 13,1% der EFV+FTC/TDF-Patienten auf. Hinsichtlich Schlaflosigkeit besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (RAL+FTC/TDF vs. EFV+FTC/TDF: 12,1% vs. 11,0%; OR [95%-KI]: 1,11 [0,66;1,87], $p=0,6813$; RR [95%-KI]: 1,10 [0,70;1,74], $p=0,6813$). Auch hinsichtlich dem Auftreten von Depressionen kann kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (RAL+FTC/TDF vs. EFV+FTC/TDF: 7,5% vs. 8,9%; OR [95%-KI]: 0,83 [0,45;1,52], $p=0,5470$; RR [95%-KI]: 0,84 [0,48;1,47], $p=0,5472$).

Bei *Erkrankungen der Haut und des Unterhautzellgewebes (SOC)* trat Hautausschlag unter einer Behandlung mit RAL+FTC/TDF (6,0%) statistisch signifikant seltener auf als unter einer Behandlung mit EFV+FTC/TDF (12,1%) (OR [95%-KI]: 0,47 [0,26;0,86], $p=0,0148$; RR [95%-KI]: 0,50 [0,29;0,88], $p=0,0155$).

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.*

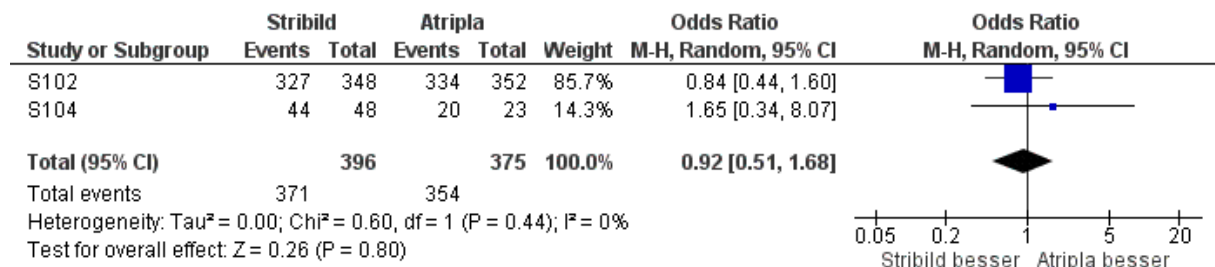


Abbildung 18: Meta-Analyse für die Gesamtrate unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zur **Gesamtrate unerwünschter Ereignisse** zeigt - bei homogener Datenlage ($p=0,44$; $I^2=0\%$) - keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen. Das OR [95%-KI] beträgt 0,92 [0,51;1,68] mit $p=0,80$ (RR [95%-KI]: 0,99 [0,96;1,03], $p=0,68$). Eine Sensitivitätsanalyse, unter Verwendung eines FEM, zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,92 [0,51;1,68], $p=0,79$).

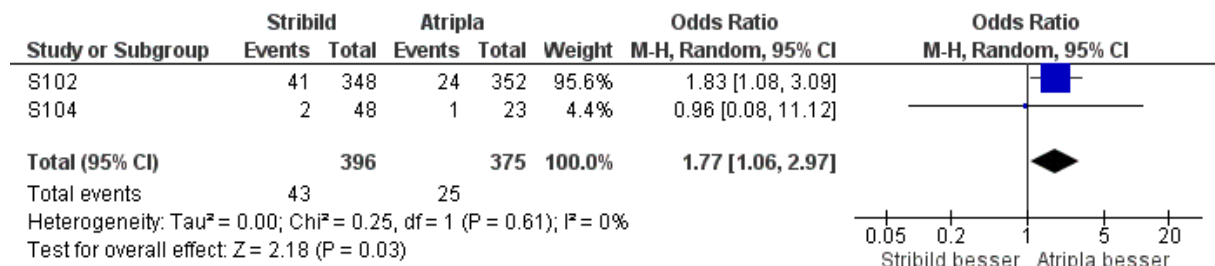


Abbildung 19: Meta-Analyse für die Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse der **schwerwiegenden unerwünschten Ereignisse** zeigt keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,77 [1,06;2,97], $p=0,03$; RR [95%-KI]: 1,69 [1,05;2,70], $p=0,03$). Dabei ist die Datenlage homogen ($p=0,61$; $I^2=0\%$) und eine Sensitivitätsanalyse, unter Verwendung des FEM, zeigt robuste Ergebnisse (OR [95%-KI]: 1,77 [1,06;2,97], $p=0,03$).

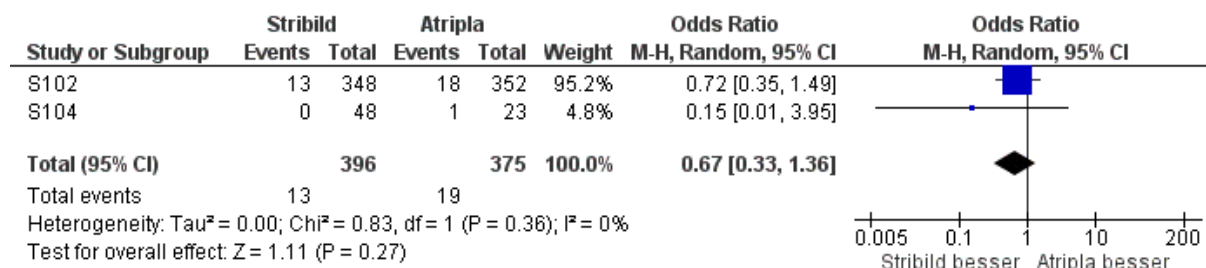


Abbildung 20: Meta-Analyse für Therapieabbrüche aufgrund unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zu **Therapieabbrüchen aufgrund unerwünschter Ereignisse** zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,67 [0,33;1,36], p=0,27; RR [95%-KI]: 0,68 [0,34;1,35], p=0,27). Es besteht keine bedeutsame Heterogenität (p=0,36; I²=0%) zwischen den Studien, weshalb die Sensitivitätsanalyse, unter Verwendung eines FEM, robuste Ergebnisse zeigt (OR [95%-KI]: 0,66 [0,33;1,34], p=0,25).

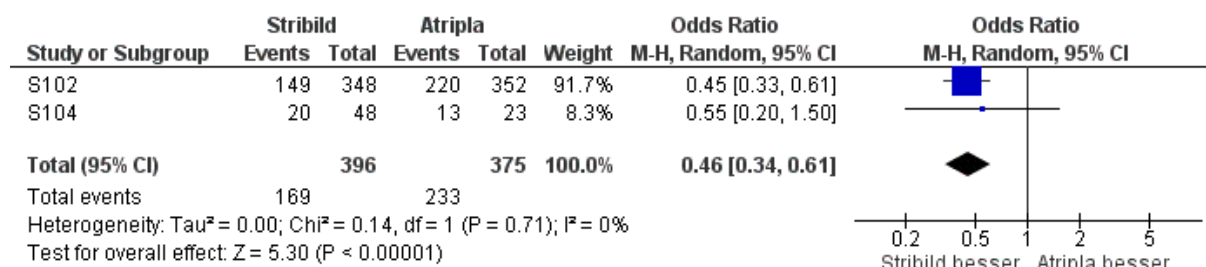


Abbildung 21: Meta-Analyse für neurologische & psychiatrische Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Hinsichtlich **neurologischer und psychiatrischer Ereignisse** besteht in der Meta-Analyse zu Woche 48 ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,46 [0,34;0,61], p<0,00001; RR [95%-KI]: 0,69 [0,60;0,79], p<0,00001). Neurologische & psychiatrische Ereignisse traten unter EVG/COBI/FTC/TDF somit statistisch signifikant seltener auf als unter EFV/FTC/TDF. Die Studien zeigen keine bedeutsame Heterogenität untereinander (p=0,71; I²=0%) und eine Sensitivitätsanalyse, unter Verwendung eines FEM, zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,46 [0,34;0,61], p<0,00001).

In der folgenden Tabelle werden die PT, die unter neurologischen & psychiatrischen (Studie S102 und S104) bzw. neuropsychiatrischen Ereignissen (STARTMRK-Studie) betrachtet werden, dargestellt. In der mittleren Spalte sind die PT dargestellt, die in beiden direkten Vergleichen unter der jeweiligen Endpunktdefinition eingeschlossen wurden.

Tabelle 4-161: Definition und Überschneidungen der neurologischen & psychiatrischen Ereignisse bzw. neuropsychiatrischen Ereignisse

PTs, die in die EVG/COBI/FTC/TDF-Studien eingeschlossen wurden (unter neurologische & psychiatrische Ereignisse)	PTs, die in die EVG/COBI/FTC/TDF-Studien und die RAL+FTC/TDF-Studie eingeschlossen wurden (Überschneidungen)	PTs, die in die RAL+FTC/TDF-Studie eingeschlossen wurden (unter neuropsychiatrischen Ereignissen)
Depression	Depression	Depression
Alpträume	Alpträume	Alpträume
Verwirrheitszustände	Verwirrheitszustände	Verwirrheitszustände
Suizidgedanken	Suizidgedanken	Suizidgedanken
abnorme Träume	abnorme Träume	abnorme Träume
Suizidversuch	Suizidversuch	Suizidversuch
akute Psychose	akute Psychose	akute Psychose
Halluzinationen	Halluzinationen	Halluzinationen
vollendeter Suizid	vollendeter Suizid	vollendeter Suizid
starke Depressionen	starke Depressionen	starke Depressionen
Schwindel	Schwindel	Schwindel
Schlaflosigkeit	Schlaflosigkeit	Schlaflosigkeit
Somnolenz	Somnolenz	Somnolenz
affektive Erkrankung		Nervensystemstörungen
Agitation		psychotische Störung
Amnesie		Delirium
Angst		eingeschränktes Bewusstsein
Angststörung		akustische Halluzination
bipolare Störung		Konzentrationsstörungen
depressive Stimmung		
Desorientierung		
Aufmerksamkeitsstörungen		
euphorische Stimmung		
Kopfschmerzen		
Reizbarkeit		
verringerte Libido		
veränderte Stimmung		
Stimmungsschwankungen		
Nervosität		
Panikattacke		
Paranoia		

PTs, die in die EVG/COBI/FTC/TDF-Studien eingeschlossen wurden (unter neurologische & psychiatrische Ereignisse)	PTs, die in die EVG/COBI/FTC/TDF-Studien und die RAL+FTC/TDF-Studie eingeschlossen wurden (Überschneidungen)	PTs, die in die RAL+FTC/TDF-Studie eingeschlossen wurden (unter neuropsychiatrischen Ereignissen)
Unruhe		
Schlafstörungen		
Gleichgewichtsstörung		
Grau hinterlegte PTs wurden in der STARTMRK-Studie nach der Entblindung als Sensitivitätsanalyse hinzugefügt. Fett gedruckte PTs in der linken Spalte wurden in beiden EVG/COBI/FTC/TDF-Studien eingeschlossen, normal gedruckte nur in Studie S102.		

Grau hinterlegte PT sind die PT, die in der STARTMRK-Studie nach Entblindung zu dem Endpunkt neuropsychiatrische Ereignisse hinzugefügt wurden, um Sensitivitätsanalysen durchzuführen. Fett gedruckte PT in der linken Spalte wurden in beiden EVG/COBI/FTC/TDF-Studien eingeschlossen; normal gedruckte wurden nur in Studie S102 unter neurologischen & psychiatrischen Ereignissen eingeschlossen.

Die Operationalisierung der neurologischen & psychiatrischen bzw. neuropsychiatrischen Ereignisse unterschied sich zwischen den Studien insofern, dass in STARTMRK deutlich weniger, und auch andere PT in diese Gruppe von PT eingeschlossen wurden. Diese Inkonsistenz zwischen den Endpunktdefinitionen ist zu hoch, um einen indirekten Vergleich zu diesem Endpunkt als adäquat einzustufen. Auch unter Einschluss der PT, die nach der Entblindung hinzugefügt wurden, ist die Inkonsistenz zwischen den beiden Gruppen von PT zu hoch, als dass ein indirekter Vergleich sinnvoll wäre. Diese Daten werden hier somit nur dargestellt.

Zu Woche 96 gibt es in STARTMRK keine Angaben darüber, ob sich die angegebenen Ereignisraten auf die präspezifizierte Gruppe von PT oder die im Nachhinein erweiterte Gruppe von PT bezieht.

Da keine direkte und indirekte Evidenz vorlag, werden keine Konsistenzprüfungen durchgeführt.

Im Weiteren werden die Ergebnisse der indirekten Vergleiche zu den weiteren Endpunkten dargestellt.

Tabelle 4-162: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 48)

Unerwünschte Ereignisse (Woche 48)	OR [95%-KI]			
	Mortalität	Gesamt UEs	Gesamt SUEs	Abbrüche aufgrund UEs
Direkter Vergleich EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	0,50 [0,05;5,59]	0,92 [0,51;1,68]	1,77 [1,06;2,97]	0,67 [0,33;1,36]
Direkter Vergleich RAL+FTC/TDF vs. EFV+FTC/TDF	5,05 [0,24;105,74]	0,33 [0,16;0,70]	1,05 [0,60;1,82]	0,52 [0,23;1,18]
Indirekter Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	0,10 [0,00;4,66]	2,79 [1,08;7,20]	1,69 [0,79;3,59]	1,29 [0,44;3,80]
KI = Konfidenzintervall, OR = Odds Ratio				

Der indirekte Vergleich zur Mortalität zeigt, dass zwischen den Behandlungen EVG/COBI/FTC/TDF und RAL+FTC/TDF kein statistisch signifikanter Unterschied besteht (OR [95%-KI]: 0,10 [0,00;4,66]; RR [95%-KI]: 0,10 [0,00;4,74]). Bei der Gesamtrate unerwünschter Ereignisse besteht durch den indirekten Vergleich ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 2,79 [1,08;7,20]). Bei Verwendung eines anderen Effektmaßes, wie dem relativen Risiko, besteht der statistisch signifikante Unterschied zwischen den beiden Behandlungen nur noch knapp (RR [95%-KI]: 1,06 [1,003;1,13]). Hinsichtlich der schwerwiegenden unerwünschten Ereignisse besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,69 [0,79;3,59]; RR [95%-KI]: 1,63 [0,82;3,24]). Und auch für die Therapieabbrüche aufgrund unerwünschter Ereignisse kann kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (OR [95%-KI]: 1,29 [0,44;3,80]; RR [95%-KI]: 1,28 [0,45;3,67]).

Tabelle 4-163: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 96)

Unerwünschte Ereignisse (Woche 96)	OR [95%-KI]			
	Mortalität	Gesamt UEs	Gesamt SUEs	Abbrüche aufgrund UEs
Direkter Vergleich EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	0,50 [0,05;5,59]	0,90 [0,38;2,14]	1,85 [1,17;2,93]	0,70 [0,37;1,33]
Direkter Vergleich RAL+FTC/TDF vs. EFV+FTC/TDF	7,10 [0,37;138,10]	0,45 [0,18;1,12]	1,21 [0,74;1,98]	0,64 [0,29;1,38]
Indirekter Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	0,07 [0,00;3,10]	2,00 [0,57;7,04]	1,53 [0,78;3,00]	1,09 [0,40;3,00]

Unerwünschte Ereignisse (Woche 96)	OR [95%-KI]			
	Mortalität	Gesamt UEs	Gesamt SUEs	Abbrüche aufgrund UEs
KI = Konfidenzintervall, OR = Odds Ratio				

Der indirekte Vergleich zur Mortalität zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF. Das OR [95%-KI] beträgt 0,07 [0,00;3,10] (RR [95%-KI]: 0,07 [0,00;3,20]). Auch hinsichtlich der Gesamtrate unerwünschter Ereignisse kann, im Gegensatz zu Woche 48, kein statistisch signifikanter Unterschied zwischen den Behandlungen mehr beobachtet werden (OR [95%-KI]: 2,00 [0,57;7,04]; RR [95%-KI]: 1,03 [0,99;1,07]). Die Gesamtrate schwerwiegender unerwünschter Ereignisse trat ebenso in beiden Behandlungen (EVG/COBI/FTC/TDF vs. RAL+FTC/TDF) statistisch nicht signifikant unterschiedlich häufig auf (OR [95%-KI]: 1,53 [0,78;3,00]; RR [95%-KI]: 1,46 [0,81;2,62]). Auch der indirekte Vergleich zu Therapieabbrüchen aufgrund unerwünschter Ereignisse zeigt keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,09 [0,40;3,00]; RR [95%-KI]: 1,11 [0,43;2,88]).

Tabelle 4-164: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 96, Fortsetzung I)

Unerwünschte Ereignisse (Woche 96)	OR [95%-KI]				
	Diarrhoe	Übelkeit	Fatigue	Schwindel	Kopfschmerzen
Direkter Vergleich EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	1,08 [0,76;1,53]	1,58 [1,07;2,32]	0,86 [0,56;1,32]	0,22 [0,14;0,36]	1,50 [0,97;2,31]
Direkter Vergleich RAL+FTC/TDF vs. EFV+FTC/TDF	0,64 [0,42;0,96]	1,13 [0,70;1,84]	0,55 [0,30;0,99]	0,15 [0,09;0,25]	0,91 [0,62;1,34]
Indirekter Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	1,69 [0,98;2,90]	1,40 [0,75;2,60]	1,56 [0,75;3,26]	1,47 [0,73;2,94]	1,65 [0,92;2,94]
KI = Konfidenzintervall, OR = Odds Ratio					

Der indirekte Vergleich zu Woche 96 zu Diarrhoe zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF und RAL+FTC/TDF (OR [95%-KI]: 1,69 [0,98;2,90]; RR [95%-KI]: 1,51 [0,99;2,31]). Auch hinsichtlich Übelkeit kann kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (OR [95%-KI]: 1,40 [0,75;2,60]; RR [95%-KI]: 1,29 [0,76;2,19]). Hinsichtlich dem Auftreten von Fatigue besteht zwischen den beiden Behandlungen kein statistisch signifikanter Unterschied (OR [95%-KI]: 1,56 [0,75;3,26]; RR [95%-KI]: 1,52 [0,79;2,91]). Das OR [95%-KI] zu Schwindel beträgt 1,47 [0,73;2,94] (RR [95%-KI]: 1,27 [0,71;2,29]), was keinen statistisch

signifikanten Unterschied zwischen den Behandlungen anzeigt. Und auch zu Kopfschmerzen besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,65 [0,92;2,94]; RR [95%-KI]: 1,53 [0,94;2,47]).

Tabelle 4-165: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Woche 96, Fortsetzung II)

Unerwünschte Ereignisse (Woche 96)	OR [95%-KI]			
	Abnorme Träume	Schlaflosigkeit	Depression	Hautausschlag
Direkter Vergleich EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	0,47 [0,32;0,69]	0,63 [0,41;0,99]	0,85 [0,55;1,32]	0,50 [0,30;0,82]
Direkter Vergleich RAL+FTC/TDF vs. EFV+FTC/TDF	0,53 [0,30;0,94]	1,11 [0,66;1,87]	0,83 [0,45;1,52]	0,47 [0,26;0,86]
Indirekter Vergleich EVG/COBI/FTC/TDF vs. RAL+FTC/TDF	0,89 [0,45;1,76]	0,57 [0,29;1,12]	1,02 [0,48;2,17]	1,06 [0,49;2,32]
KI = Konfidenzintervall, OR = Odds Ratio				

Hinsichtlich abnormer Träume ist der Unterschied zwischen den Behandlungen statistisch nicht signifikant (OR [95%-KI]: 0,89 [0,45;1,76]; RR [95%-KI]: 0,96 [0,53;1,75]). Und auch hinsichtlich Schlaflosigkeit kann kein statistisch signifikanter Unterschied zwischen den Behandlungen festgestellt werden (OR [95%-KI]: 0,57 [0,29;1,12]; RR [95%-KI]: 0,61 [0,34;1,10]). Außerdem besteht zu Woche 96 weder hinsichtlich von Depressionen (OR [95%-KI]: 1,02 [0,48;2,17]; RR [95%-KI]: 1,04 [0,53;2,04]) noch von Hautausschlag (OR [95%-KI]: 1,06 [0,49;2,32]; RR [95%-KI]: 1,08 [0,53;2,21]) ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen EVG/COBI/FTC/TDF und RAL+FTC/TDF.

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Da in den EVG/COBI/FTC/TDF-Studien nur Subgruppenanalysen für den primären Endpunkt (Virologisches Ansprechen mittels Snapshot-Algorithmus) durchgeführt wurden, und nicht für andere Analyse-Algorithmen des virologischen Ansprechens, können hier keine indirekten Vergleiche für Subgruppen durchgeführt werden, da für die STARTMRK keine vergleichbaren Subgruppenergebnisse veröffentlicht wurden.

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 4-166: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Andere Aspekte, die Verzerrungen verursachen können
			Patient	Behandler		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-167: 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 4-168: 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 4.3.1.3.2.

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*

In den folgenden Abschnitten werden unterstützende Informationen aus einer nicht-randomisierten Studie dargestellt. Diese Studie ist gemäß der in Abschnitt 4.2.2 angegebenen Kriterien nicht geeignet, die Fragestellung gemäß 4.2.1 zu beantworten. Sie wird hier dennoch aufgeführt, um die Evidenz der Übertragbarkeit der Daten zu EVG/COBI/FTC/TDF auf vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, zu unterstützen. Ihre Resultate werden allerdings nicht für die Bewertung des medizinischen Nutzens oder die Quantifizierung des Zusatznutzens berücksichtigt.

Die Studie wurde in der Registersuche zu EVG/COBI/FTC/TDF als weitere Studie im Indikationsgebiet identifiziert. Es handelt sich um eine Studie des pharmazeutischen Unternehmers.

Nachfolgend werden als weitere Unterlagen Daten aus der oben erwähnten nicht-randomisierten Studie im betrachteten Indikationsgebiet mit dem zu bewertenden Arzneimittel dargestellt und bewertet.

4.3.2.3.1.1 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.

Es werden nur Studien des pharmazeutischen Unternehmers mit dem zu bewertenden Arzneimittel berücksichtigt, die entweder mit dem Single-Tablet-Regimen (EVG/COBI/FTC/TDF) oder aber allen vier Komponenten in Kombination (EVG+COBI+FTC+TDF) in einem Studienarm durchgeführt wurden [59]. Studien, die nur mit einzelnen, aber nicht allen vier Komponenten in Kombination durchgeführt wurden, werden hier nicht gelistet, unabhängig davon, ob sie mit der Marketing Authorisation eingereicht wurden. Zudem werden in der folgenden Tabelle nur die Studien im entsprechenden Anwendungsgebiet aufgenommen (Infektion mit dem Humanen Immundefizienzvirus 1 (HIV-1) bei Erwachsenen im Alter von 18 Jahren und darüber, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind).

Tabelle 4-169: Liste der Studien des pharmazeutischen Unternehmers – weitere Untersuchungen

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
GS-US-236-0123	nein	laufend	48 Wochen (ggf. Erweiterung um 48 Wochen für Patienten in deren Land nach 48 Wochen Behandlung noch nicht kommerziell verfügbar ist oder die kein bestätigtes virologisches Ansprechen hatten)	Switch zu Elvitegravir 150 mg/ Cobicistat 150 mg/ Emtricitabin 200 mg/ Tenofoviridisoproxil 245 mg

Geben Sie an, welchen Stand die Information in Tabelle 4-3 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.

Tabelle 4-169 gibt einen Überblick über die Studie des pharmazeutischen Unternehmers im beschriebenen Anwendungsgebiet, die nicht bereits in Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt wird. Die Studie ist eine nichtrandomisierte, nichtkontrollierte Studie mit dem zu bewertenden Arzneimittel in dem in diesem Dossier beschriebenen Anwendungsgebiet.

Der Stand der Information in Tabelle 4-169 ist der 30. April 2013.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-3 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-170: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – weitere Untersuchungen

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
GS-US-236-0123	Dies ist eine nicht-randomisierte Studie, für die bisher lediglich Daten zu Woche 12 veröffentlicht wurden.*
* Da diese Studie nicht zur Nutzenbewertung, wohl aber zur Unterstützung der Übertragbarkeit der Daten aus Modul 4A herangezogen wird, wird sie in diese Tabelle aufgenommen, gleichzeitig werden allerdings Daten zu dieser Studie dargestellt.	

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

4.3.2.3.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

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

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

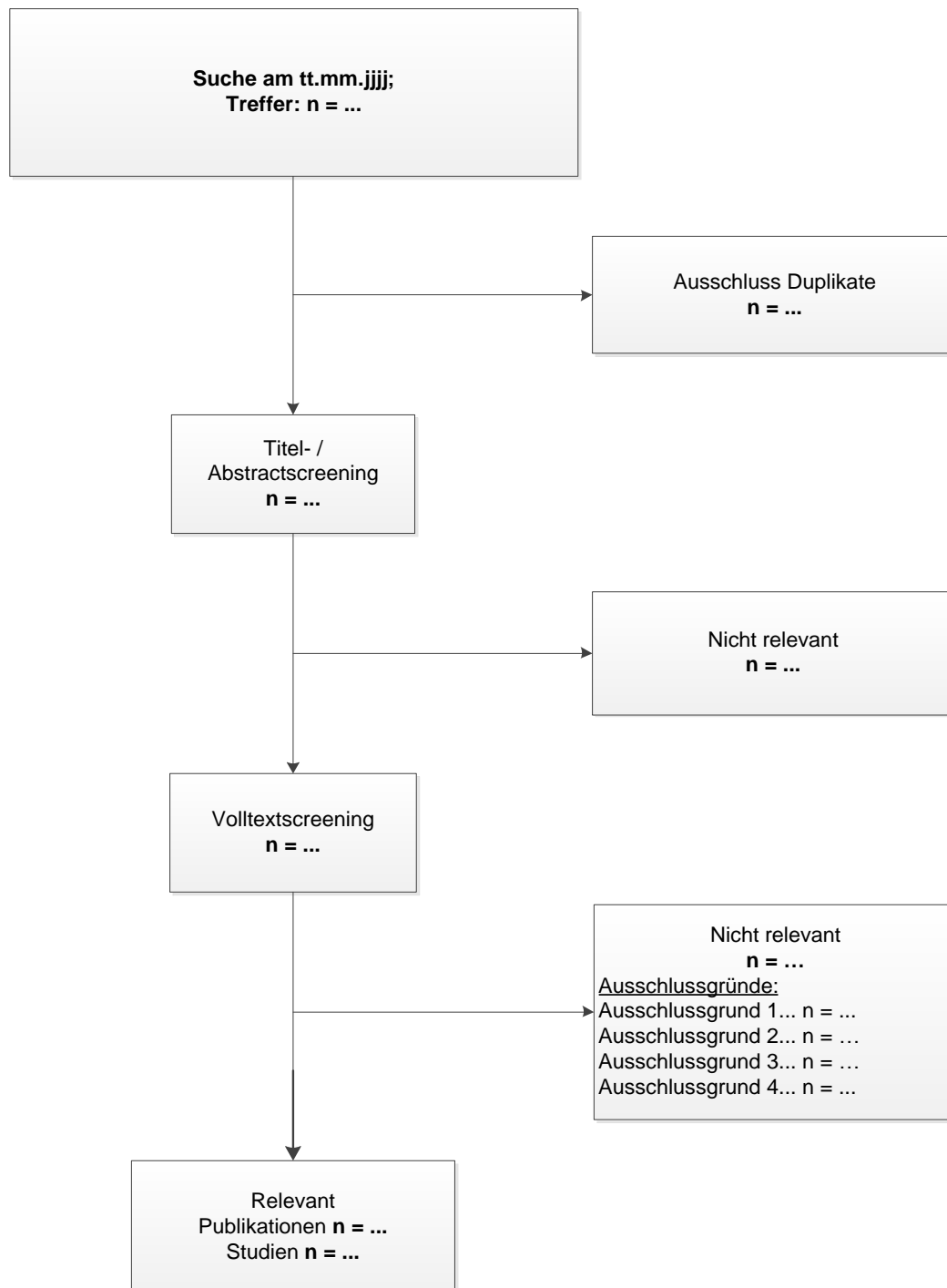


Abbildung 22: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die in Abschnitt 4.3.2.3 dargestellten Ergebnisse dienen nur der Unterstützung der Fragestellung gemäß 4.2.1 und werden für die eigentliche Beantwortung der in 4.2.1 spezifizierten Fragestellung und somit zur Quantifizierung des Zusatznutzens von EVG/COBI/FTC/TDF gegenüber den zweckmäßigen Vergleichstherapien EFV/FTC/TDF,

ATV/r+FTC/TDF und RAL+FTC/TDF nicht herangezogen. Aufgrund dessen wurde keine bibliographische Literaturrecherche durchgeführt.

4.3.2.3.1.3 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 4-3) 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 4-171: Relevante Studien aus der Suche in Studienregistern – weitere Untersuchungen

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)
<Studie 1>	clinicaltrials.gov [Zitate], clinicalstudyresults.org [Zitate]	ja	nein

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

4.3.2.3.1.4 Resultierender Studienpool: weitere Untersuchungen

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 4-172: Studienpool – weitere Untersuchungen

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						
GS-US-236-0123 ^d	nein	ja	nein	nein (Studienprotokoll [122])	ja [123, 124]	ja [125, 126]
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.</p> <p>b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>d: Die in den Studienpool aufgenommene Studie wird nicht zur Nutzenableitung herangezogen; sie dient lediglich zur Unterstützung der Übertragbarkeit der Daten aus Modul 4A.</p>						

Zu Studie GS-US-236-0123 liegt noch kein Studienbericht vor. Es wird daher im Weiteren zur Beschreibung der Studiencharakteristika, des Verzerrungspotenzials und der Ergebnisse auf Informationen aus dem Studienprotokoll, erste Veröffentlichungen und erste Daten zu Woche 12 zu dieser Studie zurückgegriffen [122, 125-127]. Hier sei nochmal betont, dass die Ergebnisse dieser Studie nicht zur Nutzenbewertung herangezogen werden, sondern lediglich die Übertragbarkeit der Daten der nicht-vorbehandelten Patienten auf vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, unterstützen sollen.

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.

Zu weiteren Unterlagen wurde auf der Basis der Studienregistersuche eine nicht-randomisierte, nicht-kontrollierte Studie identifiziert: GS-US-236-0123.

4.3.2.3.2.1 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 4-173: Charakterisierung der eingeschlossenen Studien – weitere Untersuchungen

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
GS-US-236-0123	Einarmig, Switch-Studie, offen	Virussupprimierte HIV-1-infizierte Erwachsene	EVG/COBI/FTC/TDF (n=48)	Run-in: 21 Tage Behandlung: 48 Wochen (ggf. Erweiterung um 48 Wochen für Patienten in deren Land EVG/COBI/FTC/TDF nach 48 Wochen Behandlung noch nicht kommerziell verfügbar ist und die kein bestätigtes virologisches Versagen hatten)	USA 2/2012-10/2013	<u>Primärer Endpunkt:</u> Virologisches Ansprechen zu Woche 12 <u>Sekundäre Endpunkte:</u> Virologisches Ansprechen zu Woche 24 und 48 Unerwünschte Ereignisse

Tabelle 4-174: Charakterisierung der Interventionen – weitere Untersuchungen

Studie	EVG/COBI/FTC/TDF	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in- Phase etc.</i>
GS-US-236-0123	150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin, 245 mg Tenofoviridisoproxil einmal täglich (EVG/COBI/FTC/TDF)	Voraussetzung: ≥6 Monate Vorbehandlung mit RAL+FTC/TDF Screening: 21 Tage Behandlung: 48 Wochen (ggf. Erweiterung um 48 Wochen für Patienten in deren Land EVG/COBI/FTC/TDF nach 48 Wochen Behandlung noch nicht kommerziell verfügbar ist und die kein bestätigtes virologisches Versagen hatten) Follow-Up: 30 Tage

Tabelle 4-175: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD)	Geschlecht w/m (%)	Baseline BMI (kg/m²) Mittelwert (SD)	Dauer der Vorbehandlung vor dem Switch (Jahre) Median
GS-US-236-0123 EVG/COBI/FTC/TDF	48	44 (8,6)	4,2/95,8	27,7 (4,16)	2,5
BMI = Body Mass Index, SD = Standardabweichung					

Tabelle 4-176: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)

Studie Gruppe	Rasse n (%)					
	Kaukasisch	Asiatisch	Schwarz oder afro- amerikanisch	Indianer/ Ureinwohner Alaskas	Hawaiianer/ Pazifik- insulaner	Andere
GS-US-236-0123 EVG/COBI/FTC/TDF	40 (83,3)	1 (2,1)	7 (14,6)	0 (0,0)	0 (0,0)	0 (0,0)
n: Anzahl Patienten in der Kategorie						

Tabelle 4-177: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)

Studie Gruppe	Ethnische Abstammung n (%)	
	Hispanisch/Latino	Nicht hispanisch/Latino
GS-US-236-0123		

Studie Gruppe	Ethnische Abstammung n (%)	
	Hispanisch/Latino	Nicht hispanisch/Latino
EVG/COBI/FTC/TDF	10 (20,8)	38 (79,2)
n: Anzahl der Patienten in der Kategorie		

Tabelle 4-178: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)

Studie Gruppe	Screening HIV-1-RNA Kopien/mL n (%)		Baseline HIV-1-RNA Kopien/mL n (%)	
	<50 Kopien/mL	≥50 Kopien/mL	<50 Kopien/mL	≥50 Kopien/mL
	GS-US-236-0123 EVG/COBI/FTC/TDF	48 (100,0)	0 (0,0)	46 (95,8)
n: Anzahl der Patienten in der Kategorie				

Tabelle 4-179: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)

Studie Gruppe	CD4-Zellzahl zu Baseline (/ μ L) Mittelwert (SD)	CD4-Zellzahl zu Baseline (/ μ L) n (%)				
		≤50	51-≤200	201-≤350	351-≤500	>500
GS-US-236-0123 EVG/COBI/FTC/TDF	711 (265,9)	0 (0,0)	1 (2,1)	4 (8,3)	6 (12,5)	37 (77,1)
n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung						

Tabelle 4-180: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)

Studie Gruppe	HIV Krankheitsstatus n (%)			eGFR nach Cockcroft-Gault (mL/min) Mittelwert (SD)
	Asymptomatisch	Symptomatische HIV Infektionen	Aids	
GS-US-236-0123 EVG/COBI/FTC/TDF	45 (93,8)	1 (2,1)	2 (4,2)	110,4 (21,75)
Aids = Erworbenes Immundefektsyndrom (Acquired immunodeficiency syndrome), eGFR = Geschätzte glomeruläre Filtrationsrate (estimated Glomerular Filtration Rate), HIV = Humanes Immundefizienz-Virus, n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung				

Tabelle 4-181: Charakterisierung der Studienpopulationen – weitere Untersuchungen (Safety Analyse-Set)

Studie Gruppe	Adhärenz-Rate n (%)				Adhärenz-Rate Mittelwert (SD)
	≥70%	≥80%	≥90%	≥95%	
Woche 12					
GS-US-236-0123					
EVG/COBI/FTC/TDF	48 (100,0)	48 (100,0)	47 (97,9)	45 (93,8)	99 (2,3)
Adhärenz-Rate war wie folgt definiert: 100 x (Gesamtanzahl eingenommener Tabletten / Gesamtanzahl verschriebener Tabletten)					
n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung					

Tabelle 4-182: Charakterisierung der Studienpopulationen (Patientendisposition) – weitere Untersuchungen

Studie Disposition	EVG/COBI/FTC/TDF n (%)
S102	
Randomisiert	48
Behandlung erhalten	48 (100,0)
Woche 12 in Behandlung beendet	48 (100,0)
Woche 12 in Studie beendet	48 (100,0)
n: Anzahl der Patienten in der Kategorie	

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

Design

Studie GS-US-236-0123 war eine einarmige, offene Switch-Studie, in der Patienten, die zuvor mindestens sechs Monate virologisch stabil mit einer Initialtherapie bestehend aus RAL+FTC/TDF behandelt worden sind, und deren Viruslast für mindestens sechs Monate vor Screening unter der Nachweisgrenze lag (gemäß des lokal verwendeten HIV-Test; mindestens zweimal mit dem gleichen Test bestätigt) eingeschlossen wurden. Außerdem durften die Patienten nach dem Erreichen des virologischen Ansprechens keine zwei bestätigten HIV-1 RNA-Level über der Nachweisgrenze aufweisen.

Nach einer maximal 21-tägigen Screening-Phase wurden alle Patienten von ihrer bisherigen Therapie mit RAL+FTC/TDF auf EVG/COBI/FTC/TDF umgestellt und wurden für 48 Wochen mit EVG/COBI/FTC/TDF behandelt. Bei kommerzieller Verfügbarkeit traten die Patienten anschließend in eine 30-tägige Follow-Up-Phase ein (vor Ort oder per Telefon). War EVG/COBI/FTC/TDF noch nicht kommerziell verfügbar, hatten die Patienten die Möglichkeit entweder nach 30 Tagen eine Follow-Up-Untersuchung zu absolvieren oder EVG/COBI/FTC/TDF weiterhin für maximal 48 Wochen zu erhalten, bis EVG/COBI/FTC/TDF kommerziell verfügbar ist oder Gilead das klinische Programm zu

diesem Produkt stoppt. Es konnten nur Patienten ohne virologisches Versagen in die Extensionsphase eintreten.

Die Patienten erhielten 150 mg Elvitegravir, 150 mg Cobicistat, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil als STR. Dies sollte einmal täglich oral mit einer Mahlzeit um ungefähr die gleiche Tageszeit eingenommen werden.

Die Studie wurde in 20 Zentren in den USA durchgeführt.

Studienpopulation

Es wurden insgesamt 48 Patienten in die Studien eingeschlossen. Alle Patienten beendeten 12 Wochen der Behandlung in Behandlung (siehe Tabelle 4-182). Im Durchschnitt waren die Patienten 44 Jahre alt; 95,8% der Patienten waren männlich, während lediglich 4,2% der Patienten weiblich waren. Der Baseline-BMI lag bei 27,7 kg/m² und der Median der Erkrankungsdauer lag bei 2,5 Jahren (siehe Tabelle 4-175).

14,6% der eingeschlossenen Patienten waren schwarz oder afroamerikanisch, 2,1% der Patienten waren asiatisch und 83,3% der Patienten in der Studie waren kaukasisch (siehe Tabelle 4-176); außerdem waren 20,8% der Patienten hispanisch bzw. Latino (siehe Tabelle 4-177).

Zum Screening wiesen alle 48 (100,0%) Patienten eine HIV-1-RNA Viruslast <50 Kopien/mL auf. Zu Baseline lag die HIV-1 RNA Viruslast bei zwei (4,2%) Patienten über der Nachweisgrenze von 50 Kopien/mL (siehe Tabelle 4-178). Die CD4-Zellzahl zu Baseline lag bei 711 Zellen/μL; lediglich ein Patient (2,1%) wies zu Baseline eine CD4-Zellzahl zwischen 51 und ≤200 Zellen/μL auf. Der Großteil der Patienten (37 [77,1%]) wies zu Baseline eine CD4-Zellzahl >500 Zellen/μL auf (siehe Tabelle 4-179).

93,8% der Patienten in der Studie waren zu Baseline asymptomatisch, ein Patient (2,1%) zeigte symptomatische HIV-1 Infektionen und zwei Patienten (4,2%) waren an Aids erkrankt.

Die eGFR nach Cockcroft-Gault lag im Mittel bei 110,4 mL/min (siehe Tabelle 4-180).

Alle Patienten wiesen eine Adhärenzrate ≥80% auf. ≥90% adhärenz waren 97,9% der Patienten und eine Adhärenz ≥95% wiesen 93,8% der Patienten auf. Die Adhärenzrate bis Woche 12 lag im Mittel bei 99% (siehe Tabelle 4-181).

Die Dauer der HIV-Infektion vor Switch lag im Median bei 2,5 Jahren

Einschlusskriterien waren unter anderem eine HIV-1 RNA Viruslast <50 Kopien/mL zum Screening, eine adäquate Nierenfunktion, definiert durch eine eGFR ≥90 mL/min gemäß Cockcroft-Gault Formel. Vor Beginn der Initialtherapie musste ein Resistenztest durchgeführt worden sein, es durften keine Resistenzen gegen eine der Studiensubstanzen (inklusive, jedoch nicht beschränkt auf K65R und M184V) vorliegen. Aus der Studie ausgeschlossen wurden unter anderem Patienten, für die innerhalb 21 Tagen vor Screening neue Aids-

definierende Erkrankungen diagnostiziert wurden, die schwanger waren oder die eine akute oder chronische Hepatitis B oder Hepatitis C Ko-Infektion aufwiesen.

Aufgrund falscher Angaben des Patienten wurde ein Patient, der nicht alle Einschlusskriterien erfüllte, in die Studie eingeschlossen. Der Patient erfüllte nicht das Einschlusskriterium nach dem die einzige Vorbehandlung der Patienten die Initialbehandlung sein musste, sondern wurde vor der Vorbehandlung mit RAL+FTC/TDF schon drei Monate mit EFV/FTC/TDF behandelt.

4.3.2.3.2 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 4-183: Verzerrungspotenzial auf Studienebene – weitere Untersuchungen

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
GS-US-236-0123	nein*	nein*	nein*	nein*	unklar	nein	hoch**
* nicht anwendbar, da die Studie ein einarmiges Design hatte							
** Gilt aufgrund des Studiendesigns generell als potenziell hoch verzerrt							

Begründen Sie für jede Studie die abschließende Einschätzung.

Da dies eine einarmige Switch-Studie ist, wurde keine Randomisierungssequenz erzeugt, es gab keine Verdeckung der Gruppenzuteilung und weder Patienten noch Behandler waren verblindet. Es kann nicht eingeschätzt werden ob es Hinweise auf ergebnisgesteuerte Berichterstattung gibt, da noch kein Studienbericht vorliegt. Sonstige das Verzerrungspotenzial beeinflussende Punkte wurden nicht gefunden. Da unkontrollierte Studien jedoch generell als potenziell hoch verzerrt gelten, wird das Verzerrungspotenzial für diese Studie als hoch eingestuft.

Es gilt zu beachten, dass alle Angaben auf dem Studienprotokoll, den ersten Veröffentlichungen und ersten Daten zu Woche 12 [122, 125-127] basieren.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

4.3.2.3.3.1 Virologisches Ansprechen – 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 4-184: Operationalisierung von virologisches Ansprechen

Studie	Operationalisierung
GS-US-236-0123	Anteil an Patienten mit Viruslast <50 Kopien/mL, zu Woche 12, 24 bzw. 48. Virologisches Ansprechen wurde mittels Snapshot-Algorithmus erfasst.

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.

Tabelle 4-185: Bewertung des Verzerrungspotenzials für virologisches Ansprechen in weitere Untersuchungen

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver- zerrungspotenzial be- einflussende Punkte
GS-US-236-0123	nein	nicht zutreffend, da unkontrollierte Studie	nein	nein

Der Endpunkt virologisches Ansprechen war definiert als Abfall der Viruslast unter die Nachweisgrenze von 50 Kopien/mL, erfasst mittels FDA Snapshot-Algorithmus. Der primäre Endpunkt war definiert als Anteil an Patienten, die nach 12 Wochen Behandlung virologisches Ansprechen erreicht haben. Sekundäre Endpunkte waren der Anteil der Patienten mit virologischem Ansprechen nach 24 bzw. 48 Wochen. Da jedoch bisher nur erste Daten nach 12 Wochen Behandlung veröffentlicht wurden, wird hier nur der Anteil an Patienten mit virologischem Ansprechen nach 12 Wochen Behandlung betrachtet.

Da dies eine einarmige Studie ist, gab es keine Verblindung der Endpunkterheber. Eine adäquate Umsetzung des ITT-Prinzips ist hier nicht zutreffend, da keine randomisierte Gruppenzuteilung stattfand. Es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung und auch sonstige das Verzerrungspotenzial beeinflussende Punkte fehlen. Da unkontrollierte Studien jedoch generell potenziell hoch verzerrt sind, wird das Verzerrungspotenzial für diesen Endpunkt als hoch bewertet.

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.

Nach 12 Wochen Behandlung mit EVG/COBI/FTC/TDF zeigten 48 Patienten von 48 (100%) virologisches Ansprechen gemäß FDA Snapshot-Algorithmus.

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 Veränderung der CD4-Zellzahl– 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 4-186: Operationalisierung von Veränderung der CD4-Zellzahl

Studie	Operationalisierung
GS-US-236-0123	Mittlere Veränderung der CD4-Zellzahl ab Baseline bis Woche 12, 24 bzw. 48

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.

Tabelle 4-187: Bewertung des Verzerrungspotenzials für Veränderung der CD4-Zellzahl in weitere Untersuchungen

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver- zerrungspotenzial be- einflussende Punkte
GS-US-236-0123	nein	nicht zutreffend, da unkontrollierte Studie	nein	nein

Der Endpunkt Veränderung der CD4-Zellzahl war definiert als mittlere Veränderung der CD4-Zellzahl ab Baseline bis Woche 12, 24 bzw. 48. Da bisher nur erste Daten nach 12 Wochen Behandlung veröffentlicht wurden, wird hier nur die mittlere Veränderung der CD4-Zellzahl nach 12 Wochen Behandlung betrachtet.

Da dies eine einarmige Studie ist, gab es keine Verblindung der Endpunkterheber. Eine adäquate Umsetzung des ITT-Prinzips ist hier nicht zutreffend, da keine randomisierte Gruppenzuteilung stattfand. Es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung und auch sonstige das Verzerrungspotenzial beeinflussende Punkte fehlen. Da unkontrollierte Studien jedoch generell potenziell hoch verzerrt sind, wird das Verzerrungspotenzial für diesen Endpunkt als hoch bewertet.

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.

Zu Baseline wiesen die Patienten im Mittel eine CD4-Zellzahl von 711 Zellen/ μ L auf mit einer Standardabweichung von 265,9. Bis Woche 12 stieg die CD4-Zellzahl um durchschnittlich 33 Zellen/ μ L (SD: 171,9) auf 744 Zellen/ μ L (SD: 286,3) an. Der p-Wert für die Veränderung der CD4-Zellzahl ab Baseline resultiert aus einem Wilcoxon Signed-Rank Test und beträgt 0,25, was keinen statistisch signifikanten Unterschied zu Woche 12 im Vergleich zu Baseline anzeigt.

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

4.3.2.3.3 Unerwünschte Ereignisse– 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 4-188: Operationalisierung von unerwünschten Ereignissen

Studie	Operationalisierung
GS-US-236-0123	<p><u>Mortalität</u> Anteil verstorbenen Patienten im Behandlungszeitraum bis Woche 12, 24 bzw. 48 (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage)</p> <p><u>Gesamtrate unerwünschter Ereignisse</u> Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 12, 24 bzw. 48</p> <p><u>Schwerwiegende unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 12, 24 bzw. 48</p> <p><u>Unerwünschte Ereignisse Grad 3-4</u> Anteil an Patienten mit mindestens einem unerwünschten Ereignis Grad 3-4 (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 12, 24 bzw. 48</p> <p><u>Therapieabbrüche aufgrund unerwünschter Ereignisse</u> Anteil an Patienten die die Studie aufgrund eines unerwünschten Ereignisses abgebrochen haben (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 12, 24</p>

Studie	Operationalisierung
	bzw. 48 <u>Häufigste unerwünschte Ereignisse</u> Anteil an Patienten mit mindestens einem häufigsten unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 12, 24 bzw. 48. Es werden die unerwünschten Ereignisse betrachtet, die bei mindestens 10% der Patienten auftraten. Häufigste unerwünschte Ereignisse, welche in der vorliegenden Fragestellung relevant sind, waren Erkrankungen des Gastrointestinaltrakts (SOC), allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), Erkrankungen des Nervensystems (SOC), psychiatrische Erkrankungen (SOC) und Erkrankungen der Haut und des Unterhautzellgewebes (SOC). Diese wurden jeweils präspezifiziert als SOC erfasst. Darunter werden die jeweils am häufigsten auftretenden unerwünschten Ereignisse (Auftreten bei $\geq 10\%$ der Patienten in einem der beiden Behandlungsarme) als PTs betrachtet.

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.

Tabelle 4-189: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse in weitere Untersuchungen

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver- zerrungspotenzial be- einflussende Punkte
GS-US-236-0123	nein	Nicht zutreffend, da unkontrollierte Studie	nein	nein

Da dies eine einarmige Studie ist, gab es keine Verblindung der Endpunkterheber. Eine adäquate Umsetzung des ITT-Prinzips ist hier nicht zutreffend, da keine randomisierte Gruppenzuteilung stattfand. Es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung und auch sonstige das Verzerrungspotenzial beeinflussende Punkte fehlen. Da unkontrollierte Studien jedoch generell potenziell hoch verzerrt sind, wird das Verzerrungspotenzial für diesen Endpunkt als hoch bewertet.

Unter unerwünschten Ereignissen wurden verschiedene sicherheitsbezogene Ereignisse betrachtet. Diese waren folgende:

- Anteil verstorbener Patienten (Mortalität)
- Anteil an Patienten mit mindestens einem unerwünschten Ereignis jeglicher Art (Gesamtrate unerwünschter Ereignisse)

- Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Schwerwiegende unerwünschte Ereignisse)
- Anteil an Patienten mit mindestens einem unerwünschten Ereignis Grad 3-4 (Unerwünschte Ereignisse Grad 3-4)
- Anteil an Patienten die die Therapie aufgrund unerwünschter Ereignisse vorzeitig abbrechen (Therapieabbrüche aufgrund unerwünschter Ereignisse)

Es werden die in der Studie am häufigsten beobachteten unerwünschten Ereignisse dargestellt. Dabei werden die folgenden SOC als relevant erachtet:

- Erkrankungen des Gastrointestinaltrakts
- Allgemeine Erkrankungen und Beschwerden am Verabreichungsort
- Erkrankungen des Nervensystems
- Psychiatrische Erkrankungen
- Erkrankungen der Haut und des Unterhautzellgewebes
- Augenerkrankungen (nur bei ATV)

Darunter werden jeweils die am häufigsten auftretenden PTs beobachtet. Dies sind die Erkrankungen, die bei mindestens 10% der Patienten auftraten.

Da bisher nur erste Daten nach 12 Wochen Behandlung veröffentlicht wurden, wird hier nur das Auftreten unerwünschter Ereignisse nach 12 Wochen Behandlung betrachtet.

Zur Erfassung der häufigsten unerwünschten Ereignisse mittels SOC wurde die MedDRA Version 15.1 verwendet.

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.

Tabelle 4-190: Ergebnisse für unerwünschte Ereignisse aus weiteren Untersuchungen (Safety Analyse-Set)

Endpunkt Studie	EVG/COBI/FTC/TDF	
	N	n (%)
Mortalität		
GS-US-236-0123	48	0 (0,0)
Gesamtrate unerwünschter Ereignisse		
GS-US-236-0123	48	36 (75,0)
Schwerwiegende unerwünschte Ereignisse		
GS-US-236-0123	48	0 (0,0)
Therapieabbrüche aufgrund unerwünschter Ereignisse		
GS-US-236-0123	48	0 (0,0)
Unerwünschte Ereignisse Grad 3-4		
GS-US-236-0123	48	0 (0,0)

N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis

In Studie GS-US-236-0123 gab es keine Todesfälle und es traten weder schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4 noch Therapieabbrüche aufgrund unerwünschter Ereignisse auf. Unerwünschte Ereignisse jeglicher Art traten bei 36 Patienten (75,0%) auf und elf Patienten (22,9%) davon hatten ein unerwünschtes Ereignis Grad 2.

Tabelle 4-191: Ergebnisse für häufigste unerwünschte Ereignisse (SOC, Grenze: $\geq 10\%$) aus weiteren Untersuchungen (Safety Analyse-Set)

Endpunkt Studie	EVG/COBI/FTC/TDF	
	N	n (%)
Erkrankungen des Gastrointestinaltrakts		
GS-US-236-0123	48	12 (25,0)
<i>Diarrhoe</i>		
GS-US-236-0123	48	5 (10,4)
Psychiatrische Erkrankungen		
GS-US-236-0123	48	11 (22,9)
<i>Schlaflosigkeit</i>		
GS-US-236-0123	48	5 (10,4)

N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis

Bei 12 (25,0%) Patienten traten Erkrankungen des Gastrointestinaltrakts auf. Davon hatten fünf (10,4%) Patienten Diarrhoe.

Psychiatrische Erkrankungen traten bei elf (22,9%) der Patienten auf. Davon war das bei den meisten Patienten auftretende Ereignis Schlaflosigkeit (fünf [10,4%] Patienten).

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

4.3.2.3.4 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Zum Zeitpunkt der Einreichung des vorliegenden Dossiers lagen noch keine Daten zu Subgruppen aus der betrachteten Studie GS-US-236-0123 vor.

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.

Ergebnisse aus dem indirekten Vergleich von EVG/COBI/FTC/TDF vs. RAL+FTC/TDF

Der Nutzen bzw. Zusatznutzen der Behandlung mit Elvitegravir/Cobicistat/Emtricitabin/Tenofoviridisoproxil (EVG/COBI/FTC/TDF) wird hier anhand eines indirekten Vergleichs gezeigt, der EVG/COBI/FTC/TDF in der zugelassenen Dosierung mit der zweckmäßigen Vergleichstherapie, Raltegravir (RAL) in Kombination mit Tenofoviridisoproxil und Emtricitabin (FTC/TDF), über den Brückenkompator, Efavirenz (EFV) in Kombination mit Tenofoviridisoproxil und Emtricitabin (FTC/TDF), vergleicht.

Die in den eingeschlossenen Studien untersuchte Studienpopulation besteht aus antiretroviral nicht-vorbehandelten erwachsenen Patienten mit einer HIV-1-Infektion. Aufgrund der Übertragbarkeit der Daten dieser Patienten auf vorbehandelte HIV-1 infizierte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, treffen die Ergebnisse dieser Studie auch auf die im vorliegenden Modul 4B behandelte Population zu.

In die Bewertung des Zusatznutzens sind folgende Endpunkte eingeflossen: Virologisches Ansprechen, Veränderung der CD4-Zellzahl, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschter Ereignisse, Studienabbrüche aufgrund unerwünschter Ereignisse und häufigste unerwünschte Ereignisse (Erkrankungen des Gastrointestinaltrakts

[SOC], Allgemeine Erkrankungen und Beschwerden am Verabreichungsort [SOC], Erkrankungen des Nervensystems [SOC], Psychiatrische Erkrankungen [SOC], Erkrankungen der Haut und des Unterhautzellgewebes [SOC]).

Virologisches Ansprechen (M=F)

Hinsichtlich des virologischen Ansprechens (erfasst mittels Missing=Failure-Algorithmus) besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 1,34 [0,87;2,05], $p=0,18$). Und auch im Vergleich von RAL+FTC/TDF mit EFV+FTC/TDF kann kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (OR [95%-KI]: 1,37 [0,87;2,16], $p=0,1743$). Insgesamt ergibt sich somit für den indirekten Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF kein statistisch signifikanter Unterschied hinsichtlich virologischem Ansprechen (OR [95%-KI]: 0,98 [0,52;1,83]). Auch zu Woche 96 besteht weder in den EVG/COBI/FTC/TDF-Studien noch in der RAL+FTC/TDF-Studie ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI] EVG/COBI/FTC/TDF vs. EFV/FTC/TDF: 1,26 [0,83;1,90], $p=0,27$; RAL+FTC/TDF vs. EFV+FTC/TDF: 1,14 [0,75;1,72], $p=0,5404$). Der indirekte Vergleich von EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,11 [0,61;1,99]).

Veränderung der CD4-Zellzahl

Zu Woche 48 besteht in den EVG/COBI/FTC/TDF-Studien ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (MWD [95%-KI] aus Meta-Analyse: 36,26 [12,42;60,10], $p=0,003$). Allerdings zeigt auch der Vergleich zwischen RAL+FTC/TDF und EFV+FTC/TDF einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF (MWD [95%-KI]: 26 [4;47], $p=0,0184$). Durch den indirekten Vergleich kann kein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF beobachtet werden (MWD [95%-KI]: 10,26 [-21,84;42,36]). Zu Woche 96 besteht weder zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF (MWD [95%-KI]: 22 [-10;54], $p=0,19$) noch zwischen RAL+FTC/TDF und EFV+FTC/TDF (MWD [95%-KI]: 15 [-13;42], $p=0,2745$) ein statistisch signifikanter Unterschied zwischen den jeweiligen Behandlungen. Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt, dass hinsichtlich der Veränderung der CD4-Zellzahl kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht (MWD [95%-KI]: 7 [-35,19;49,19]).

Mortalität

Hinsichtlich der Mortalitätsrate besteht zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,50 [0,05;5,59], $p=0,5769$). Und auch zwischen RAL+FTC/TDF und EFV+FTC/TDF besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 5,05 [0,24;105,74], $p=0,2964$). Der indirekte Vergleich von EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt, dass zwischen diesen beiden Behandlungen kein statistisch signifikanter Unterschied besteht (OR [95%-KI]: 0,10 [0,00;4,66]). Auch zu Woche 96

besteht weder zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF (OR [95%-KI]: 0,50 [0,05;5,59], $p=0,5769$) noch zwischen RAL+FTC/TDF und EFV+FTC/TDF (OR [95%-KI]: 7,10 [0,37;138,10], $p=0,1955$) ein statistisch signifikanter Unterschied. Der indirekte Vergleich zeigt, dass kein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF besteht (OR [95%-KI]: 0,07 [0,00;3,10]).

Gesamtrate unerwünschter Ereignisse

Zu Woche 48 besteht zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI] aus Meta-Analyse: 0,92 [0,51;1,68], $p=0,80$). Allerdings besteht zwischen RAL+FTC/TDF und EFV+FTC/TDF ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF (OR [95%-KI]: 0,33 [0,16;0,70], $p=0,002$). Der indirekte Vergleich zeigt, dass hinsichtlich der Vermeidung unerwünschter Ereignisse ein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zuungunsten von EVG/COBI/FTC/TDF besteht (OR [95%-KI]: 2,79 [1,08;7,20]). Wählt man hier als Effektmaß das relative Risiko, besteht der statistisch signifikante Unterschied zwischen den beiden Behandlungen nur noch knapp (RR [95%-KI]: 1,06 [1,003;1,13]). Zu Woche 96 besteht kein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF (OR [95%-KI]: 0,90 [0,38;2,14], $p=0,8041$). Und auch zwischen RAL+FTC/TDF und EFV+FTC/TDF kann kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (OR [95%-KI]: 0,45 [0,18;1,12], $p=0,086$). Der zu Woche 48 bestehende statistisch signifikante Unterschied aus dem indirekten Vergleich zuungunsten von EVG/COBI/FTC/TDF besteht zu Woche 96 nicht mehr (OR [95%-KI]: 2,00 [0,57;7,04]).

Schwerwiegende unerwünschte Ereignisse

Zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF besteht zu Woche 48 hinsichtlich des Auftretens schwerwiegender unerwünschter Ereignisse ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 1,77 [1,06;2,97], $p=0,03$). Zwischen RAL+FTC/TDF und EFV+FTC/TDF ist der Unterschied nicht statistisch signifikant (OR [95%-KI]: 1,05 [0,60;1,82], $p=0,888$). Der indirekte Vergleich zeigt, dass zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF kein statistisch signifikanter Unterschied besteht (OR [95%-KI]: 1,69 [0,79;3,59]). Zu Woche 96 besteht zwar zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF ein statistisch signifikanter Unterschied (OR [95%-KI]: 1,85 [1,17;2,93], $p=0,0083$), doch zwischen RAL+FTC/TDF und EFV+FTC/TDF kann kein statistisch signifikanter Unterschied zwischen den Behandlungen festgestellt werden (OR [95%-KI]: 1,21 [0,74;1,98], $p=0,457$). Wie in Modul 4A ausgeführt, ist der Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF dabei primär auf einen Unterschied in der SOC Infektionen und parasitäre Erkrankungen und dabei insbesondere die Diagnosen Cellulitis und Appendicitis zurückzuführen – diese wurden vom jeweiligen Prüfarzt als nicht in Zusammenhang mit der Studienmedikation stehend beurteilt. Rechnet man beide Diagnosen heraus, geht der statistisch signifikante Unterschied verloren. Der indirekte Vergleich von EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt, dass hinsichtlich der Vermeidung

schwerwiegender unerwünschter Ereignisse kein statistisch signifikanter Unterschied zwischen den Behandlungen besteht (OR [95%-KI]: 1,53 [0,78;3,00]).

Therapieabbrüche aufgrund unerwünschter Ereignisse

Therapieabbrüche aufgrund unerwünschter Ereignisse traten bis Woche 48 zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF statistisch nicht signifikant unterschiedlich häufig aus (OR [95%-KI] aus Meta-Analyse: 0,67 [0,33;1,36], $p=0,27$). Und auch zwischen RAL+FTC/TDF und EFV+FTC/TDF kann dazu kein statistisch signifikanter Unterschied festgestellt werden (OR [95%-KI]: 0,52 [0,23;1,18], $p=0,159$). Der indirekte Vergleich zeigt keinen statistisch signifikanten Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF (OR [95%-KI]: 1,29 [0,44;3,80]). Auch zu Woche 96 besteht weder zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF (OR [95%-KI]: 0,70 [0,37;1,33], $p=0,2782$) noch zwischen RAL+FTC/TDF und EFV+FTC/TDF (OR [95%-KI]: 0,64 [0,29;1,38], $p=0,333$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,09 [0,40;3,00]).

Neurologische & psychiatrische bzw. neuropsychiatrische Ereignisse

Zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF besteht zu Woche 48 ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,46 [0,34;0,61], $p<0,00001$). Die Sensitivitätsanalyse, unter Ausschluss der Studie S104 zeigt robuste Ergebnisse (OR [95%-KI]: 0,45 [0,33;0,61], $p<0,001$). Im Vergleich zwischen RAL+FTC/TDF und EFV+FTC/TDF besteht zu Woche 48 ebenso ein statistisch signifikanter Unterschied zwischen den Behandlungen, der zugunsten RAL+FTC/TDF ausgeprägt ist (OR [95%-KI] für den präspezifizierten Endpunkt: 0,52 [0,34;0,81], $p=0,0044$). Eine Sensitivitätsanalyse, unter Einschluss weiterer PTs in die Gruppierung, zeigt ebenso einen statistisch signifikanten Unterschied zugunsten von RAL+FTC/TDF (OR [95%-KI]: 0,25 [0,17;0,36], $p<0,0001$). Da Art und Anzahl der eingeschlossenen PTs unter den in den Studien S102, S104 und STARTMRK betrachteten Endpunkten stark voneinander abwichen, kann kein indirekter Vergleich zu diesem Endpunkt durchgeführt werden.

Gastrointestinale Erkrankungen (SOC)

Diarrhoe

Zu Woche 96 besteht zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,08 [0,76;1,53], $p=0,6613$). Allerdings besteht zwischen RAL+FTC/TDF und EFV+FTC/TDF ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von RAL+FTC/TDF (OR [95%-KI]: 0,64 [0,42;0,96], $p=0,0315$). Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt jedoch, dass kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht (OR [95%-KI]: 1,69 [0,98;2,90]).

Übelkeit

Hinsichtlich Übelkeit zu Woche 96 besteht zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,58 [1,07;2,32], $p=0,0213$). Zwischen RAL+FTC/TDF und EFV+FTC/TDF besteht jedoch kein statistisch signifikanter Unterschied (OR [95%-KI]: 1,13 [0,70;1,84], $p=0,6102$). Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt, dass kein statistisch signifikanter Unterschied zwischen den Behandlungen besteht (OR [95%-KI]: 1,40 [0,75;2,60]).

Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)

Fatigue

Zu Woche 96 ist der Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF nicht statistisch signifikant (OR [95%-KI]: 0,86 [0,56;1,32], $p=0,4810$). Zwischen RAL+FTC/TDF und EFV+FTC/TDF ist der Unterschied statistisch signifikant zugunsten von RAL+FTC/TDF (OR [95%-KI]: 0,55 [0,30;0,99], $p=0,0454$). Der indirekte Vergleich zeigt, dass hinsichtlich der Vermeidung von Fatigue kein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF besteht (OR [95%-KI]: 1,56 [0,75;3,26]).

Erkrankungen des Nervensystems (SOC)

Schwindel

Zu Woche 96 kann hinsichtlich Schwindel ein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF zugunsten von EVG/COBI/FTC/TDF beobachtet werden (OR [95%-KI]: 0,22 [0,14;0,36], $p<0,00001$). Auch der Vergleich zwischen RAL+FTC/TDF und EFV+FTC/TDF zeigt einen statistisch signifikanten Unterschied zugunsten von RAL+FTC/TDF zwischen den Behandlungen (OR [95%-KI]: 0,15 [0,09;0,25], $p<0,00001$). Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt, dass hinsichtlich der Vermeidung von Schwindel kein statistisch signifikanter Unterschied zwischen den Behandlungen besteht (OR [95%-KI]: 1,47 [0,73;2,94]).

Kopfschmerzen

Hinsichtlich des Auftretens von Kopfschmerzen besteht zu Woche 96 kein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF (OR [95%-KI]: 1,50 [0,97;2,31], $p=0,0702$). Und auch zwischen RAL+FTC/TDF und EFV+FTC/TDF kann kein statistisch signifikanter Unterschied festgestellt werden (OR [95%-KI]: 0,91 [0,62;1,34], $p=0,6365$). Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,65 [0,92;2,94]).

Psychiatrische Erkrankungen (SOC)

Abnorme Träume

Zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF besteht hinsichtlich dem Auftreten von abnormen Träumen zu Woche 96 ein statistisch signifikanter Unterschied zugunsten von

EVG/COBI/FTC/TDF (OR [95%-KI]: 0,47 [0,32;0,69], p=0,0001). Zwischen RAL+FTC/TDF und EFV+FTC/TDF kann ein statistisch signifikanter Unterschied zugunsten von RAL+FTC/TDF festgestellt werden (OR [95%-KI]: 0,53 [0,30;0,94], p=0,0294). Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,89 [0,45;1,76]).

Schlaflosigkeit

Hinsichtlich des Auftretens von Schlaflosigkeit besteht zu Woche 96 ein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,63 [0,41;0,99], p=0,0428). Der Unterschied zwischen RAL+FTC/TDF und EFV+FTC/TDF hingegen zeigt keine statistische Signifikanz (OR [95%-KI]: 1,11 [0,66;1,87], p=0,6813). Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,57 [0,29;1,12]).

Depression

Nach 96 Wochen Behandlung kann weder zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF (OR [95%-KI]: 0,85 [0,55;1,32], p=0,4668) noch zwischen RAL+FTC/TDF und EFV+FTC/TDF (OR [95%-KI]: 0,83 [0,45;1,52], p=0,5470) ein statistisch signifikanter Unterschied beobachtet werden. Der indirekte Vergleich zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt, dass zwischen den Behandlungen kein statistisch signifikanter Unterschied besteht (OR [95%-KI]: 1,02 [0,48;2,17]).

Erkrankungen der Haut und des Unterhautzellgewebes (SOC)

Hautausschlag

Nach 96 Wochen zeigt sich zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF hinsichtlich der Vermeidung von Hautausschlag ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,50 [0,30;0,82], p=0,0066). Auch zwischen RAL+FTC/TDF und EFV+FTC/TDF besteht ein statistisch signifikanter Unterschied, der sich zugunsten RAL+FTC/TDF ausprägt (OR [95%-KI]: 0,47 [0,26;0,86], p=0,0148). Der indirekte Vergleich der beiden Behandlungen EVG/COBI/FTC/TDF und RAL+FTC/TDF zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,06 [0,49;2,32]).

Ergebnisse aus der Studie zur Unterstützung der Übertragbarkeit der Daten (GS-US-236-0123)

Virologisches Ansprechen (Snapshot)

Nach 12 Wochen Behandlung mit EVG/COBI/FTC/TDF wiesen alle Patienten (48 [100%]) virologisches Ansprechen auf.

Veränderung der CD4-Zellzahl

Nachdem die Patienten zu Baseline schon durch die Vorbehandlung recht hohe CD4-Zellzahlen aufwiesen (durchschnittlich 711 Zellen/ μ L), stieg die durchschnittliche CD4-Zellzahl bis Woche 12 noch um im Mittel 33 Zellen/ μ L auf 744 Zellen/ μ L an. Der Unterschied zu Baseline war jedoch nicht statistisch signifikant ($p=0,25$).

Unerwünschte Ereignisse

In Studie GS-US-236-0123 gab es keine Todesfälle und es traten weder schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4 noch Therapieabbrüche aufgrund unerwünschter Ereignisse auf. Unerwünschte Ereignisse jeglicher Art traten bei 36 Patienten (75,0%) auf und elf Patienten (22,9%) davon hatten ein unerwünschtes Ereignis Grad 2.

Bei 12 (25,0%) Patienten traten Erkrankungen des Gastrointestinaltrakts auf. Davon hatten fünf (10,4%) Patienten Diarrhoe.

Psychiatrische Erkrankungen traten bei elf (22,9%) der Patienten auf. Davon war das bei den meisten Patienten auftretende Ereignis Schlaflosigkeit (fünf [10,4%] Patienten).

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.

Es konnte keine Studie gefunden werden, die sich zur Bewertung des Zusatznutzen von EVG/COBI/FTC/TDF in der Population der vorbehandelten erwachsenen, HIV-1 infizierten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, geeignet hätte.

Aufgrund der Übertragbarkeit der Daten zu nicht-vorbehandelten Patienten auf vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, wurden hier die Studien zu nicht-vorbehandelten Patienten eingeschlossen.

Für den Vergleich von EVG/COBI/FTC/TDF mit EFV/FTC/TDF findet hier somit die Beurteilung der Aussagekraft der Nachweise aus Modul 4A Anwendung.

EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF

Die in die Nutzenbewertung von EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF eingeschlossene Studie ist eine randomisierte, kontrollierte Studien und entspricht demnach der Evidenzstufe Ib. Die Studie enthält einen direkten Vergleich des zu bewertenden Arzneimittels EVG/COBI/FTC/TDF mit der zweckmäßigen Vergleichstherapie ATV/r+FTC/TDF. Die Vollständigkeit des Studienpools wurde durch eine umfassende systematische Studienregistersuche bestätigt (siehe Abschnitt 4.2.3.3 und Anhang 4-B).

Die Studie wurde in einem double-dummy Parallelgruppendesign durchgeführt und wurde auf Studienebene mit einem niedrigen Verzerrungspotenzial bewertet. Auf Endpunktebene bestand für alle Endpunkte ein niedriges Verzerrungspotenzial. Es wurden post-hoc Subgruppenanalysen für jeden eingeschlossenen Endpunkt durchgeführt, um den Anforderungen des G-BA an ein Nutzendossier gerecht zu werden und um ein möglichst umfassendes Bild über den Nutzen und Zusatznutzen des zu bewertenden Arzneimittels zu geben.

Da eine randomisierte und kontrollierte Studie zur Bewertung des Zusatznutzens herangezogen wurde, sind grundsätzlich Hinweise für einen Zusatznutzen ableitbar.

Alle in die Nutzenbewertung eingeschlossenen Endpunkte sind valide und patientenrelevante Endpunkte, die auch von u.a. dem IQWiG und der EMA als solche anerkannt wurden. Die Verwendung von Surrogatendpunkten wurde in Abschnitt 4.5.4 begründet.

EVG/COBI/FTC/TDF vs. RAL+FTC/TDF

Alle in die vorliegende Nutzenbewertung eingeschlossenen Studien sind randomisierte kontrollierte Studien und entsprechen demnach der Evidenzstufe Ib. Das Verzerrungspotenzial wurde auf Studienebene jeweils als niedrig eingeschätzt, da keine Faktoren gefunden werden konnten, die auf ein hohes Verzerrungspotenzial schließen ließen.

Auf Endpunktebene wurden alle Endpunkte, abgesehen von neurologischen & psychiatrischen Ereignissen in Studie S104, als niedrig verzerrt bewertet. Somit sind valide Schlüsse mit hoher Ergebnissicherheit möglich. Für wichtige unerwünschte Ereignisse (neurologische & psychiatrische Ereignisse) konnte das Verzerrungspotenzial in Studie S104 aufgrund fehlender Präspezifikation nicht sicher ausgeschlossen werden. Dies wird in der Ableitung des Zusatznutzens berücksichtigt.

Alle verwendeten Endpunkte sind valide und in der Indikation etablierte Endpunkte.

Unterschiede in den Baseline-Charakteristika, wie zum Beispiel dass in der STARTMRK mehr Patienten mit Ausgangsviruslast >100.000 Kopien/mL eingeschlossen wurden, woraus vermutlich auch die geringere durchschnittliche CD4-Zellzahl zu Baseline resultiert, werden in der Einschätzung der Aussagesicherheit berücksichtigt.

Grundsätzlich eignet sich dieser indirekte Vergleich um Hinweise auf einen Zusatznutzen abzuleiten. Für Endpunkte, deren Verzerrungspotenzial nicht sicher ausgeschlossen werden

konnte, wird die Ergebnissicherheit gegebenenfalls herabgestuft und bei der Einschätzung der Wahrscheinlichkeit des Zusatznutzens berücksichtigt.

Zur Unterstützung der Übertragbarkeit der Daten zu den nicht-vorbehandelten Patienten auf die vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, wurden 12-Wochen Ergebnisse einer einarmigen Studie präsentiert.

Diese einarmige Studie entspricht der Evidenzklasse IIa. Da nicht-kontrollierte Studien generell potenziell hoch verzerrt sind, wurde das Verzerrungspotenzial dieser Studie sowohl auf Studienebene als auch auf Endpunktebene als hoch eingestuft.

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.

Bei den Therapieprinzipien der ART gilt es als generelles Ziel der Therapie, durch die Hemmung der HIV-Replikation infektionsbedingte Symptome zu unterdrücken, Krankheitsprogression zu vermindern, eine immunologische Rekonstitution zu erreichen und die chronische Immunaktivierung zu reduzieren. Nur eine hohe virologische Wirksamkeit der ART und eine dauerhafte Suppression der Viruslast auf unter 50 RNA-Kopien/mL können Resistenzentwicklung und konsekutives Therapieversagen nachhaltig verhindern [2].

Grundsätzlich unterscheidet man in der ART zwei Patientenpopulationen, für die die übergeordneten Therapieprinzipien (dauerhafte Hemmung der Replikation und virale Suppression auf <50 Kopien/mL) gleichermaßen gültig sind:

1. Patienten, die noch keine antiretrovirale Therapie erhalten (nicht-vorbehandelte Patienten):
2. Patienten, die antiretroviral behandelt werden (vorbehandelte Patienten)

In den Empfehlungen zur Umstellung antiretroviral vorbehandelter Patienten wird in allen nationalen und internationalen Leitlinien zwischen zwei Subpopulationen, bzw. Indikationen zur Umstellung unterschieden [2-6]:

- Supprimierte Patienten: Umstellung aufgrund von Nebenwirkungen oder zur Therapievereinfachung und Unterstützung, bzw. Verbesserung der Adhärenz
- Patienten von mit virologischem Versagen: Umstellung aufgrund des virologischen Versagens und konsekutiver Virämie

Diese Unterscheidung ergibt sich aus dem entscheidenden Kriterium, das die Grundvoraussetzung für die virologische Wirksamkeit einer ART darstellt: Dem Vorliegen Resistenz-assoziiierter Mutationen gegen die einzusetzenden Substanzen.

Die Unterscheidung zwischen Patienten mit und ohne Resistenz-assoziierten Mutationen ist von elementarer Bedeutung: Ein vorbehandelter Patient, dessen Virus keine Resistenz-assoziierten Mutationen gegen antiretrovirale Substanzen aufweist, entspricht hinsichtlich des Resistenz-Status einem nicht-vorbehandelten Patienten. Damit ist dieser Patient im Hinblick auf die entsprechenden antiretroviralen Substanzen als funktionell nicht-vorbehandelt zu definieren - womit beide Populationen (nicht-vorbehandelt und vorbehandelt ohne Resistenzen) als gleichgestellt zu betrachten sind. Vor dem Hintergrund der Gleichstellung der Population können klinische Daten zu nicht-vorbehandelten Patienten somit auf vorbehandelte Patienten ohne Resistenzen extrapoliert und damit übertragen werden [7].

Die Rationale der Übertragbarkeit klinischer Daten von nicht-vorbehandelten Patienten auf vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, lässt sich zudem wie folgt belegen:

1. Die DHHS-Leitlinien führen in den Empfehlungen zur Umstellung supprimierter Patienten explizit aus: „Für Regime, die zur Therapievereinfachung eingesetzt werden, sollte grundsätzlich eine hohe Wirksamkeit bei antiretroviral nicht-vorbehandelten Patienten nachgewiesen worden sein, oder es sollte für den Patienten anhand seiner individuellen Vorbehandlung und seines Resistenzprofils eine hohe Wirksamkeit vorhergesagt werden können“ [4]. Gemäß der DHHS-Leitlinie ist somit davon auszugehen, dass die nachgewiesene Wirksamkeit eines antiretroviralen Regimes bei nicht-vorbehandelten Patienten auf suppressierte vorbehandelte Patienten extrapoliert und damit übertragen werden kann.
2. Die Übertragbarkeit der Daten ist hinsichtlich der Wirksamkeit und des Verträglichkeitsprofils auch rein mechanistisch und zwar sowohl pharmakologisch als auch biologisch nachvollziehbar:

Eine Initialtherapie muss ausreichend virologisch wirksam sein, um das Virus erfolgreich zu kontrollieren und die Ausgangsviruslast unabhängig von ihrer Höhe (etwa 50% der Patienten weisen eine Ausgangsviruslast von >100.000 Kopien/mL auf) unter die Nachweisgrenze zu senken. Antiretrovirale Therapieregime, die zur Umstellung supprimierter Patienten eingesetzt werden, hingegen müssen lediglich ausreichend virologisch wirksam sein, um die Viruslast unter der Nachweisgrenze zu halten. Entsprechend muss die virologische Potenz einer antiretroviralen Initialtherapie mindestens so hoch sein wie die eines Regimes, auf das ein suppressierter Patient umgestellt wird. Vor diesem Hintergrund ist davon auszugehen, dass klinische Daten zur virologischen Wirksamkeit, die bei nicht-vorbehandelten Patienten erhoben werden, direkt auf die Umstellung vorbehandelter Patienten extrapolierbar sind, sofern diese keine Resistenzen gegen die Substanzen aufweisen.

Unter Therapie auftretende Nebenwirkungen sind primär Substanz- und nicht Virus-assoziiert, woraus schlussgefolgert werden kann, dass die Höhe der Viruslast zu Beginn einer bestimmten Therapie keinen Einfluss auf die auftretenden Nebenwirkungen hat. Entsprechend wird auch im Bericht zum Sicherheitsprofil in den Fachinformationen [8-35] zu den verschiedenen antiretroviralen Substanzen nicht zwischen vorbehandelten und nicht-vorbehandelten Patienten unterschieden – ganz im Gegenteil beruht das in den Fachinformationen berichtete Sicherheitsprofil stets auf Basis des gesamten klinischen Entwicklungsprogramms.

3. Die EMA hat für Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil sowohl die Zulassung für nicht-vorbehandelte Patienten, als auch die Zulassung für die Patienten erteilt, die keine Resistenzen gegen einen der drei enthaltenen antiretroviralen Substanzen aufweisen. Da diese Zulassung ausschließlich auf Basis von klinischen Daten zu nicht-vorbehandelten Patienten eingereicht wurde, muss geschlussfolgert werden, dass die EMA die beiden Population auf Basis der dargestellten wissenschaftlichen Rationale funktionell gleichgestellt und die Daten entsprechend extrapoliert hat. Dies wird auch durch ein von der EMA erstelltes Konzeptpapier zur Extrapolation von klinischen Daten klar unterstützt: Die angestrebte Extrapolation ist definiert als Deduktion von Studienergebnissen in

bestimmten Subpopulation (Quellenpopulation) auf andere Subpopulationen (Zielpopulation) – analog gilt dies für verwandte Erkrankungen oder Substanzen. Zielsetzung der Extrapolation von Daten ist zum einen die Reduktion des Bedarfs der zusätzlichen Datengenerierung für die Zielpopulation, und zum anderen die Vermeidung unnötiger Studien und Allokation von Ressourcen in die klinischen Studien, die tatsächlich notwendig sind. Grundvoraussetzung für die Übertragbarkeit ist eine pharmakologische und biologische Rationale, die für Quellen- und Zielpopulation anhand von Erkrankung, Medikation (Wirkmechanismus, Pharmakokinetik und Pharmakodynamik) sowie Anwendbarkeit von Wirksamkeits- und Sicherheitsendpunkten bewertet werden sollte [7]. Für die beiden hier vorliegenden Populationen der nicht-vorbehandelten Quellenpopulation und der Zielpopulation von Patienten ohne Resistenz-assoziierte Mutationen, zeigt sich die geforderte pharmakologische und biologische Rationale deutlich bei der vollständigen Übereinstimmung der genannten Parameter. Vor diesem Hintergrund ist mit Erteilung der Zulassung also davon auszugehen, dass die Europäische Zulassungsbehörde die Übertragbarkeit der Daten von den nicht-vorbehandelten Patienten auf die vorbehandelten Patienten ohne Resistenz-assoziierte Mutationen als gegeben ansieht.

In logischer Konsequenz bedeutet die dargestellte grundsätzliche Übertragbarkeit der Daten, dass auch ein in den klinischen Studien zu nicht-vorbehandelten Patienten gefundener Zusatznutzen für auf die vorbehandelte Patientenpopulation, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, zu übertragen ist.

Die Übertragbarkeit der Daten wird zusätzlich gestützt durch die Ergebnisse der einarmigen Studie GS-US-236-0123, in welcher Patienten, die mindestens sechs Monate mit RAL+FTC/TDF vorbehandelt wurden, auf EVG/COBI/FTC/TDF umgestellt wurden. Nach 12 Wochen Behandlung mit EVG/COBI/FTC/TDF zeigten alle 48 (100%) Patienten virologisches Ansprechen. Hinsichtlich der CD4-Zellzahl zeigten die Patienten eine numerische, aber nicht signifikante Verbesserung (Baseline: 711 Zellen/ μ L, zu Woche 12: 744 Zellen/ μ L). In der Studie traten weder Todesfälle, noch schwerwiegende unerwünschte Ereignisse oder Therapieabbrüche aufgrund unerwünschter Ereignisse auf. 36 (75,0%) Patienten hatten unerwünschte Ereignisse; davon wurde bei elf Patienten (22,9%) das unerwünschte Ereignis als Grad 2 eingestuft. Unerwünschte Ereignisse höheren Grads traten nicht auf. 12 (25,0%) Patienten wiesen Erkrankungen des Gastrointestinaltrakts auf; hier war Diarrhoe das am häufigsten auftretende Ereignis aus dieser SOC (fünf [10,4%] Patienten). Psychiatrische Erkrankungen traten bei elf (22,9%) Patienten auf, mit Schlaflosigkeit als am häufigsten auftretendes Ereignis (fünf [10,4%] Patienten). Für eine detailliertere Übersicht der Ergebnisse aus dieser Studie siehe Abschnitt 4.3.2.3.

EVG/COBI/FTC/TDF vs. EFV/FTC/TDF

Aus den oben aufgeführten Gründen besteht hier ein **Beleg** für einen **geringen** Zusatznutzen durch EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF, für die Population der erwachsenen mit HIV-1 infizierten vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei

antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind. Zu weiteren Details und Ausführungen des Zusatznutzens von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF siehe Modul 4A.

EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF

Der Zusatznutzen von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF lässt sich auf Endpunktebene sowie insgesamt wie folgt ableiten.

Es wurde eine randomisierte aktiv-kontrollierte Studie, S103, in die Nutzenbewertung eingeschlossen, welche EVG/COBI/FTC/TDF mit der zweckmäßigen Vergleichstherapie ATV/r+FTC/TDF vergleicht. Es wurden dazu folgende patientenrelevante Endpunkte untersucht:

- Virologisches Ansprechen, definiert als Abfall der Viruslast unter die Nachweisgrenze von 50 Kopien/mL
- Zeit bis zum Verlust des virologischen Ansprechens
- Veränderung der CD4-Zellzahl
- Unerwünschte Ereignisse
 - Mortalität
 - Gesamtrate unerwünschter Ereignisse
 - Gesamtrate schwerwiegender unerwünschter Ereignisse
 - Therapieabbrüche aufgrund unerwünschter Ereignisse
 - Unerwünschte Ereignisse Grad 3-4
 - Neurologische & psychiatrische Ereignisse
 - Hautausschläge
 - Renale Ereignisse
 - Knochenfrakturen
 - Häufigste unerwünschte Ereignisse

Zur Ableitung des Zusatznutzens werden Ergebnisse zu Woche 48 und Woche 96 herangezogen. In der folgenden Tabelle werden die Ergebnisse auf Endpunktebene und die als fazitrelevant eingestuften Hinweise bzw. Belege auf Effektmodifikation dargestellt. Basierend auf diesen Ergebnissen wird das Ausmaß des Zusatznutzens und dessen Wahrscheinlichkeit auf Endpunktebene, unter Angabe der Population, abgeleitet. Im Falle eines statistisch signifikanten Unterschieds zu lediglich einem der betrachteten Zeitpunkte (zum Beispiel zu Woche 48 statistisch signifikanter Unterschied, zu Woche 96 jedoch nicht mehr), wird die Wahrscheinlichkeit des Zusatznutzens für diesen Endpunkt herabgestuft.

Tabelle 4-192: Ausmaß des Zusatznutzen von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF auf Endpunktebene

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
Virologisches Ansprechen Snapshot TLOVR	<p><u>Woche 48:</u> OR: 1,30 [0,82;2,06], p=0,22 89,5% vs. 86,8%</p> <p><u>Woche 96:</u> OR: 1,08 [0,73; 1,59], p=0,70 83,3% vs. 82,3%</p> <p><u>Woche 48:</u> OR: 1,11 [0,73;1,69], p=0,55 86,1% vs. 84,8%</p> <p><u>Woche 96:</u> OR: 1,27 [0,88;1,82], p=0,19 81,3% vs. 77,5%</p> <p><i>Interaktionstest: p=0,06</i> Therapieadhärenz <95%: OR: 0,80 [0,44;1,45], p=0,4568 Therapieadhärenz ≥95%: OR: 1,69 [1,04;2,74], p=0,0337</p>	<p>Kein Zusatznutzen oder Schaden ableitbar (Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF konnte gezeigt werden)</p> <p>Kein Zusatznutzen oder Schaden ableitbar</p>
Zeit bis zum Verlust des virologischen Ansprechens	<p><u>Woche 48:</u> HR: 0,87 [0,59;1,27], p=0,48</p> <p><u>Woche 96:</u> HR: 0,81 [0,58;1,12], p=0,21</p> <p><i>Interaktionstest: p=0,04</i> Therapieadhärenz <95%: HR: 1,19 [0,72;1,95], p=0,49 Therapieadhärenz ≥95%: HR: 0,60 [0,39;0,94], p=0,025</p>	Kein Zusatznutzen oder Schaden ableitbar
Veränderung der CD4-Zellzahl	<p><u>Woche 48:</u> MWD: -6 [-31;18], p=0,61</p> <p><u>Woche 96:</u> MWD: -8 [-35;19], p=0,56</p> <p><i>Interaktionstest: p=0,04</i></p>	<p>Endpunktkategorie: Aids-definierende Erkrankungen/Tod</p> <p><u>≤350 Zellen/μL:</u> Ausmaß des Zusatznutzen/Schaden: Geringer Schaden</p>

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
	<p>Männlich: MWD: 7,00 [-21,43;35,43], p=0,6290</p> <p>Weiblich: MWD: 107,00 [-211,21;-2,79], p=0,0442</p> <p><i>Interaktionstest: p=0,02</i></p> <p>≤350 Zellen/μL: MWD: -40,00 [-77,16;-2,84], p=0,0350</p> <p>>350 Zellen/μL: MWD: 25,00 [-15,32;65,32], p=0,2274</p>	<p>Wahrscheinlichkeit: Anhaltspunkt</p>
Unerwünschte Ereignisse		
Mortalität	<p><u>Woche 48:</u> OR: 0,14 [0,01;2,77], p=0,1980 0,0% vs. 0,8%</p> <p><u>Woche 96:</u> OR: 0,14 [0,01;2,77], p=0,1980 0,0% vs. 0,8%</p>	Kein Zusatznutzen oder Schaden ableitbar
Gesamtrate UEs	<p><u>Woche 48:</u> OR: 0,71 [0,40;1,26], p=0,2423 91,5% vs. 93,8%</p> <p><u>Woche 96:</u> OR: 0,69 [0,33;1,47], p=0,3376 95,2% vs. 96,6%</p>	Kein Zusatznutzen oder Schaden ableitbar
Gesamtrate SUEs	<p><u>Woche 48:</u> OR: 0,83 [0,48;1,43], p=0,5043 7,4% vs. 8,7%</p> <p><u>Woche 96:</u> OR: 0,65 [0,41;1,03], p=0,0684 9,6% vs. 14,1%</p> <p><i>Interaktionstest: p=0,19</i></p> <p>≤100.000 Kopien/mL: OR: 0,84 [0,46;1,53], p=0,5731</p> <p>>100.000 Kopien/mL: OR: 0,45 [0,21;0,94], p=0,0325</p>	<p>Endpunktkategorie: Schwerwiegende Nebenwirkungen</p> <p><u>>100.000 Kopien/mL:</u> Ausmaß des Zusatznutzen/Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Anhaltspunkt</p>
Therapieabbrüche aufgrund UEs	<p><u>Woche 48:</u> OR: 0,72 [0,35;1,48], p=0,3689 3,7% vs. 5,1%</p>	Kein Zusatznutzen oder Schaden ableitbar

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
	<u>Woche 96:</u> OR: 0,71 [0,36;1,39], p=0,3151 4,2% vs. 5,9%	
UES vom Grad 3-4	<u>Woche 48:</u> OR: 0,93 [0,60;1,45], p=0,7607 12,7% vs. 13,5% <i>Interaktionstest: p=0,0002</i> ≤100.000 Kopien/mL: OR: 1,95 [1,08;3,52], p=0,0262 >100.000 Kopien/mL: OR: 0,32 [0,15;0,67],p=0,0025 <u>Woche 96:</u> OR: 0,89 [0,60;1,33], p=0,5648 15,6% vs. 17,2% <i>Interaktionstest: p=0,002</i> ≤100.000 Kopien/mL: OR: 1,48 [0,89;2,46], p=0,1268 >100.000 Kopien/mL: OR: 0,37 [0,18;0,74], p=0,0050	Endpunktkategorie: Schwerwiegende Nebenwirkungen <u>>100.000 Kopien/mL:</u> Ausmaß des Zusatznutzen/Schaden: Erheblicher Zusatznutzen Wahrscheinlichkeit: Hinweis <u>≤100.000 Kopien/mL:</u> Ausmaß des Zusatznutzen/Schaden: Geringer Schaden Wahrscheinlichkeit: Anhaltspunkt
Neurologische & psychiatrische Ereignisse	<u>Woche 48:</u> OR: 0,94 [0,69;1,27], p=0,69 34,6% vs. 36,1% <u>Woche 96:</u> OR: 0,95 [0,71;1,28], p=0,76 41,4% vs. 42,5%	Kein Zusatznutzen oder Schaden ableitbar
Hautausschläge	<u>Woche 48:</u> OR: 0,99 [0,67;1,45], p=1,00 17,8% vs. 18,0% <u>Woche 96:</u> OR: 0,93 [0,65;1,32], p=0,72 21,5% vs. 22,8%	Kein Zusatznutzen oder Schaden ableitbar
Renale Ereignisse	<u>Woche 48:</u> OR: n.b. 0,0% vs. 0,0% <u>Woche 96:</u> OR: 9,15 [0,49;170,67], p=0,061	Kein Zusatznutzen oder Schaden ableitbar

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
	1,1% vs. 0,0%	
Knochenfrakturen	<p><u>Woche 48:</u> OR: 0,50 [0,12;2,01], p=0,51 0,8% vs. 1,7%</p> <p><u>Woche 96:</u> OR: 0,28 [0,09;0,86], p=0,029 1,1% vs. 3,9%</p> <p><i>Interaktionstest: p=0,09:</i> Männlich: OR: 0,20 [0,06;0,71], p=0,0125 Weiblich: OR: 4,16 [0,16;105,81], p=0,3882</p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Anhaltspunkt</p>
Erkrankungen des Gastrointestinaltrakts (SOC)	<p><u>Woche 48:</u> OR: 0,88 [0,66;1,19], p=0,4103 53,5% vs. 56,6%</p> <p><i>Interaktionstest: p=0,01</i> Kaukasisch: OR: 1,12 [0,79;1,58], p=0,5288 Nicht-kaukasisch: OR: 0,46 [0,25;0,84], p=0,0113</p> <p><u>Woche 96:</u> OR: 0,79 [0,58;1,07], p=0,1279 58,1% vs. 63,7%</p> <p><i>Interaktionstest: p=0,02</i> Kaukasisch: OR: 1,00 [0,70;1,42], p=0,9823 Nicht-kaukasisch: OR: 0,42 [0,23;0,78], p=0,0057</p> <p><i>Interaktionstest: p=0,13</i> ≤100.000 Kopien/mL: OR: 0,65 [0,44;0,97], p=0,0328 >100.000 Kopien/mL: OR: 1,04 [0,65;1,67], p=0,8599</p> <p><i>Interaktionstest: p=0,07</i> ≤350 Zellen/μL:</p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Nicht-kaukasische Patienten:</u> Ausmaß des Zusatznutzen/Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis</p> <p><u>≤100.000 Kopien/mL:</u> Ausmaß des Zusatznutzen/Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Anhaltspunkt</p> <p><u>>350 Zellen/μL:</u> Ausmaß des Zusatznutzen/Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Anhaltspunkt</p>

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
	OR: 1,06 [0,68;1,63], p=0,8029 >350 Zellen/µL: OR: 0,61 [0,40;0,92], p=0,0202	
Diarrhoe	<u>Woche 48:</u> OR: 0,74 [0,53;1,05], p=0,0891 21,8% vs. 27,3% <i>Interaktionstest: p=0,15</i> ≤100.000 Kopien/mL: OR: 0,59 [0,37;0,94], p=0,0257 >100.000 Kopien/mL: OR: 0,99 [0,59;1,67], p=0,9627 <i>Interaktionstest: p=0,01</i> ≤350 Zellen/µL: OR: 1,18 [0,71;1,95], p=0,5280 >350 Zellen/µL: OR: 0,49 [0,30;0,80], p=0,0042 <u>Woche 96:</u> OR: 0,73 [0,53;1,01], p=0,0611 24,9% vs. 31,3% <i>Interaktionstest: p=0,18</i> ≤350 Zellen/µL: OR: 0,93 [0,57;1,49], p=0,7546 >350 Zellen/µL: OR: 0,59 [0,37;0,93], p=0,0240	Endpunktkategorie: Nicht- schwerwiegende Nebenwirkungen <u>>350 Zellen/µL:</u> Ausmaß des Zusatznutzen/Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Hinweis
Übelkeit	<u>Woche 48:</u> OR: 1,03 [0,71;1,49], p=0,8952 19,8% vs. 19,4% <u>Woche 96:</u> OR: 0,99 [0,69;1,42], p=0,9574 21,0% vs. 21,1%	Kein Zusatznutzen oder Schaden ableitbar
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)	<u>Woche 48:</u> OR: 1,04 [0,74;1,45], p=0,8297 27,2% vs. 26,5% <u>Woche 96:</u> OR: 0,81 [0,59;1,11], p=0,1844 28,9% vs. 33,5%	Kein Zusatznutzen oder Schaden ableitbar

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
<i>Fatigue</i>	<u>Woche 48:</u> OR: 1,14 [0,74;1,75], p=0,5615 14,2% vs. 12,7% <u>Woche 96:</u> OR: 0,92 [0,61;1,39], p=0,7020 15,0% vs. 16,1%	Kein Zusatznutzen oder Schaden ableitbar
Erkrankungen des Nervensystems (SOC)	<u>Woche 48:</u> OR: 0,98 [0,70;1,37], p=0,8991 25,8% vs. 26,2% <u>Woche 96:</u> OR: 0,91 [0,66;1,25], p=0,5565 31,7% vs. 33,8%	Kein Zusatznutzen oder Schaden ableitbar
<i>Kopfschmerzen</i>	<u>Woche 48:</u> OR: 1,25 [0,81;1,92], p=0,3114 15,0% vs. 12,4% <u>Woche 96:</u> OR: 1,14 [0,76;1,71], p=0,5156 16,7% vs. 14,9%	Kein Zusatznutzen oder Schaden ableitbar
Psychiatrische Erkrankungen (SOC)	<u>Woche 48:</u> OR: 0,91 [0,64;1,30], p=0,6143 21,2% vs. 22,8% <i>Interaktionstest: p=0,004</i> Männlich: OR: 0,76 [0,53;1,11], p=0,1597 Weiblich: OR: 5,35 [1,49;19,19], p=0,0101 <i>Interaktionstest: p=0,01</i> Therapieadhärenz <95%: OR: 0,42 [0,21;0,84], p=0,0147 Therapieadhärenz ≥95%: OR: 1,19 [0,78;1,82], p=0,4096 <u>Woche 96:</u> OR: 0,90 [0,66;1,25], p=0,5423 28,6% vs. 30,7% <i>Interaktionstest: p=0,009</i>	Kein Zusatznutzen oder Schaden ableitbar

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
	Männlich: OR: 0,78 [0,56;1,10], p=0,1581 Weiblich: OR: 4,16 [1,25;13,82], p=0,0202	
<i>Abnorme Träume</i>	<u>Woche 48:</u> OR: 0,86 [0,39;1,88], p=0,7005 3,4% vs. 3,9%	Kein Zusatznutzen oder Schaden ableitbar
<i>Schlaflosigkeit</i>	<u>Woche 48:</u> OR: 1,68 [0,91;3,08], p=0,0958 8,2% vs. 5,1%	Kein Zusatznutzen oder Schaden ableitbar
<i>Depression</i>	<u>Woche 48:</u> OR: 0,87 [0,47;1,61], p=0,6508 5,7% vs. 6,5% <u>Woche 96:</u> OR: 0,82 [0,50;1,32], p=0,4077 9,6% vs. 11,5%	Kein Zusatznutzen oder Schaden ableitbar
Erkrankungen der Haut und des Unterhautzellgewebes (SOC)	<u>Woche 48:</u> OR: 0,95 [0,69;1,32], p=0,7743 27,8% vs. 28,8% <u>Woche 96:</u> OR: 0,82 [0,60;1,11], p=0,2008 33,1% vs. 37,7%	Kein Zusatznutzen oder Schaden ableitbar
<i>Hautausschlag</i>	<u>Woche 48:</u> OR: 1,15 [0,64;2,09], p=0,6365 7,1% vs. 6,2% <u>Woche 96:</u> OR: 0,94 [0,55;1,59], p=0,8049 8,2% vs. 8,7%	Kein Zusatznutzen oder Schaden ableitbar
Augenerkrankungen (SOC)	<u>Woche 48:</u> OR: 0,19 [0,11;0,33], p<0,00001 5,1% vs. 22,0% <i>Interaktionstest: p=0,02</i> <40 Jahre: OR: 0,11 [0,05;0,23], p<0,00001 ≥40 Jahre: OR: 0,42 [0,18;0,93], p=0,0334	Endpunktkategorie: Nicht- schwerwiegende Nebenwirkungen <u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
	<p><i>Interaktionstest: p=0,15</i></p> <p>Kaukasisch: OR: 0,24 [0,13;0,44], p<0,00001</p> <p>Nicht-kaukasisch: OR: 0,09 [0,02;0,31], p=0,0001</p> <p><i>Interaktionstest: p=0,16</i></p> <p>≤100.000 Kopien/mL: OR: 0,13 [0,06;0,29], p<0,00001</p> <p>>100.000 Kopien/mL: OR: 0,28 [0,14;0,58], p=0,0007</p> <p><i>Interaktionstest: p=0,18</i></p> <p>Therapieadhärenz <95%: OR: 0,08 [0,02;0,35], p=0,0009</p> <p>Therapieadhärenz ≥95%: OR: 0,23 [0,13;0,42], p<0,00001</p> <p><u>Woche 96:</u> OR: 0,27 [0,17;0,43], p<0,00001 7,9% vs. 23,9%</p> <p><i>Interaktionstest: p=0,15</i></p> <p><40 Jahre: OR: 0,21 [0,11;0,37], p<0,00001</p> <p>≥40 Jahre: OR: 0,41 [0,20;0,83], p=0,0138</p> <p><i>Interaktionstest: p=0,04</i></p> <p>Kaukasisch: OR: 0,36 [0,22;0,60], p=0,0001</p> <p>Nicht-kaukasisch: OR: 0,09 [0,02;0,31], p=0,0001</p> <p><i>Interaktionstest: p=0,14</i></p> <p>Therapieadhärenz <95%: OR: 0,13 [0,04;0,40], p=0,0003</p> <p>Therapieadhärenz ≥95%: OR: 0,33 [0,20;0,55], p<0,0001</p>	
Skleraler Ikterus	<p><u>Woche 48:</u> OR: 0,03 [0,01;0,14], p<0,00001 0,6% vs. 14,4%</p>	<p>Endpunktkategorie: Nicht- schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u></p>

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. ATV/r+FTC/TDF	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
	<u>Woche 96:</u> OR: 0,03 [0,01;0,14], p<0,00001 0,6% vs. 14,4%	Ausmaß des Zusatznutzen/Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis
Fett gedruckte Effektschätzer basieren auf einer Meta-Analyse der Studien S102 und S104 KI = Konfidenzintervall, MWD = Mittelwertdifferenz, OR = Odds Ratio, UE = unerwünschtes Ereignis, SUE = schwerwiegendes unerwünschtes Ereignis		

Ableitung des Zusatznutzens auf Endpunktebene

Virologisches Ansprechen

Zur Bewertung des Zusatznutzens wurde der Endpunkt virologisches Ansprechen mittels zweier Imputationsalgorithmen der FDA, Snapshot und TLOVR, untersucht. Unter Verwendung des Snapshot-Algorithmus besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,30 [0,82;2,06], p=0,22). Allerdings zeigt die Studie S103 die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze. Dies trifft sowohl auf die ITT- als auch die PP-Population zu. Es besteht zu Woche 96 kein statistischer Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,08 [0,73; 1,59], p=0,70), jedoch kann auch hier, in der ITT- und der PP-Population, die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze gezeigt werden.

Hinsichtlich des virologischen Ansprechens, erfasst mittels TLOVR, zeigen sich ähnliche Ergebnisse: Der Unterschied zwischen den Behandlungen zu Woche 48 ist nicht statistisch signifikant (OR [95%-KI]: 1,11 [0,73;1,69], p=0,55). Die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze wurde jedoch gezeigt. Auch zu Woche 96 besteht die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze; ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht jedoch nicht (OR [95%-KI]: 1,27 [0,88;1,82], p=0,19). Hier besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Therapieadhärenz (p=0,06; I²=72,7%). Da Subgruppenanalysen, stratifiziert nach Therapieadhärenz, allgemein als nicht aussagekräftig bewertet werden, ergibt sich hieraus keine Konsequenz für die Ableitung eines Zusatznutzens.

Zusammenfassend lässt sich feststellen, dass hinsichtlich des virologischen Ansprechens der Nutzen von EVG/COBI/FTC/TDF äquivalent zu dem von ATV/r+FTC/TDF ist. Zusätzlich dazu konnte die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber ATV/r+FTC/TDF zur 12%-Grenze sowohl zu Woche 48 als auch zu Woche 96 gezeigt werden. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Zeit bis zum Verlust des virologischen Ansprechens

Weder zu Woche 48 (HR [95%-KI]: 0,87 [0,59;1,27], $p=0,48$) noch zu Woche 96 (HR [95%-KI]: 0,81 [0,58;1,12], $p=0,21$) kann hinsichtlich der Zeit bis zum Verlust des virologischen Ansprechens ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen festgestellt werden. Somit besteht für EVG/COBI/FTC/TDF hinsichtlich der Zeit bis zum Verlust des virologischen Ansprechens ein äquivalenter Nutzen wie für ATV/r+FTC/TDF. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Es besteht jedoch ein Beleg für Effektmodifikation durch Therapieadhärenz ($p=0,04$; $I^2=75,2\%$). Da Subgruppenanalysen, stratifiziert nach Therapieadhärenz, allgemein als nicht aussagekräftig bewertet werden, ergibt sich hieraus keine Konsequenz für die Ableitung eines Zusatznutzens.

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Veränderung der CD4-Zellzahl

Weder zu Woche 48 (MWD [95%-KI]: -6 [-31;18], $p=0,61$) noch zu Woche 96 (MWD [95%-KI]: -8 [-35;19], $p=0,56$) zeigt sich ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen.

Zu Woche 96 besteht jedoch ein Beleg für Effektmodifikation durch Geschlecht ($p=0,04$; $I^2=76,6\%$). Klinische Daten zeigen, dass es weder hinsichtlich der virologischen noch hinsichtlich der immunologischen Response einen geschlechtsspezifischen Unterschied gibt [128]. Zudem hat bereits der G-BA hinsichtlich der Response festgestellt, dass „...vor dem Hintergrund der bisherigen Erfahrungen in der HIV-Therapie ein geschlechtsspezifischer Unterschied virusbiologisch nicht plausibel und medizinisch nicht begründet erscheint...“, hat dieser Beleg für Effektmodifikation [129] keine Auswirkungen auf das Gesamtfazit zum Zusatznutzen. Außerdem gibt es zu Woche 96 einen Beleg für Effektmodifikation durch CD4-Zellzahl zu Baseline ($p=0,02$; $I^2=81,5\%$). Demnach besteht in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline von ≤ 350 Zellen/ μL eine MWD [95%-KI] von -40,00 [-77,16;-2,84] mit $p=0,0350$.

Zusammenfassend lässt sich hier, da der Beleg für Effektmodifikation lediglich zu Woche 96 besteht, ein Anhaltspunkt für einen geringen Schaden für die Patienten, die zu Baseline eine CD4-Zellzahl ≤ 350 Zellen/ μL aufweisen, ableiten. Die klinische Relevanz des Schadens ist von untergeordneter Bedeutung, da der für das Risiko der Entwicklung Aids-definierender Erkrankungen relevante Bereich bei < 200 Zellen/ μL liegt [130, 131] – die CD4-Zellzahl zu Woche 96 lag in beiden Behandlungsarmen jedoch bereits deutlich über 200 CD4-Zellen/ μL (EVG/COBI/FTC/TDF: 488 CD4-Zellen/ μL ; ATV/r+FTC/TDF: 539 CD4-Zellen/ μL). Von klinischem Interesse ist zu diesem Zeitpunkt vielmehr die absolute Veränderung der CD4-Zellzahl versus Baseline, weniger jedoch der Behandlungsunterschied hinsichtlich der Veränderung der CD4-Zellzahl bei einem sowieso guten Immunstatus.

Mortalität

Hinsichtlich der Mortalität besteht weder zu Woche 48 (OR [95%-KI]: 0,14 [0,01;2,77], $p=0,1980$) noch zu Woche 96 (OR [95%-KI]: 0,14 [0,01;2,77], $p=0,1980$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Gesamtrate unerwünschter Ereignisse

Hinsichtlich der Gesamtrate unerwünschter Ereignisse besteht weder zu Woche 48 (OR [95%-KI]: 0,71 [0,40;1,26], $p=0,2423$) noch zu Woche 96 (OR [95%-KI]: 0,69 [0,33;1,47], $p=0,3376$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Gesamtrate schwerwiegender unerwünschter Ereignisse

Hinsichtlich der Gesamtrate schwerwiegender unerwünschter Ereignisse besteht weder zu Woche 48 (OR [95%-KI]: 0,83 [0,48;1,43], $p=0,5043$) noch zu Woche 96 (OR [95%-KI]: 0,65 [0,41;1,03], $p=0,0684$) ein statistisch signifikanter Unterschied zwischen den Behandlungen.

Allerdings besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,19$; $I^2=41,5\%$). Demnach besteht in der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL, ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Das OR [95%-KI] in dieser Gruppe beträgt 0,45 [0,21;0,94], $p=0,0325$.

Zusammenfassend lässt sich hier für Patienten mit Ausgangsviruslast von >100.000 Kopien/mL ein Anhaltspunkt für einen geringen Zusatznutzen ableiten.

Therapieabbrüche aufgrund unerwünschter Ereignisse

Hinsichtlich der Therapieabbrüche aufgrund unerwünschter Ereignisse besteht weder zu Woche 48 (OR [95%-KI]: 0,72 [0,35;1,48], $p=0,3689$) noch zu Woche 96 (OR [95%-KI]: 0,71 [0,36;1,39], $p=0,3151$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Unerwünschte Ereignisse Grad 3-4

Es bestehen zu Woche 48 (OR [95%-KI]: 0,93 [0,60;1,45], $p=0,7607$) und 96 (OR [95%-KI]: 0,89 [0,60;1,33], $p=0,5648$) keine statistisch signifikanten Unterschiede zwischen den Behandlungen.

Zu beiden Zeitpunkten bestehen jedoch Belege für Effektmodifikation durch Ausgangsviruslast (Woche 48: $p=0,0002$; $I^2=92,9\%$, Woche 96: $p=0,002$; $I^2=90,1\%$). Demnach besteht in der Gruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL zu beiden Zeitpunkten ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: Woche 48: 0,32 [0,15;0,67], $p=0,0025$; Woche 96: 0,37 [0,18;0,74], $p=0,0050$). Außerdem besteht zu Woche 48 in der Gruppe der Patienten mit

Ausgangsviruslast ≤ 100.000 Kopien/mL ein statistisch signifikanter Unterschied zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,95 [1,08;3,52], $p=0,0262$).

Zusammenfassend lässt sich hier für die Gruppe von Patienten mit Ausgangsviruslast >100.000 Kopien/mL ein Hinweis für einen erheblichen Zusatznutzen von EVG/COBI/FTC/TDF ableiten. Es besteht außerdem ein Anhaltspunkt für einen geringen Schaden durch EVG/COBI/FTC/TDF in der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL.

Während sich in der Subgruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL bei den UE Grad 3-4 ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF zeigt, findet sich in der Subgruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL ein ebenfalls statistisch signifikanter Unterschied zuungunsten von EVG/COBI/FTC/TDF.

Berücksichtigt man jedoch die Wahrscheinlichkeit und das auf der Effektstärke basierende Ausmaß des sich ergebenden Zusatznutzens, so zeigt sich in der Subgruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL auf Basis der Konfidenzintervalle und Messzeitpunkte (Woche 48 und Woche 96) ein Hinweis für einen erheblichen Zusatznutzen. Der statistisch signifikante Unterschied bei den Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL hingegen wurde lediglich zu Woche 48 gefunden, zu Woche 96 zeigte sich kein Unterschied zwischen den beiden Armen. Entsprechend ist für die Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL hier lediglich ein Anhaltspunkt für einen geringen Zusatznutzen anzuleiten.

Da sich der gefundene Effekt medizinisch nicht schlüssig begründen lässt, und sich hier maßgebliche Unterschiede hinsichtlich Effektstärke, Wahrscheinlichkeit und Ausmaß des Zusatznutzens zeigen, wird davon ausgegangen, dass dem gefundenen Hinweis auf einen erheblichen Zusatznutzen bei Patienten mit Ausgangsviruslast >100.000 Kopien/mL (statistische Signifikanz zu Woche 48 und 96) ein ungleich höheres Gewicht zuzumessen ist als dem Anhaltspunkt für einen geringen Zusatznutzen bei Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL.

Neurologische & psychiatrische Ereignisse

Hinsichtlich neurologischer & psychiatrischer Ereignisse besteht weder zu Woche 48 (OR [95%-KI]: 0,94 [0,69;1,27], $p=0,69$) noch zu Woche 96 (OR [95%-KI]: 0,95 [0,71;1,28], $p=0,76$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Hautausschläge

Hinsichtlich dem Auftreten von Hautausschlägen besteht weder zu Woche 48 (OR [95%-KI]: 0,99 [0,67;1,45], $p=1,00$) noch zu Woche 96 (OR [95%-KI]: 0,93 [0,65;1,32], $p=0,72$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Renale Ereignisse

Hinsichtlich renalere Ereignisse besteht weder zu Woche 48 (OR [95%-KI] nicht berechenbar) noch zu Woche 96 (OR [95%-KI]: 9,15 [0,49;170,67], $p=0,061$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Knochenfrakturen

Zu Woche 48 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,50 [0,12;2,01], $p=0,51$). Zu Woche 96 kann jedoch ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF festgestellt werden (OR [95%-KI]: 0,28 [0,09;0,86], $p=0,029$).

Es besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Geschlecht ($p=0,09$; $I^2=65,7\%$). In der Gruppe der männlichen Patienten ergibt sich ein OR [95%-KI] von 0,20 [0,06;0,71] mit $p=0,0125$, was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF anzeigt, und in der Gruppe der Frauen beträgt das OR [95%-KI] 4,16 [0,16;105,81] mit $p=0,3882$. Da in der Gruppe der Frauen nur sehr wenig Patienten eingeschlossen sind und lediglich ein Patient eine Knochenfraktur erlitt, ist diese Interaktion mit Vorsicht zu interpretieren.

Pathologische Frakturen stellen die symptomatische Endstrecke Knochenmineraldichte-reduzierender Zustandsbilder wie Osteopenie und Osteoporose dar. Hierbei handelt es je nach Genese um eine Erhöhung osteoklastischer Aktivität oder einer Verminderung osteoblastischer Aktivität mit entsprechender mechanisch-inadäquater Mineralfineinstruktur des Knochens. Eine initiale leichtgradige Reduktion der Knochenmineraldichte mit Stabilisierung nach 24 oder 48 Wochen, je nach Messpunkt ist nach Einleitung einer antiretroviralen HIV Therapie für nahezu alle Therapieregime auch in kontrollierten klinischen Studien beschrieben. Ursächlich hierfür werden immunrekonstitutive Prozesse im Rahmen des Wiederanstiegs der CD4-Zellzahl beschrieben. Die beschriebenen Reduktionen bewegen sich im niedrigen einstelligen Prozentbereich und bewegen sich nicht in Dimensionen, in denen pathologische Frakturen zum Beispiel im Rahmen der postmenstruellen Osteoporose oder maligner Zustandsbilder regelmäßig beschrieben werden.

Der beschriebene Unterschied zwischen EVG/COBI/FTC/TDF und ATV/r+FTC/TDF nach 96 Wochen ist weder mit der o.g. Pathogenese noch aufgrund des kurzen Beobachtungszeitraumes schlüssig erklärbar. Eine für Frakturen ursächliche Knochendichtereduktion über 96 Wochen hätte zudem einen erheblichen Anstieg entsprechender Serumparameter, wie zum Beispiel der alkalischen Phosphatase und oder der Lactat-Dehydrogenase verursachen müssen, der in der Studie aber nicht beobachtet werden konnte.

Zusammenfassend lässt sich hier für EVG/COBI/FTC/TDF, da der statistisch signifikante Unterschied lediglich zu Woche 96 besteht, ein Anhaltspunkt für einen geringen Zusatznutzen in Gesamtpopulation feststellen.

Erkrankungen des Gastrointestinaltrakts (SOC)

Weder zu Woche 48 (OR [95%-KI]: 0,88 [0,66;1,19], $p=0,4103$) noch zu Woche 96 (OR [95%-KI]: 0,79 [0,58;1,07], $p=0,1279$) kann in der Gesamtpopulation ein statistisch signifikanter Unterschied zwischen den Behandlungen festgestellt werden.

Zu Woche 48 und 96 besteht jeweils ein Beleg für Effektmodifikation durch Rasse (Woche 48: $p=0,01$; $I^2=84,1\%$, Woche 96: $p=0,02$; $I^2=82,5\%$). Demnach besteht in der nicht-kaucasischen Population jeweils ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: Woche 48: 0,46 [0,25;0,84], $p=0,0113$; Woche 96: 0,42 [0,23;0,78], $p=0,0057$). Außerdem bestehen zu Woche 96 Hinweise auf Effektmodifikation durch Ausgangsviruslast ($p=0,13$; $I^2=55,8\%$) und CD4-Zellzahl zu Baseline ($p=0,07$; $I^2=69,2\%$). Es besteht in der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,65 [0,44;0,97], $p=0,0328$). Außerdem gibt es in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,61 [0,40;0,92], $p=0,0202$).

Der in der nicht-kaucasischen Subgruppe gefundene Unterschied bei den Erkrankungen des Gastrointestinaltrakts (SOC) zugunsten von EVG/COBI/FTC/TDF lässt sich insbesondere auch auf Basis der vorliegenden Evidenz (eher höhere Nebenwirkungsraten unter Nicht-Kaukasiern) nicht schlüssig begründen und wird für die Nutzenableitung nicht weiter berücksichtigt.

Diarrhö

Zu keinem der hier betrachteten Zeitpunkte besteht, trotz numerischem Vorteil von EVG/COBI/FTC/TDF zu beiden Zeitpunkten, ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: Woche 48: 0,74 [0,53;1,05], $p=0,0891$; Woche 96: 0,73 [0,53;1,01], $p=0,0611$).

Zu Woche 48 gibt es einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ($p=0,15$; $I^2=51,9\%$). Demnach besteht in der Gruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,59 [0,37;0,94], $p=0,0257$). Außerdem besteht zu Woche 48 ($p=0,01$; $I^2=83,2\%$) ein Hinweis und zu Woche 96 ($p=0,18$; $I^2=44,3\%$) ein Beleg für Effektmodifikation durch CD4-Zellzahl zu Baseline. Das OR [95%-KI] in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline >350 Zellen/ μL beträgt dabei zu Woche 48 0,49 [0,30;0,80] und $p=0,0042$ und zu Woche 96 0,59 [0,37;0,93], $p=0,0240$.

Zusammenfassend lässt sich hier ein Hinweis auf einen geringen Zusatznutzen durch EVG/COBI/FTC/TDF für Patienten mit CD4-Zellzahl zu Baseline von >350 Zellen/ μL ableiten.

Übelkeit

Hinsichtlich des Auftretens von Übelkeit besteht weder zu Woche 48 (OR [95%-KI]: 1,03 [0,71;1,49], $p=0,8952$) noch zu Woche 96 (OR [95%-KI]: 0,99 [0,69;1,42], $p=0,9574$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)

Hinsichtlich allgemeiner Erkrankungen und Beschwerden am Verabreichungsort (SOC) besteht weder zu Woche 48 (OR [95%-KI]: 1,04 [0,74;1,45], $p=0,8297$) noch zu Woche 96 (OR [95%-KI]: 0,81 [0,59;1,11], $p=0,1844$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Fatigue

Hinsichtlich des Auftretens von Fatigue besteht weder zu Woche 48 (OR [95%-KI]: 1,14 [0,74;1,75], $p=0,5615$) noch zu Woche 96 (OR [95%-KI]: 0,92 [0,61;1,39], $p=0,7020$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Erkrankungen des Nervensystems (SOC)

Hinsichtlich Erkrankungen des Nervensystems (SOC) besteht weder zu Woche 48 (OR [95%-KI]: 0,98 [0,70;1,37], $p=0,8991$) noch zu Woche 96 (OR [95%-KI]: 0,91 [0,66;1,25], $p=0,5565$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Kopfschmerzen

Hinsichtlich des Auftretens von Kopfschmerzen besteht weder zu Woche 48 (OR [95%-KI]: 1,25 [0,81;1,92], $p=0,3114$) noch zu Woche 96 (OR [95%-KI]: 1,14 [0,76;1,71], $p=0,5156$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Psychiatrische Erkrankungen (SOC)

Hinsichtlich psychiatrischer Erkrankungen (SOC) besteht weder zu Woche 48 (OR [95%-KI]: 0,91 [0,64;1,30], $p=0,6143$) noch zu Woche 96 (OR [95%-KI]: 0,90 [0,66;1,25], $p=0,5423$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Es besteht jedoch zu Woche 48 und 96 jeweils ein Beleg für Effektmodifikation durch Geschlecht (Woche 48: $p=0,004$; $I^2=87,8\%$, Woche 96: $p=0,009$; $I^2=85,4\%$). Aufgrund der geringen Fallzahl der weiblichen Studienteilnehmer sowie der hohen Zahl an möglichen fälschlicherweise als signifikant eingestuften Interaktionen, ist die hier gefundene Interaktion mit Vorsicht zu beurteilen. Außerdem besteht ein Beleg für Effektmodifikation zu Woche 48 durch Therapieadhärenz ($p=0,01$; $I^2=84,2\%$). Da Subgruppenanalysen, stratifiziert nach Therapieadhärenz, allgemein als nicht aussagekräftig bewertet werden, ergibt sich hieraus keine Konsequenz für die Ableitung eines Zusatznutzens.

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Abnorme Träume

Hinsichtlich abnormer Träume besteht zu Woche 48 (OR [95%-KI]: 0,86 [0,39;1,88], $p=0,7005$) kein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Schlaflosigkeit

Hinsichtlich des Auftretens von Schlaflosigkeit besteht zu Woche 48 (OR [95%-KI]: 1,68 [0,91;3,08], $p=0,0958$) kein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Depression

Hinsichtlich des Auftretens von Depressionen besteht weder zu Woche 48 (OR [95%-KI]: 0,87 [0,47;1,61], $p=0,6508$) noch zu Woche 96 (OR [95%-KI]: 0,82 [0,50;1,32], $p=0,4077$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Erkrankungen der Haut und des Unterhautzellgewebes (SOC)

Hinsichtlich Erkrankungen der Haut und des Unterhautzellgewebes besteht weder zu Woche 48 (OR [95%-KI]: 0,95 [0,69;1,32], $p=0,7743$) noch zu Woche 96 (OR [95%-KI]: 0,82 [0,60;1,11], $p=0,2008$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Hautausschlag

Hinsichtlich des Auftretens von Hautausschlag besteht weder zu Woche 48 (OR [95%-KI]: 1,15 [0,64;2,09], $p=0,6365$) noch zu Woche 96 (OR [95%-KI]: 0,94 [0,55;1,59], $p=0,8049$) ein statistisch signifikanter Unterschied zwischen den Behandlungen. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Augenerkrankungen (SOC)

Es besteht hinsichtlich des Auftretens von Augenerkrankungen sowohl zu Woche 48 (OR [95%-KI]: 0,19 [0,11;0,33], $p<0,00001$) als auch zu Woche 96 (OR [95%-KI]: 0,27 [0,17;0,43], $p<0,00001$) ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF. Unter einer Behandlung mit ECG/COBI/FTC/TDF traten somit deutlich weniger Augenerkrankungen auf als unter einer Behandlung mit ATV/r+FTC/TDF.

Es bestehen zu beiden Zeitpunkten Hinweise bzw. Belege auf Effektmodifikation (Woche 48: Alter [$p=0,02$; $I^2=81,9\%$], Rasse [$p=0,15$; $I^2=51,5\%$], Ausgangsviruslast [$p=0,16$; $I^2=50,1\%$] und Therapieadhärenz [$p=0,18$; $I^2=43,7\%$]; Woche 96: Alter [$p=0,15$; $I^2=51,6\%$], Rasse [$p=0,04$; $I^2=76,7\%$] und Therapieadhärenz [$p=0,14$; $I^2=53,1\%$]). Da dies jedoch ausschließlich quantitative Interaktionen sind, wirken sie sich nicht auf das Gesamtfazit zum Zusatznutzen aus.

Zusammenfassend lässt sich hier ein beträchtlicher Zusatznutzen durch EVG/COBI/FTC/TDF für die Gesamtpopulation ableiten. Und da dieser Zusatznutzen zu beiden Zeitpunkten besteht, wird er mit der Wahrscheinlichkeit eines Hinweises bewertet.

Skleraler Ikterus

Es besteht zu Woche 48 ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,03 [0,01;0,14], $p < 0,00001$). Und auch zu Woche 96 ist der Unterschied zwischen den Behandlungen statistisch signifikant zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,03 [0,01;0,14], $p < 0,00001$).

Zusammenfassend lässt sich hier Hinweis auf einen beträchtlichen Zusatznutzen für die Gesamtpopulation ableiten. Da dieser Endpunkt aber durch die SOC der Augenerkrankungen abgedeckt wird, wird dieser Endpunkt nicht zusätzlich in der Ableitung eines gesamten Zusatznutzens bzw. Schadens berücksichtigt.

Gesamtfazit zum Zusatznutzen

Ziel einer ART ist es die Viruslast dauerhaft unter die Nachweisgrenze von 50 Kopien/mL abzusenken und damit die infektionsbedingten Symptome zu unterdrücken, die Krankheitsprogression zu vermindern und eine Immunrekonstitution zu erreichen [2]. Da Nebenwirkungen – insbesondere dabei Nebenwirkungen, die den HIV-Infizierten hinsichtlich seiner Lebensqualität und/oder im Hinblick auf eine mögliche Stigmatisierung, maßgeblich beeinträchtigen - einen erheblichen Einfluss auf die dringend notwendige hohe Therapieadhärenz, und damit auch auf das Risiko von Resistenzen und virologischem Versagen, haben, ist eine Minimierung der Nebenwirkungen einer ART von immenser Wichtigkeit für den dauerhaften Therapieerfolg [4, 49, 50].

In der vorliegenden Nutzenbewertung wurde anhand von patientenrelevanten und validen Endpunkten gezeigt, dass EVG/COBI/FTC/TDF, bei gleichwertiger Wirksamkeit wie ATV/r+FTC/TDF (Nichtunterlegenheit zur 12%-Grenze wurde gezeigt), eine sicherere Behandlungsoption darstellt als die zweckmäßige Vergleichstherapie ATV/r+FTC/TDF. Durch EVG/COBI/FTC/TDF kann der durch die zweckmäßige Vergleichstherapie ATV/r+FTC/TDF ausgelöste patientenrelevante Schaden vermindert werden.

Es konnte gezeigt werden, dass EVG/COBI/FTC/TDF eine gleichwertige bzw. nichtunterlegene Wirksamkeit wie ATV/r+FTC/TDF besitzt. Dies zeigt sich anhand der Endpunkte virologisches Ansprechen, Zeit bis zum Verlust des virologischen Ansprechens und Veränderung der CD4-Zellzahl.

Bei sicherheitsbezogenen Endpunkten zeigten sich unter EVG/COBI/FTC/TDF Vorteile hinsichtlich der Vermeidung von schwerwiegenden unerwünschten Ereignissen, unerwünschten Ereignissen Grad 3-4, Knochenfrakturen, Erkrankungen des Gastrointestinaltrakts (darunter Diarrhoe) und Augenerkrankungen (darunter skleraler Ikterus). Bei der Ableitung eines Gesamtfazits zum Zusatznutzen werden die Gruppen von

PTs (neurologische & psychiatrische Ereignisse und Hautausschläge) nicht als Zusatzinformation betrachtet, da sie durch die standardisierten SOC Erkrankungen des Nervensystem und Psychiatrische Erkrankungen abgedeckt werden. Sie tragen somit kein Gewicht zur Ermittlung eines gesamten Zusatznutzens bei.

Bei der Verträglichkeit ist nicht nur der rein medizinische Aspekt von Bedeutung, sondern bei bestimmten Nebenwirkungen im gleichen Ausmaß auch der psychologische: Die für den Wirkstoff ATV typische Nebenwirkung skleraler Ikterus stellt eine morphologische Veränderung und damit einen kaum zu verbergenden, äußerlich sichtbaren Effekt dar, der zu Fragen hinsichtlich der Ursachen und somit zur unbeabsichtigten Bekanntgabe einer HIV-Infektion führen kann. Dies ist besonders in der homosexuellen Community, die gut vernetzt und zum Thema HIV gut informiert ist, von hoher Relevanz. Da unerwünschte Ereignisse wie ein skleraler Ikterus – zumal dieser sich nicht verbergen lässt - im Kreis der mit HIV-infizierten Bevölkerung direkt stigmatisierend wirken können, ist es von besonderem Interesse, Ereignisse dieser Art weitgehend zu vermeiden. Die Stigmatisierung eines HIV-infizierten Patienten kann auch weitere Erkrankungen, beispielsweise Depressionen, zur Folge haben [132]. Gerade für Depressionen wiederum konnte in vielen Studien gezeigt werden, dass diese einen negativen Einfluss auf die Adhärenz haben, was in der Folge zu einem Wiederanstieg der Viruslast mit konsekutivem virologischen Therapieversagen, immunologischer Verschlechterung und erhöhter Morbidität und Mortalität führen kann [36, 37]. Ebenso kann eine solche morphologische Veränderung direkt zu einer verminderten Adhärenz und sogar zu einer Beendigung der Therapie von Seiten des Patienten führen. Dieses Risiko ist bekannt durch die Lipodystrophie, eine Fettumverteilungsstörung, die von einigen, mittlerweile obsoleten antiretroviralen Substanzen verursacht wird [133]: Die Effekte der Lipodystrophie sind sehr gut untersucht und die negativen Auswirkungen dieses körperlichen Effekts der ART lassen sich im Wesentlichen auch auf andere sichtbare körperliche Veränderungen wie den skleralen Ikterus übertragen. Dem Stigma folgt oftmals Diskriminierung [134] - Stigma und Diskriminierung in Bezug auf eine HIV-Infektion können viele Aspekte des täglichen Lebens negativ beeinflussen und auch in die soziale Isolation führen [135-137]. Die Folgen können sowohl beruflich als auch privat zum Tragen kommen [135-139]. Körperliche Veränderungen durch HIV bzw. die ART können psychologische Effekte auf die betroffene Person haben, zu der neben Depressionen auch eine Beeinträchtigung des Selbstbildes und des Selbstwertgefühls gehören [140, 141].

Insgesamt haben beeinträchtigende Nebenwirkungen generell einen negativen Einfluss auf die Therapieadhärenz, was das Risiko der Bildung von Resistenzen und virologischem Versagen mit konsekutiver klinischer Progression erhöht. Da jedoch gerade in der Behandlung der HIV-Infektion eine hohe Therapieadhärenz unabdingbar für einen Behandlungserfolg ist und einen direkten Einfluss auf Morbidität und Mortalität haben, spielt die Vermeidung von unerwünschten Ereignissen, welche die Therapieadhärenz negativ beeinflussen könnten, eine zentrale Rolle [4, 49, 50].

Der hier beobachtete Vorteil durch EVG/COBI/FTC/TDF hinsichtlich der Nebenwirkungen im Allgemeinen und Augenerkrankungen im Besonderen ist als weitestgehende Vermeidung

solch stigmatisierender Nebenwirkungen zu verstehen. Dies hat nicht nur einen Einfluss auf die Lebensqualität der Patienten, sondern auch auf den Behandlungserfolg. Allgemein sinkt durch die bessere Verträglichkeit insgesamt das Risiko eines durch den Patienten gewünschten Therapiewechsels, wodurch die Dauerhaftigkeit des Erstregimes positiv beeinflusst wird. Dies ist angesichts einer Zahl von 25-30% der Patienten, die im ersten Jahr ihre ART wechseln, ein erstrebenswertes Ziel [43, 44]. Das langfristige Verbleiben auf einem Regime erhält die bei einem nötigen Therapiewechsel weiterhin zur Verfügung stehenden Behandlungsoptionen. Frühe und häufige Therapiewechsel schränken die weiteren Behandlungsoptionen immer weiter ein und zudem erhöhen sich Komplexität und Kosten der nachfolgend eingesetzten Regime, was wiederum das Risiko der Nicht-Adhärenz steigert. Des Weiteren steigt mit ansteigender Viruslast auch das Risiko für die Transmission des Virus, wobei eine ART durch die Übertragung resistenter Viren von Anfang an komplizierter und im Zweifel auch teurer wird [4]. Darüber hinaus zeigen Daten, dass die Dauer der viralen Suppression einen Prädiktor für eine dauerhafte virale Suppression darstellt, d. h. mit der Dauer der viralen Suppression sinkt die Wahrscheinlichkeit für einen virologischen Rebound [46].

Neben den positiven sicherheitsbezogenen Aspekten konnten auch einige Nachteile von EVG/COBI/FTC/TDF beobachtet werden. Dies zeigte sich bei der Veränderung der CD4-Zellzahl bei Patienten mit CD4-Zellzahl bei Baseline ≤ 350 Zellen/ μL und unerwünschten Ereignissen Grad 3-4 bei Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL. Der Anhaltspunkt für einen geringen Schaden hinsichtlich der Veränderung der CD4-Zellzahl durch EVG/COBI/FTC/TDF besteht nur in der Subgruppe der Patienten, die zu Baseline eine CD4-Zellzahl ≤ 350 Zellen/ μL aufwiesen. Da dieser Schaden lediglich in einer Subgruppe besteht und auch nur mit der Wahrscheinlichkeit eines Anhaltspunkts bewertet wird, wirkt sich dies nicht auf die allgemeine Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF aus, die als äquivalent bzw. nichtunterlegen zu ATV/r+FTC/TDF eingeschätzt wird. Die klinische Relevanz des Schadens ist von untergeordneter Bedeutung, da der für das Risiko der Entwicklung Aids-definierender Erkrankungen relevante Bereich bei < 200 Zellen/ μL liegt [130, 131] – die CD4-Zellzahl zu Woche 96 lag in beiden Behandlungsarmen jedoch bereits deutlich über 200 CD4-Zellen/ μL (EVG/COBI/FTC/TDF: 488 CD4-Zellen/ μL ; ATV/r+FTC/TDF: 539 CD4-Zellen/ μL). Von klinischem Interesse ist zu diesem Zeitpunkt vielmehr die absolute Veränderung der CD4-Zellzahl versus Baseline, weniger jedoch der Behandlungsunterschied hinsichtlich der Veränderung der CD4-Zellzahl bei einem sowieso guten Immunstatus.

Dem Anhaltspunkt für einen geringen Schaden hinsichtlich der Veränderung der CD4-Zellzahl in der Subgruppe mit CD4-Zellzahl zu Baseline ≤ 350 Zellen/ μL stehen ein Anhaltspunkt für einen geringen Zusatznutzen in der Gruppe der Patienten mit Ausgangsviruslast > 100.000 Kopien/mL hinsichtlich der Vermeidung schwerwiegender unerwünschter Ereignisse, und darüber hinaus ein Anhaltspunkt für einen geringen Zusatznutzen hinsichtlich der Vermeidung gastrointestinaler Erkrankungen in der weniger stark erkrankten Population (Ausgangsviruslast ≤ 100.000 Kopien/mL bzw. CD4-Zellzahl zu Baseline > 350 Zellen/ μL) gegenüber. Da die CD4-Zellzahl zu Woche 96 generell oberhalb

des kritischen Bereichs von <200 Zellen/ μL (CD4-Zellzahl bei Patienten mit CD4-Zellzahl >350 Zellen/ μL zu Baseline zu Woche 96: 488 Zellen/ μL unter COBI/EVG/FTC/TDF versus 539 CD4-Zellen/ μL unter ATV/r+FTC/TDF) liegt, stellt die Verminderung schwerwiegender unerwünschter Ereignisse in der Patientengruppe mit Ausgangsviruslast >100.000 Kopien/mL sowie die Verminderung gastrointestinaler Ereignisse in der weniger stark erkrankten Population (Ausgangsviruslast ≤ 100.000 Kopien/mL bzw. CD4-Zellzahl zu Baseline >350 Zellen/ μL) einen wesentlich patientenrelevanteren Aspekt dar.

Der Anhaltspunkt für einen geringen Schaden hinsichtlich unerwünschter Ereignisse Grad 3-4 durch EVG/COBI/FTC/TDF in der Subgruppe der Patienten mit Ausgangsviruslast ≤ 100.000 Kopien/mL zu Woche 48 steht einem Hinweis auf einen erheblichen Zusatznutzen durch EVG/COBI/FTC/TDF bei diesem Endpunkt in der Subgruppe der Patienten mit Ausgangsviruslast >100.000 Kopien/mL zu Woche 48 und 96 gegenüber. Die Gegenläufigkeit der Effekte in den beiden Viruslast-Strata ist medizinisch nicht plausibel. Insgesamt hebt der geringe Schaden den erheblichen Zusatznutzen durch EVG/COBI/FTC/TDF nicht auf – im Gegenteil hat der Hinweis auf erheblichen Zusatznutzen sowohl im Hinblick auf Effektstärke, Wahrscheinlichkeit und Ausmaß ein deutlich höheres Gewicht.

In der Gesamtschau des auf Endpunktebene abgeleiteten Zusatznutzens (siehe Tabelle 4-193) verbleibt nach der Nutzen-Schaden-Abwägung ein positiver Zusatznutzen für EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF, dessen Ausmaß als **gering** eingestuft wird. Die Wahrscheinlichkeit dieses Zusatznutzens wird als **Hinweis** eingestuft.

Tabelle 4-193: Gegenüberstellung von Zusatznutzen und Schaden von EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF

Zusatznutzen durch EVG/COBI/FTC/TDF <u>Format:</u> Endpunkt Population: Wahrscheinlichkeit und Ausmaß des ZN	Schaden durch EVG/COBI/FTC/TDF <u>Format:</u> Endpunkt Population: Wahrscheinlichkeit und Ausmaß des Schadens
<ul style="list-style-type: none"> • Gesamtrate SUEs >100.000 Kopien/mL: Anhaltspunkt für geringen ZN • UEs Grad 3-4 >100.000 Kopien/mL: Hinweis auf erheblichen ZN • Knochenfrakturen Gesamt: Anhaltspunkt für geringen ZN • Erkrankungen des Gastrointestinaltrakts (SOC) Nicht-kaukasisch: Anhaltspunkt für beträchtlichen ZN ≤100.000 Kopien/mL: Anhaltspunkt für geringen ZN >350 Zellen/μL: Anhaltspunkt für geringen ZN <i>Diarrhoe</i> >350 Zellen/μL: Hinweis auf geringen ZN • Augenerkrankungen (SOC) Gesamt: Hinweis auf beträchtlichen ZN <i>Skleraler Ikterus</i> Gesamt: Hinweis auf beträchtlichen ZN 	<ul style="list-style-type: none"> • Veränderung der CD4-Zellzahl ≤350 Zellen/μL: Anhaltspunkt für geringen Schaden • UEs Grad 3-4 ≤100.000 Kopien/mL: Anhaltspunkt für geringen Schaden
ZN = Zusatznutzen	

EVG/COBI/FTC/TDF vs. RAL+FTC/TDF

Gegenüber RAL+FTC/TDF lässt sich der Zusatznutzen von EVG/COBI/FTC/TDF auf Endpunktebene sowie insgesamt wie folgt ableiten.

Es wurden zwei randomisierte aktiv-kontrollierte Studien, die EVG/COBI/FTC/TDF mit EFV/FTC/TDF verglichen, und eine randomisierte kontrollierte Studie, welche RAL+FTC/TDF mit EFV+FTC/TDF verglich, in die Nutzenbewertung eingeschlossen. Es wurden dazu folgende patientenrelevante Endpunkte untersucht:

- Virologisches Ansprechen, definiert als Abfall der Viruslast unter die Nachweisgrenze von 50 Kopien/mL
- Veränderung der CD4-Zellzahl
- Unerwünschte Ereignisse

- Mortalität
- Gesamtrate unerwünschter Ereignisse
- Gesamtrate schwerwiegender unerwünschter Ereignisse
- Therapieabbrüche aufgrund unerwünschter Ereignisse
- Häufigste unerwünschte Ereignisse

Zur Ableitung des Zusatznutzens werden Ergebnisse zu Woche 48 und Woche 96 herangezogen. In der folgenden Tabelle werden die Ergebnisse auf Endpunktebene und die als fazitrelevant eingestuft Hinweise bzw. Belege auf Effektmodifikation dargestellt. Basierend auf diesen Ergebnissen wird das Ausmaß des Zusatznutzens und dessen Wahrscheinlichkeit auf Endpunktebene, unter Angabe der Population, abgeleitet. Im Falle eines statistisch signifikanten Unterschieds zu lediglich einem der betrachteten Zeitpunkte (zum Beispiel zu Woche 48 statistisch signifikanter Unterschied, zu Woche 96 jedoch nicht mehr), wird die Wahrscheinlichkeit des Zusatznutzens für diesen Endpunkt herabgestuft.

Tabelle 4-194: Ausmaß des Zusatznutzen von EVG/COBI/FTC/TDF gegenüber RAL+FTC/TDF auf Endpunktebene

Endpunkt	Effektschätzer [95%-KI]	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
Virologisches Ansprechen (M=F)	<u>Woche 48:</u> OR: 0,98 [0,52;1,83] <u>Woche 96:</u> OR: 1,11 [0,61;1,99]	Kein Zusatznutzen oder Schaden ableitbar
Veränderung der CD4-Zellzahl	<u>Woche 48:</u> MWD: 10,26 [-21,84;42,36] <u>Woche 96:</u> MWD: 7 [-35,19;49,19]	Kein Zusatznutzen oder Schaden ableitbar
Unerwünschte Ereignisse		
Mortalität	<u>Woche 48:</u> OR: 0,10 [0,00;4,66] <u>Woche 96:</u> OR: 0,07 [0,00;3,10]	Kein Zusatznutzen oder Schaden ableitbar
Gesamtrate UEs	<u>Woche 48:</u> OR: 2,79 [1,08;7,20] <u>Woche 96:</u> OR: 2,00 [0,57;7,04]	Kein Zusatznutzen oder Schaden ableitbar

Endpunkt	Effektschätzer [95%-KI]	Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens
Gesamtrate SUEs	<u>Woche 48:</u> OR: 1,69 [0,79;3,59] <u>Woche 96:</u> OR: 1,53 [0,78;3,00]	Kein Zusatznutzen oder Schaden ableitbar
Therapieabbrüche aufgrund UEs	<u>Woche 48:</u> OR: 1,29 [0,44;3,80] <u>Woche 96:</u> OR: 1,09 [0,40;3,00]	Kein Zusatznutzen oder Schaden ableitbar
<i>Diarrhoe</i>	<u>Woche 96:</u> OR: 1,69 [0,98;2,90]	Kein Zusatznutzen oder Schaden ableitbar
<i>Übelkeit</i>	<u>Woche 96:</u> OR: 1,40 [0,75;2,60]	Kein Zusatznutzen oder Schaden ableitbar
<i>Fatigue</i>	<u>Woche 96:</u> OR: 1,56 [0,75;3,26]	Kein Zusatznutzen oder Schaden ableitbar
<i>Schwindel</i>	<u>Woche 96:</u> OR: 1,47 [0,73;2,94]	Kein Zusatznutzen oder Schaden ableitbar
<i>Kopfschmerzen</i>	<u>Woche 96:</u> OR: 1,65 [0,92;2,94]	Kein Zusatznutzen oder Schaden ableitbar
<i>Abnorme Träume</i>	<u>Woche 96:</u> OR: 0,89 [0,45;1,76]	Kein Zusatznutzen oder Schaden ableitbar
<i>Schlaflosigkeit</i>	<u>Woche 96:</u> OR: 0,57 [0,29;1,12]	Kein Zusatznutzen oder Schaden ableitbar
<i>Depression</i>	<u>Woche 96:</u> OR: 1,02 [0,48;2,17]	Kein Zusatznutzen oder Schaden ableitbar
<i>Hautausschlag</i>	<u>Woche 96:</u> OR: 1,06 [0,49;2,32]	Kein Zusatznutzen oder Schaden ableitbar
KI = Konfidenzintervall, MWD = Mittelwertdifferenz, OR = Odds Ratio, UE = unerwünschtes Ereignis, SUE = schwerwiegendes unerwünschtes Ereignis		

Ableitung des Zusatznutzens auf Endpunktebene

Virologisches Ansprechen (M=F)

Um valide Schlüsse zu ermöglichen, wurde für die Studien S102 und S104 für den indirekten Vergleich hinsichtlich des Endpunkts virologisches Ansprechen der Imputationsalgorithmus Missing=Failure angewandt. Der indirekte Vergleich zu Woche 48 zeigt, dass zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF kein statistisch signifikanter Unterschied besteht. Das OR [95%-KI] beträgt 0,98 [0,52;1,83]. Zu Woche 96 zeigt der indirekte Vergleich ebenso

keinen statistisch signifikanten Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF (OR [95%-KI]: 1,11 [0,61;1,99]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Veränderung der CD4-Zellzahl

Im indirekten Vergleich zu Woche 48 zeigt sich kein statistisch signifikanter Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF (MWD [95%-KI]: 10,26 [-21,84;42,36]). Auch der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen EVG/COBI/FTC/TDF und RAL+FTC/TDF (MWD [95%-KI]: 7 [-35,19;49,19]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Mortalität

Der indirekte Vergleich zu Woche 48 zeigt, dass es hinsichtlich Mortalität keinen statistisch signifikanten Unterschied zwischen den Behandlungen gibt (OR [95%-KI]: 0,10 [0,00;4,66]). Und auch zu Woche 96 ergibt der indirekte Vergleich, dass keine statistisch signifikanten Unterschiede zwischen EVG/COBI/FTC/TDF und RAL+FTC/DTF bestehen (OR [95%-KI]: 0,07 [0,00;3,10]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Gesamtrate unerwünschter Ereignisse

Zu Woche 48 ergibt der indirekte Vergleich einen statistisch signifikanten Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 2,79 [1,08;7,20]). Zu Woche 96 besteht zwischen den Behandlungen EVG/COBI/FTC/TDF und RAL+FTC/TDF kein statistisch signifikanter Unterschied (OR [95%-KI] aus indirektem Vergleich: 2,00 [0,57;7,04]). Bei der Bewertung der klinischen Relevanz des Ergebnisses muss zum einen berücksichtigt werden, dass der Unterschied lediglich zu Woche 48 und nicht mehr zu Woche 96 bestand – dies vor dem Hintergrund, dass es keinen Unterschied zwischen den Behandlungen hinsichtlich der Therapieabbrüche aufgrund von unerwünschten Ereignissen gegeben hat (OR [95%-KI]: 1,29 [0,44;3,80]). Verwendet man als Effektmaß das relative Risiko, ergibt sich ein Schätzer, dessen Konfidenzintervall die durch das IQWiG definierte Schwelle [142], ab wann ein geringer Schaden abgeleitet werden sollte, nicht unterschreitet.

Da hier laut IQWiG-Methodik kein Schaden ableitbar ist, die unerwünschten Ereignisse zudem offenkundig nicht zu einer erhöhten Abbruchrate geführt haben und bei einer lebenslang notwendigen Therapie vor allem die Langzeit-Daten klinisch relevant sind, ist das gefundene Ergebnis zu Woche 48 von minorer Bedeutung.

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Gesamtrate schwerwiegender unerwünschter Ereignisse

Zu Woche 48 besteht zwischen den Behandlungen EVG/COBI/FTC/TDF und RAL+FTC/TDF kein statistisch signifikanter Unterschied (OR [95%-KI] aus indirektem Vergleich: 1,69 [0,79;3,59]). Und auch zu Woche 96 ergibt der indirekte Vergleich keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,53 [0,78;3,00]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Therapieabbrüche aufgrund unerwünschter Ereignisse

Weder zu Woche 48 (OR [95%-KI]: 1,29 [0,44;3,80]) noch zu Woche 96 (OR [95%-KI]: 1,09 [0,40;3,00]) kann ein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden.

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Diarrhoe

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 1,69 [0,98;2,90]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Übelkeit

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 1,40 [0,75;2,60]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Fatigue

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 1,56 [0,75;3,26]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Schwindel

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 1,47 [0,73;2,94]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Kopfschmerzen

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 1,65 [0,92;2,94]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Abnorme Träume

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 0,89 [0,45;1,76]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Schlaflosigkeit

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 0,57 [0,29;1,12]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Depression

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 1,02 [0,48;2,17]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Hautausschlag

Der indirekte Vergleich zu Woche 96 zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen EVG/COBI/FTC/TDF (OR [95%-KI]: 1,06 [0,49;2,32]).

Ein Zusatznutzen oder Schaden ist nicht ableitbar.

Gesamtfazit zum Zusatznutzen

Ziel einer antiretroviralen Therapie ist es die Viruslast dauerhaft unter die Nachweisgrenze von 50 Kopien/mL abzusenken und die Symptome der Erkrankung zu minimieren. Gleichzeitig sollte die Behandlung mit möglichst wenigen Nebenwirkungen behaftet sein. Da Nebenwirkungen, im Bereich HIV vor allem stigmatisierende Nebenwirkungen, einen starken Einfluss auf die Therapieadhärenz, und damit auch auf das Risiko von Resistenzen und virologischem Versagen, haben, ist eine Minimierung der Nebenwirkungen einer ART von immenser Wichtigkeit für einen Therapieerfolg [4, 49, 50].

In der vorliegenden Nutzenbewertung wurde anhand von patientenrelevanten und validen Endpunkten gezeigt, dass EVG/COBI/FTC/TDF sowohl eine gleichwertige Wirksamkeit als auch ein äquivalentes Sicherheitsprofil aufweist wie RAL+FTC/TDF.

Mit Raltegravir wurde im Jahr 2008 der erste Vertreter der Klasse der Integrase-Inhibitoren eingeführt. Von besonderer Bedeutung ist der neue Wirkmechanismus dieser Klasse, der zur Folge hat, dass auch Patienten mit (Primär)-Resistenzen gegenüber verfügbaren antiretroviralen Substanzen erfolgreich behandelt werden können. Darüber hinaus zeichnen sich Integrase-Inhibitoren durch eine sehr gute antiretrovirale Wirksamkeit mit distinkt rascher Absenkung der Viruslast und entsprechend rascher Immunrestitution bei

gleichzeitig sehr guter Verträglichkeit aus. Dabei weisen sowohl die rasche Absenkung der Viruslast als auch die Dynamik der CD4-Zellzahl klinische Implikationen auf: Die Infektiosität eines HIV-Infizierten ist primär abhängig von der Viruslast – Daten zeigen, dass die virale Suppression unter die Nachweisgrenze das Risiko für eine Transmission fast eliminiert. Je rascher also die Nachweisgrenze erreicht, desto kürzer der Zeitraum, in der HIV potentiell übertragen werden kann [143]. Die CD4-Zellzahl ist ein Surrogat für den immunologischen Status, das zuverlässig die Abschätzung des individuellen Risikos, an Aids zu erkranken, erlaubt und damit der primäre Marker für den Beginn einer antiretroviralen Therapie: Diese sollte spätestens bei einer CD4-Zellzahl von 350 CD-Zellen/ μ L begonnen werden [2]; ab einer CD4-Zellzahl von <200 Zellen/ μ L muss mit zunehmender Dauer der Immunsuppression vermehrt mit dem Auftreten von Aids-definierenden Erkrankungen gerechnet werden [74]. Damit ist die Dynamik des CD4-Zellanstiegs gerade zu Beginn einer antiretroviralen Therapie klinisch hochrelevant, da die Patienten bei schnellerem Anstieg der CD4-Zellzahl auch rascher oberhalb des potentiellen Risiko-Bereichs für die Entwicklung Aids-definierender Erkrankungen liegen. Die raschere Rekonstitution des Immunsystems ist somit direkt patientenrelevant.

Die Klasse der Integrase-Inhibitoren zeichnet sich also insbesondere durch sehr gute Wirksamkeit mit rascher Viruslast-Absenkung und sehr guter Verträglichkeit aus; zudem bedeutet der neue Wirkmechanismus mit einem anderen viralen Enzym-Target, dass auch Patienten mit Primärresistenzen gegen die anderen Substanzklassen erfolgreich mit Integrase-Inhibitoren behandelt werden können.

Obschon EVG/COBI/FTC/TDF im indirekten Vergleich sowohl eine gleichwertige Wirksamkeit als auch ein äquivalentes Sicherheitsprofil aufweist wie RAL+FTC/TDF, weist EVG/COBI/FTC/TDF durch die Formulierung erhebliche patientenrelevante Vorteile auf:

Die Einmalgabe ist ein wesentlicher Einflussfaktor auf die Adhärenz: Willig et al. [51] untersuchten Daten von 542 Patienten, die zwischen 2000 und 2007 ihre erste ART begonnen haben und diese Therapie mindestens 14 Tage lang einnahmen. Die Daten zeigen, dass die QD-Gabe der BID-Gabe im Hinblick auf die Persistenz der Therapie über die Zeit statistisch signifikant überlegen ist ($p=0,0002$, Abbildung 23).

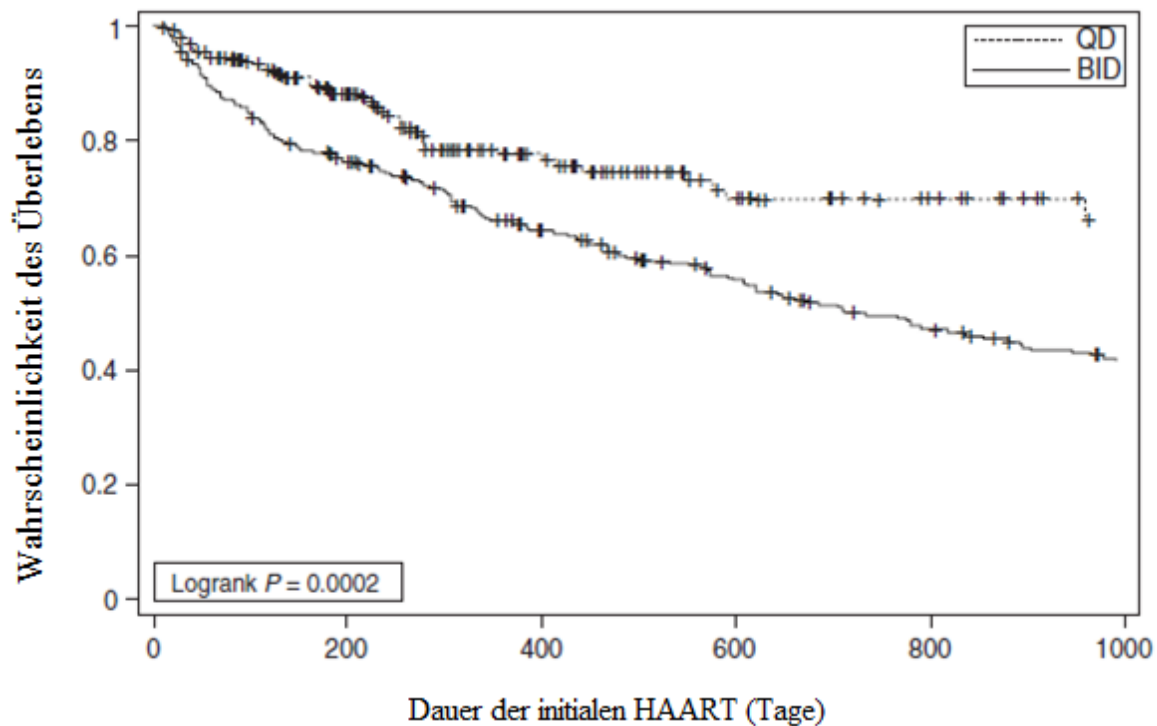


Abbildung 23: Dauerhaftigkeit (Persistenz) einer Initialtherapie als Funktion der Dosierungskomplexität [51]

Zu dem gleichen Schluss kommen Parienti et al in einer Meta-Analyse mit insgesamt elf klinischen Studien [52] ($n=3.029$): Die Adhärenz-Rate unter einem einmal täglichen Regime war signifikant höher als unter einem zweimal täglichen Regime (+2,9%, 95%-KI: [1,0%;4,8%], $p<0,003$). Bei den 5 Studien mit nicht-vorbehandelten Patienten zeigte sich ($n=1.927$) unter einmal täglichen Regimen sowohl eine signifikant bessere Adhärenz (+4,4% 95%-KI: [1,8%;7,0%], $p<0,001$) als auch ein besseres virologisches Outcome (+5,7%, 95%-KI: [0,7%;10,8%], $p<0,027$) (Abbildung 24).

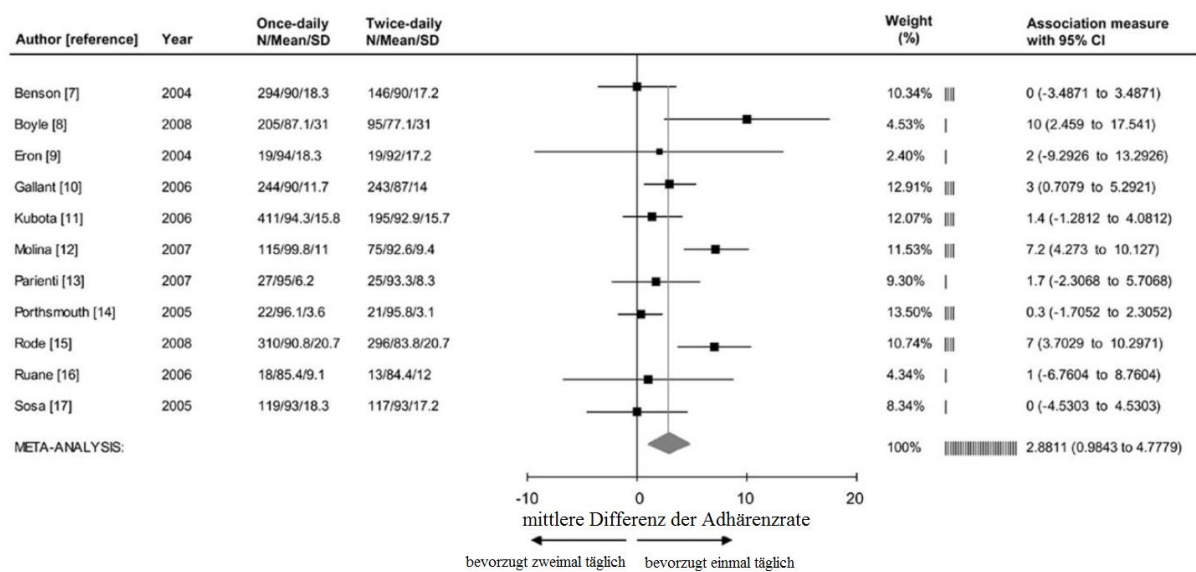


Abbildung 24: Forestplot des Effekts einer Einmal- im Vergleich zu einer Zweimal-Gabe auf die Adhärenzrate [52]

Darüber hinaus birgt die zweimal tägliche Gabe ebenso wie die Gabe von Mehrtabletten-Regimen das Risiko der selektiven Nicht-Adhärenz, das heißt der selektiven Einnahme, resp. Nicht-Einnahme bestimmter Substanzen des antiretroviralen Regimes – sei es unbeabsichtigt durch Vergessen oder beabsichtigt zur Vermeidung von Nebenwirkungen. Der Einsatz eines Single-Tablet-Regimen hat zur Folge, dass durch die Formulierung automatisch tatsächlich alle notwendigen Komponenten seines Regimes eingenommen werden - das heißt, die selektive Einnahme, resp. Nicht-Einnahme wird wirksam verhindert. Dies führt, wie im Folgenden dargestellt, zu einem verbesserten klinischen Outcome hinsichtlich der virologischen Suppression und reduzierter Hospitalisierungsraten und ist damit hochrelevant für den Patienten.

Eine Analyse klinischer Daten von 1.604 Patienten in Italien ergab in der COMPACT-Studie eine Adhärenz von $80,4 \pm 14,7\%$ für Patienten unter Single-Tablet-Regimen - unter Mehrtablettenregimen war die Adhärenz von Patienten aufgrund der beobachteten selektiven Nicht-Adhärenz (Tage ohne NRTI-Backbone oder 3. Partner) teils deutlich niedriger: Die Rate an selektiver Nicht-Adhärenz betrug 6% unter PI-basierten, 3% unter NNRTI-basierten und 13% unter Raltegravir-basierten Regimen (Abbildung 25). Die klinische Relevanz der gefundenen Ergebnisse zeigt sich in der Korrelation zwischen Adhärenz- und therapeutischer Erfolgsrate, gemessen anhand von CD4-Zellzahl (% Patienten mit CD4-Zellzahl >500 Zellen/ μL) und Viruslast (% Patienten mit <50 RNA-Kopien/mL). Der Anteil der Patienten mit Viruslast <50 Kopien/mL lag bei Patienten mit Single-Tablet-Regimen bei 96% - im Vergleich dazu bei 78% unter PI-basierten Regimen, 88% unter NNRTI-basierten Regimen und 87% unter INI (Raltegravir)-basierten Regimen. Auch die immunologischen Erfolgsraten korrelierten analog: Der Anteil der Patienten mit CD4-Zellzahl >500 Zellen/ μL lag bei Patienten mit Single-Tablet-Regimen bei 61% - im Vergleich dazu bei 44% unter PI-

basierten Regimen, 48% unter NNRTI-basierten Regimen und 42% unter INI (Raltegravir)-basierten Regimen.

Dies hatte bei Patienten mit einer Adhärenz-Rate von $\leq 80\%$ klinische Folgen: Patienten mit selektiver Nicht-Adhärenz zeigten je nach Ausmaß der selektiven Nicht-Adhärenz ein um 24%, resp. 39% erhöhtes Risiko für eine Hospitalisierung [53].

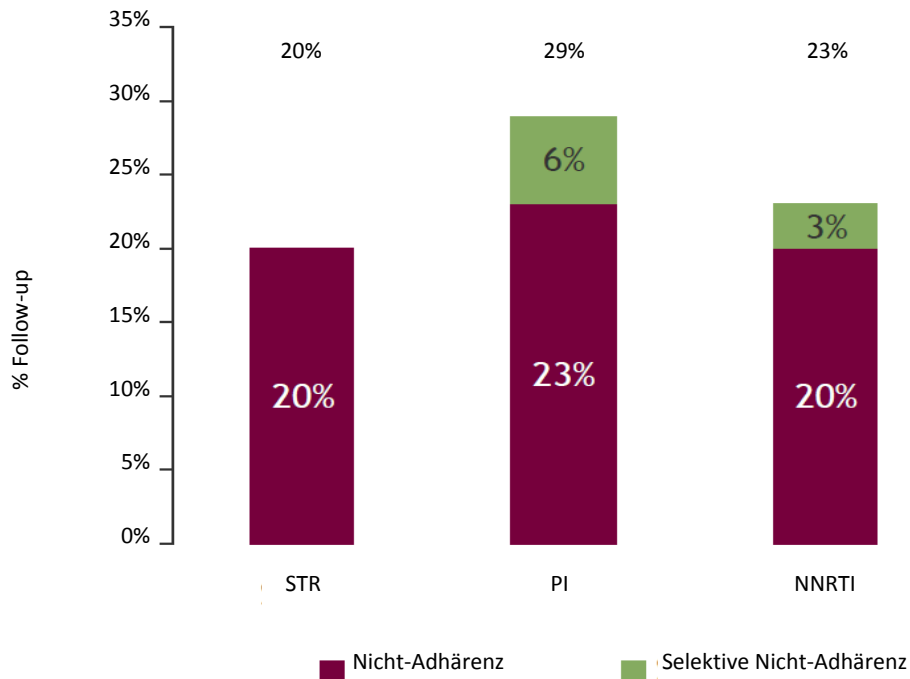


Abbildung 25: Nicht-Adhärenz bei verschiedenen ART-Regimen [53]

Bestätigende Daten liefert eine weitere italienische Kohorte: Hier war die (selbstberichtete) Nicht-Adhärenz bei Patienten unter einem Single-Tablet-Regimen deutlich niedriger als bei Patienten, die Mehrtablettentregime einnahmen [54].

Der Einfluss der selektiven Nicht-Adhärenz auf das klinische Outcome wurde auch auf Basis von Daten von 6.938 Patienten in der US-amerikanischen Medicaid Claims Datenbank (Zeitraum: Juni 2009 – Dezember 2011) untersucht: Die Daten-Analyse zeigte eine höhere Gesamt-Adhärenz für Patienten unter einem Single-Tablet-Regime (85,7%) im Vergleich zu Patienten unter einem Raltegravir-basierten (73,9%), einem PI/r-basierten (77,9%) oder einem NNRTI-basierten (79,6%) Mehrtablettentregime (>2 Tabletten/Tag) [55]. Während eine komplette Nicht-Adhärenz mit 14% bis 16% in allen vier Gruppen ähnlich war, zeigten sich erhebliche Unterschiede in der selektiven Nicht-Adhärenz: Diese lag für PI/r-basierte Regime bei 7%, bei NNRTI-basierten Regimen bei 7% und bei Raltegravir-basierten Regimen mit 13% deutlich höher. Zusammengefasst zeigten sich somit Adhärenz-Probleme bei 26% der Patienten unter RAL-basierter ART, 23% der Patienten unter PI-basierter ART und 21% der

Patienten unter NNRTI-basierter ART – demgegenüber bei lediglich 14% der Patienten unter einem STR (jeweils $p < 0,0001$ versus andere Regime). Dies zeigt sich auch klinisch: Selektive Nicht-Adhärenz war assoziiert mit einem statistisch signifikant erhöhten Risiko (43-54%) für eine Hospitalisierung [55].

In einer portugiesischen Kohorte mit 1.436 HIV-Patienten (Zeitraum: 2001 – 2011) wurden folgende Regime miteinander verglichen: QD, BID, STR (FTC/TDF/EFV), FTC/TDF+EFV, Einsatz von Fixkombinationen und keine Fixkombinationen. Während bereits die vollständige Nicht-Adhärenz bei Patienten, die keine Fixkombinationen, resp. ein zweimal tägliches Regime einnahmen mit 23%, resp. 22% am schlechtesten war, zeigte sich bereits unter der Einnahme von Fixkombinationen, resp. einem einmal täglichen Regime eine statistisch signifikante deutliche Verbesserung auf 15%, resp. 13% vollständiger Nicht-Adhärenz; diese war unter einem Single-Tablet-Regimen mit 11% am geringsten. Fasst man vollständige Nicht-Adhärenz und selektive Nicht-Adhärenz zusammen, ergeben sich klare, statistisch signifikante Vorteile für Single-Tablet-Regimen: Während die Gesamtrate von Patienten mit Adhärenzproblemen unter Einzelsubstanz-Regimen oder Regimen, bei denen Fixkombinationen eingesetzt wurden, bei 27%, resp. 18% lag, zeigten nur 11% von Patienten unter einem STR Probleme mit der Adhärenz ($p < 0,001$, resp. $p < 0,02$). Auch im Vergleich zu zweimal täglichem Mehrtablettenregimen (26%) zeigten sich statistisch signifikante Vorteile für ein Single-Tablet-Regimen ($p < 0,001$). Sogar der direkte Vergleich zwischen den substanzgleichen Formulierungen aus TDF, FTC und EFV ergab Vorteile für das STR aus FTC/TDF/EFV (Gesamt-Adhärenzrate: 89%) im Vergleich zu EFV in Kombination mit FTC/TDF (Gesamt-Adhärenzrate: 83%) ($p = 0,003$). Sowohl komplette als auch selektive Nicht-Adhärenz waren Prädiktoren für virologisches Versagen ($p < 0,001$ bzw. $p = 0,020$); eine Erhöhung der Nicht-Adhärenz um 1% resultierte in einem bis zu 4% höheren Risiko eines virologischen Versagens. Auch die Wahrscheinlichkeit für eine Hospitalisierung war bei einem Single-Tablet-Regime statistisch signifikant geringer im Vergleich zu den anderen Regimen (3,8% versus 10,3%; $p < 0,001$) [56].

Auch die Zufriedenheit der Patienten war im Falle eines Single-Tablet-Regimens am höchsten. Dabei zeigte sich, dass die Einfachheit eines Regimes – sowohl hinsichtlich der Dosisfrequenz als auch hinsichtlich der Tablettenzahl – die Patienten-Akzeptanz des Regimes, die eine Grundvoraussetzung für Adhärenz darstellt, deutlich erhöht [57].

Die Ergebnisse der oben dargestellten Analysen zeigen übereinstimmend einen Vorteil von einmal täglichem und vor allem Single-Tablet-Regimen hinsichtlich des Therapieerfolgs und möglicher klinischer Folgen (Hospitalisierung). Dieser Vorteil resultiert insbesondere auch aus der geringeren Rate an selektiver Nicht-Adhärenz, die insgesamt zu einer verbesserten Adhärenz bei Single-Tablet-Regimen gegenüber Mehrtablettenregimen führt. Die auf diese Weise verringerte selektive Nicht-Adhärenz trägt in einer Indikation, die eine lebenslange Therapie erfordert, zu einem langfristigen Therapieerfolg mit nachhaltiger viraler Suppression bei.

Insgesamt wird für EVG/COBI/FTC/TDF im Vergleich zur zweckmäßigen Vergleichstherapie RAL+FTC/TDF auf Basis der Gleichwertigkeit im indirekten Vergleich

und auf Basis der patientenrelevanten Vorteile, die sich durch die Formulierung ergeben, ein **Hinweis** auf einen **geringen** Zusatznutzen abgeleitet.

Gesamtableitung des Zusatznutzen von EVG/COBI/FTC/TDF gegenüber der zweckmäßigen Vergleichstherapie, einer individuellen antiretroviralen Therapie in Abhängigkeit der Vortherapie(en) und unter Berücksichtigung des Grundes für den Therapiewechsel, insbesondere Therapieversagen aufgrund eines virologischen Versagens und etwaig einhergehender Resistenzbildung oder aufgrund von Nebenwirkungen

Aufgrund der dargestellten Übertragbarkeit der Daten der nicht-vorbehandelten Patienten auf die vorbehandelten Patienten kann in der Gesamtschau für vorbehandelte HIV-1-Infizierte, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von Stribild assoziiert sind, für EVG/COBI/FTC/TDF ein geringer Zusatznutzen gegenüber der drei gewählten Vergleichstherapien EFV+FTC/TDF, ATV/r+FTC/TDF und RAL+FTC/TDF abgeleitet werden. EVG/COBI/FTC/TDF ist eine Dreifachtherapie, bestehend aus einem FTC/TDF-Backbone und einer Substanz aus der Substanzklasse der 3. Partner, die als komplettes Regime nicht in Kombination mit anderen antiretroviralen Arzneimitteln eingesetzt werden darf. Es ist davon auszugehen, dass die zweckmäßige Vergleichstherapie primär ebenfalls aus der Kombination eines Backbones mit einem 3. Partner besteht. Die drei gewählten Vergleichstherapien stellen jeweils eine Kombination eines FTC/TDF-Backbones mit einem 3. Partner aus einer der drei bei der Umstellung am häufigsten eingesetzten Substanzklassen NNRTI, PI und INI, dar. Sie repräsentieren somit die individualisierte zweckmäßige Vergleichstherapie und decken den mit Abstand größten Teil der hier relevanten Patientenpopulation ab.

Aufgrund der Studienanzahl, die in den Vergleich von EVG/COBI/FTC/TDF mit den ausgewählten Vertretern der zweckmäßigen Vergleichstherapie (individuelle antiretrovirale Therapie) und dem Beleg für einen geringen Zusatznutzen im Vergleich zu EFV/FTC/TDF bzw. den Hinweisen für einen geringen Zusatznutzen im Vergleich zu ATV/r+FTC/TDF und zu RAL+FTC/TDF, ist im Vergleich zur zweckmäßigen Vergleichstherapie insgesamt ein **geringer** Zusatznutzen für EVG/COBI/FTC/TDF abzuleiten; die Wahrscheinlichkeit wird als **Beleg** eingestuft.

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 4-195: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzen

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene HIV-1-infizierte vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind	Gering

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 4-196: 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.

Zum Zeitpunkt der Dossiereinreichung lagen keine direkten Vergleichsstudien zwischen dem zu bewertenden Arzneimittel (EVG/COBI/FTC/TDF) und der zweckmäßigen Vergleichstherapie (RAL+FTC/TDF) zu nicht-vorbehandelten bzw. vorbehandelten Patienten vor.

Aus diesem Grund wurde zur Bewertung des Nutzens und Zusatznutzens von EVG/COBI/FTC/TDF ein adjustierter indirekter Vergleich mit EFV/FTC/TDF als Brückenkomparator durchgeführt. Es wurden ausschließlich RCT verwendet, da diese mit der geringsten Ergebnisunsicherheit behaftet sind. Der durchgeführte adjustierte indirekte Vergleich erfüllt den aktuellen Stand der Wissenschaft hinsichtlich der Anforderungen an eine solche Analyse und ist somit aussagekräftig.

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nichtrandomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

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.

Wie in Abschnitt 4.2.1 dargestellt, wurde die Zulassung für die Patientenpopulation der vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, von der Europäischen Zulassungsbehörde auf Basis der Daten zu den nicht-vorbehandelten Patienten erteilt. Das klinische Entwicklungsprogramm zur Umstellung vorbehandelte Patienten läuft derzeit (siehe Tabelle 4-3) – es liegen jedoch noch keine Daten aus randomisierten klinischen Studien zur Population der vorbehandelten Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einer der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, vor (siehe Tabelle 4-4). Allerdings sind die vorliegenden Daten zu nicht-vorbehandelten Patienten auf die vorbehandelte Population zu übertragen:

Bei den Therapieprinzipien der ART gilt es als generelles Ziel der Therapie, durch die Hemmung der HIV-Replikation infektionsbedingte Symptome zu unterdrücken, Krankheitsprogression zu vermindern, eine immunologische Rekonstitution zu erreichen und die chronische Immunaktivierung zu reduzieren. Nur eine hohe virologische Wirksamkeit der ART und eine dauerhafte Suppression der Viruslast auf unter 50 RNA-Kopien/mL können Resistenzentwicklung und konsekutives Therapieversagen nachhaltig verhindern [2].

Grundsätzlich unterscheidet man in der ART zwei Patientenpopulationen, für die die übergeordneten Therapieprinzipien (dauerhafte Hemmung der Replikation und virale Suppression auf <50 Kopien/mL) gleichermaßen gültig sind:

1. Patienten, die noch keine antiretrovirale Therapie erhalten (nicht-vorbehandelte Patienten):
2. Patienten, die antiretroviral behandelt werden (vorbehandelte Patienten)

In den Empfehlungen zur Umstellung antiretroviral vorbehandelter Patienten wird in allen nationalen und internationalen Leitlinien zwischen zwei Subpopulationen, bzw. Indikationen zur Umstellung unterschieden [2-6]:

- Supprimierte Patienten: Umstellung aufgrund von Nebenwirkungen oder zur Therapievereinfachung und Unterstützung, bzw. Verbesserung der Adhärenz
- Patienten von mit virologischem Versagen: Umstellung aufgrund des virologischen Versagens und konsekutiver Virämie

Diese Unterscheidung ergibt sich aus dem entscheidenden Kriterium, das die Grundvoraussetzung für die virologische Wirksamkeit einer ART darstellt: Dem Vorliegen Resistenz-assoziiierter Mutationen gegen die einzusetzenden Substanzen.

Die Unterscheidung zwischen Patienten mit und ohne Resistenz-assoziierten Mutationen ist von elementarer Bedeutung: Ein vorbehandelter Patient, dessen Virus keine Resistenz-assoziierten Mutationen gegen antiretrovirale Substanzen aufweist, entspricht hinsichtlich des Resistenz-Status einem nicht-vorbehandelten Patienten. Damit ist dieser Patient im Hinblick auf die entsprechenden antiretroviralen Substanzen als funktionell nicht-vorbehandelt zu definieren - womit beide Populationen (nicht-vorbehandelt und vorbehandelt ohne Resistenzen) als gleichgestellt zu betrachten sind. Vor dem Hintergrund der Gleichstellung der Population können klinische Daten zu nicht-vorbehandelten Patienten somit auf vorbehandelte Patienten ohne Resistenzen extrapoliert und damit übertragen werden [7].

Die Rationale der Übertragbarkeit klinischer Daten von nicht-vorbehandelten Patienten auf vorbehandelte Patienten, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, lässt sich zudem wie folgt belegen:

1. Die DHHS-Leitlinien führen in den Empfehlungen zur Umstellung supprimierter Patienten explizit aus: „Für Regime, die zur Therapievereinfachung eingesetzt werden, sollte grundsätzlich eine hohe Wirksamkeit bei antiretroviral nicht-vorbehandelten Patienten nachgewiesen worden sein, oder es sollte für den Patienten anhand seiner individuellen Vorbehandlung und seines Resistenzprofils eine hohe Wirksamkeit vorhergesagt werden können“ [4]. Gemäß der DHHS-Leitlinie ist somit davon auszugehen, dass die nachgewiesene Wirksamkeit eines antiretroviralen Regimes bei nicht-vorbehandelten Patienten auf suppressierte vorbehandelte Patienten extrapoliert und damit übertragen werden kann.
2. Die Übertragbarkeit der Daten ist hinsichtlich der Wirksamkeit und des Verträglichkeitsprofils auch rein mechanistisch und zwar sowohl pharmakologisch als auch biologisch nachvollziehbar:

Eine Initialtherapie muss ausreichend virologisch wirksam sein, um das Virus erfolgreich zu kontrollieren und die Ausgangsviruslast unabhängig von ihrer Höhe (etwa 50% der Patienten weisen eine Ausgangsviruslast von >100.000 Kopien/mL auf) unter die Nachweisgrenze zu senken. Antiretrovirale Therapieregime, die zur Umstellung supprimierter Patienten eingesetzt werden, hingegen müssen lediglich ausreichend virologisch wirksam sein, um die Viruslast unter der Nachweisgrenze zu halten. Entsprechend muss die virologische Potenz einer antiretroviralen Initialtherapie mindestens so hoch sein wie die eines Regimes, auf das ein supprimierter Patient umgestellt wird. Vor diesem Hintergrund ist davon auszugehen, dass klinische Daten zur virologischen Wirksamkeit, die bei nicht-vorbehandelten Patienten erhoben werden, direkt auf die Umstellung vorbehandelter Patienten extrapolierbar sind, sofern diese keine Resistenzen gegen die Substanzen aufweisen.

Unter Therapie auftretende Nebenwirkungen sind primär Substanz- und nicht Virus-assoziiert, woraus schlussgefolgert werden kann, dass die Höhe der Viruslast zu Beginn einer bestimmten Therapie keinen Einfluss auf die auftretenden Nebenwirkungen hat. Entsprechend wird auch im Bericht zum Sicherheitsprofil in den Fachinformationen [8-35] zu den verschiedenen antiretroviralen Substanzen nicht zwischen vorbehandelten und nicht-vorbehandelten Patienten unterschieden – ganz im Gegenteil beruht das in den Fachinformationen berichtete Sicherheitsprofil stets auf Basis des gesamten klinischen Entwicklungsprogramms.

3. Die EMA hat für Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil sowohl die Zulassung für nicht-vorbehandelte Patienten, als auch die Zulassung für die Patienten erteilt, die keine Resistenzen gegen einen der drei enthaltenen antiretroviralen Substanzen aufweisen. Da diese Zulassung ausschließlich auf Basis von klinischen Daten zu nicht-vorbehandelten Patienten eingereicht wurde, muss geschlussfolgert werden, dass die EMA die beiden Population auf Basis der dargestellten wissenschaftlichen Rationale funktionell gleichgestellt und die Daten entsprechend extrapoliert hat. Dies wird auch durch ein von der EMA erstelltes Konzeptpapier zur Extrapolation von klinischen Daten klar unterstützt: Die angestrebte Extrapolation ist definiert als Deduktion von Studienergebnissen in bestimmten Subpopulation (Quellenpopulation) auf andere Subpopulationen (Zielpopulation) – analog gilt dies für verwandte Erkrankungen oder Substanzen. Zielsetzung der Extrapolation von Daten ist zum einen die Reduktion des Bedarfs der zusätzlichen Datengenerierung für die Zielpopulation, und zum anderen die Vermeidung unnötiger Studien und Allokation von Ressourcen in die klinischen Studien, die tatsächlich notwendig sind. Grundvoraussetzung für die Übertragbarkeit ist eine pharmakologische und biologische Rationale, die für Quellen- und Zielpopulation anhand von Erkrankung, Medikation (Wirkmechanismus, Pharmakokinetik und Pharmakodynamik) sowie Anwendbarkeit von Wirksamkeits- und Sicherheitsendpunkten bewertet werden sollte [7]. Für die beiden hier vorliegenden Populationen der nicht-vorbehandelten Quellenpopulation und der Zielpopulation von Patienten ohne Resistenz-assoziierte Mutationen, zeigt sich die geforderte pharmakologische und biologische Rationale deutlich bei der vollständigen Übereinstimmung der genannten Parameter. Vor diesem Hintergrund ist mit Erteilung der Zulassung also davon auszugehen, dass die Europäische Zulassungsbehörde die Übertragbarkeit der Daten von den nicht-vorbehandelten Patienten auf die vorbehandelten Patienten ohne Resistenz-assoziierte Mutationen als gegeben ansieht.

In logischer Konsequenz bedeutet die dargestellte grundsätzliche Übertragbarkeit der Daten, dass auch ein in den klinischen Studien zu nicht-vorbehandelten Patienten gefundener Zusatznutzen für auf die vorbehandelte Patientenpopulation, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von EVG/COBI/FTC/TDF assoziiert sind, zu übertragen ist.

Die Übertragbarkeit der Daten wird zusätzlich gestützt durch die Ergebnisse der einarmigen Studie GS-US-236-0123, in welcher Patienten, die mindestens sechs Monate mit

RAL+FTC/TDF vorbehandelt wurden, auf EVG/COBI/FTC/TDF umgestellt wurden. Nach 12 Wochen Behandlung mit EVG/COBI/FTC/TDF zeigten alle 48 (100%) Patienten virologisches Ansprechen. Hinsichtlich der CD4-Zellzahl zeigten die Patienten eine numerische, aber nicht signifikante Verbesserung (Baseline: 711 Zellen/ μL , zu Woche 12: 744 Zellen/ μL). In der Studie traten weder Todesfälle, noch schwerwiegende unerwünschte Ereignisse oder Therapieabbrüche aufgrund unerwünschter Ereignisse auf. 36 (75,0%) Patienten hatten unerwünschte Ereignisse; davon wurde bei elf Patienten (22,9%) das unerwünschte Ereignis als Grad 2 eingestuft. Unerwünschte Ereignisse höheren Grads traten nicht auf. 12 (25,0%) Patienten wiesen Erkrankungen des Gastrointestinaltrakts auf; hier war Diarrhoe das am häufigsten auftretende Ereignis aus dieser SOC (fünf [10,4%] Patienten). Psychiatrische Erkrankungen traten bei elf (22,9%) Patienten auf, mit Schlaflosigkeit als am häufigsten auftretendes Ereignis (fünf [10,4%] Patienten). Für eine detailliertere Übersicht der Ergebnisse aus dieser Studie siehe Abschnitt 4.3.2.3.

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 2005¹³, Molenberghs 2010¹⁴). 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 2006¹⁵) 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

¹³ Burzykowski T (Ed.: *The evaluation of surrogate endpoints*. New York: Springer; 2005.

¹⁴ 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.

¹⁵ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006¹⁶) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Aufgrund der großen Therapieerfolge, die mit der Einführung der hochwirksamen ART einhergehen, wird die Erhebung von Endpunkten wie Morbidität und Mortalität in klinischen Studien zunehmend impraktikabel. Das Fortschreiten der Erkrankung kann durch die ART stark verlangsamt werden, wodurch eine unverhältnismäßig lange Beobachtungszeit in klinischen Studien nötig wäre. Folglich ist es heutzutage üblich, Surrogatendpunkte einzusetzen, die aufgrund ihrer früheren Messbarkeit, eine wesentlich kürzere Studiendauer erlauben, um Aussagen zu Endpunkten wie Morbidität und Mortalität zu ermöglichen.

Virologisches Ansprechen

Wie bereits durch den G-BA und das IQWiG festgestellt, ist die Bedeutung der Viruslast als Endpunkt im vorliegenden Anwendungsgebiet grundsätzlich unbestritten, und der primäre Endpunkt virologisches Ansprechen wird in der Gesamtschau der nachfolgend aufgeführten Argumente als ausreichend valider Surrogatparameter für den kombinierten Endpunkt „Aids-definierende Erkrankungen/ Tod“ eingeschätzt [62]: Als Gründe für die Einschätzung des Endpunktes Viruslast als valides Surrogat stützt sich das IQWiG auf folgende Punkte:

- Arzneimittelstudien, die auf die Reduzierung der Viruslast ausgerichtet waren, führten im Behandlungsalltag zu einer dramatischen Verbesserung der Prognose für HIV-Patienten hinsichtlich Überleben und Ausbruch der Erkrankung.

¹⁶ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Stat Med* 2006; 25(2): 183-203.

- Das biologische Modell ist plausibel.
- Die Korrelation zwischen der individuellen Veränderung der Viruslast und dem Risiko für Aids-definierende Erkrankungen/ Tod ist sehr deutlich [63].
- Auch Daten auf Studienebene (Meta-Regressionen) widersprechen dieser Assoziation nicht, auch wenn eine klare Korrelation zwischen den Effekten auf das Surrogat und den Effekten auf den patientenrelevanten Endpunkt „Aids-definierende Erkrankungen/ Tod“ kaum gegeben ist. Bei diesen Studien wird der Zusammenhang wahrscheinlich durch den Einschluss vieler Nichtunterlegenheitsstudien (nahe am Nullpunkt) verwässert. Aus diesem Grund kann das Ausmaß des Zusatznutzens für den kombinierten Endpunkt „Aids-definierende Erkrankungen/ Tod“ auf Basis des Surrogats nicht quantifiziert werden.

Dieser Einschätzung wird hier gefolgt und dementsprechend stellt der Endpunkt „Virologisches Ansprechen“ ein ausreichend valides Surrogat für den patientenrelevanten Endpunkt „Aids-definierende Erkrankungen/ Tod“ dar.

Zeit bis Verlust des virologischen Ansprechens

Primäres Ziel einer ART ist nicht nur der Abfall der Viruslast unter die Nachweisgrenze, sondern auch die dauerhafte virale Suppression. Die Deutsch-Österreichischen Leitlinien definieren das Ziel der ART wie folgt: „Ziel der antiretroviralen Therapie (ART) ist es, durch die Hemmung der HIV-Replikation infektionsbedingte Symptome zu unterdrücken, die Krankheitsprogression zu vermindern, eine Rekonstitution der zellulären Immunität zu erreichen und die chronische Immunaktivierung mit ihren resultierenden Entzündungsprozessen zu reduzieren.“ [2]. Die EMA empfiehlt entsprechend die Erfassung des Erreichens und Aufrechterhaltens des virologischen Ansprechens im Rahmen klinischer Studien [61]:

„The proportion of subjects that achieves and maintains undetectable plasma HIV-RNA (i.e. < 50 copies/ml) is the preferred primary efficacy endpoint for studies in treatment naïve as well as treatment experienced populations. [...]“

Eine höhere Plasmavirämie geht mit einem rascheren Absinken der CD4-Zellzahl und einem erhöhten Progressionsrisiko einher [64]. Durch das Erreichen und Aufrechterhalten des virologischen Ansprechens, der konsekutiven viralen Suppression und der assoziierten immunologischen Rekonstitution wird das Progressionsrisiko deutlich reduziert, und das Auftreten Aids-definierender Erkrankungen oder Tod können häufig vermieden werden.

Die Zeit, über die das virologische Ansprechen und damit die virale Suppression ohne Therapiewechsel aufrechterhalten werden, beschreibt die Dauer einer erfolgreichen ART (Persistenz). Dieser Zeitraum steht für direkt erlebbare, patientenrelevante Vorteile im Sinne einer symptomfreien Zeit und ist somit ein patientenrelevanter Morbiditäts-Endpunkt. Zudem ist dieser Parameter für die lebenslang notwendige ART insofern von Bedeutung, als dass eine Therapieveränderung aufgrund von virologischem Versagen im Grunde immer den dauerhaften Verlust von Therapieoption(en) bedeutet [65]. Resistenzen schränken nicht nur

die therapeutischen Optionen ein, sondern erhöhen zudem Komplexität und Kosten der nachfolgend eingesetzten Regime für den individuellen Patienten; zudem stellt die Transmission resistenter Viren ein Problem für die öffentliche Gesundheit dar [66]. Darüber hinaus steigt das Risiko eines virologischen Rebounds mit der Komplexität des Regimes [45].

Es ist von besonderer klinischer Relevanz, dass nur durch eine **dauerhafte** Reduktion der Viruslast auf <50 Kopien/mL das Risiko für Resistenzentwicklung und Therapieversagen mit konsekutivem immunologischen Therapieversagen und Krankheitsprogression **nachhaltig** verhindert werden kann.

Gestützt wird dies durch die Ergebnisse SMART-Studie:

Eine wirksame ART ist mit substanziellen Vorteilen im Hinblick auf Mortalität und Morbidität assoziiert - allerdings auch mit potentiellen Nebenwirkungen. Basierend auf der Hypothese, Nebenwirkungen durch eine reduzierte Substanzexposition minimieren zu können, wurden in SMART zwei unterschiedliche Therapiestrategien miteinander verglichen: CD4-zellgesteuerte, episodische Einnahme versus kontinuierliche Einnahme der ART. Im Ergebnis zeigte sich, dass eine episodische ART das Risiko für Nebenwirkungen nicht reduziert, und zu einem signifikanten Anstieg des Risikos für opportunistische Infektionen und Tod führt – dies primär als Konsequenz des Abfalls der CD4-Zellzahl und dem Anstieg der Viruslast [67].

Ein langfristiges, dauerhaftes virologisches Ansprechen führt zu einer fortschreitenden immunologischen Rekonstitution, wie die über Jahre ansteigenden CD4-Zellzahlen belegen [68].

Des Weiteren belegen verschiedene Studien belegen eine Assoziation zwischen Dauer der viralen Suppression und Rebound-Rate: Die Rate virologischer Rebounds nahm über die Zeit unter ART mit viraler Suppression deutlich ab [45, 69, 70]. Zudem ist das Risiko eines virologischen Rebounds bei nicht-vorbehandelten Patienten niedriger als bei behandelten Patienten, die ihre Therapie wechseln [45].

Dauer des virologischen Ansprechens und Persistenz sind somit kritische Parameter für die Reduktion von Morbidität und Mortalität, dies bedingt durch das assoziierte immunologische Ansprechen [42], als auch Risikofaktoren für einen virologischen Rebound mit konsekutivem Therapieversagen und Krankheitsprogression.

Vor diesem Hintergrund stellt der Endpunkt, welcher die Dauer des virologischen Ansprechens abbildet, ebenfalls ein ausreichend valides Surrogat für den patientenrelevanten Endpunkt „Aids-definierende Erkrankungen/Tod“ dar.

Veränderung der CD4-Zellzahl

Die absolute CD4-Zellzahl ist das Surrogat für den immunologischen Status und der primäre Indikator für den Therapiebeginn; die Veränderung der CD4-Zellzahl unter Therapie zeigt das immunologische Ansprechen. Neben dem virologischen Ansprechen, erfasst durch die

Viruslast, sollte auch immer die CD4-Zellzahl untersucht und bewertet werden. Dies wird auch vom G-BA so empfohlen [71]:

„Der Einsatz von Surrogat-Parametern, die eine zeitnahe Bewertung des Krankheitsverlaufes ermöglichen, ist daher unumgänglich. Als Surrogatparameter zur Therapie HIV-Infizierter bewährt und etabliert sind die HI-Viruslast als Maß der Virusreplikation und die CD4-Zellzahl als Maß der ("quantitativen") Funktion des Immunsystems [...].“

Demnach werden die HI-Viruslast („...als Maß der Virusreplikation...“), wie auch die CD4-Zellzahl („...als Maß der ["quantitativen"] Funktion des Immunsystems...“) vom G-BA als valide Surrogatparameter eingestuft.

Auch die EMA stuft den Endpunkt CD4-Zellzahl als valides Surrogat ein, und empfiehlt eine gleichzeitige Erfassung von HI-Viruslast und CD4-Zellzahl [61]:

„Since the introduction of Highly Active Anti-retroviral Therapy (HAART = ART in this document), viral load and CD4+ T-cell counts have been generally accepted as surrogate markers for efficacy in studies with anti-retroviral agents. [...].“

„Effects on the CD4+ T-cell count should always be documented. [...].“

Gemäß FDA sollte die Erfassung der CD4-Zellzahl gemeinsam mit der HI-Viruslast, und ggf. den CDC Klasse C Ereignissen, erfolgen:

„Analyses evaluating changes over time in CD4 cell counts should accompany the analyses of HIV RNA. Clinical endpoint data (CDC class C events) should also be collected, analyzed, and submitted with the NDA. However, the frequency of such events is likely to be low particularly over a 24-week time period.“

Marschner et al. [72] untersuchten den Zusammenhang zwischen HI-Viruslast bzw. CD4-Zellzahl und Krankheitsprogression. Dabei zeigte sich, dass ein Anstieg der CD4-Zellzahl versus Baseline um 100 Zellen/ μ L bis Woche 24 mit einem 60% geringeren Risiko einer Krankheitsprogression verbunden ist (95%-KI: 51%-64%, $p < 0,001$).

In einer Meta-Analyse von 15 Studien zeigen Hill et al. [73] einen signifikanten Zusammenhang zwischen der durchschnittlichen Veränderung von CD4-Zellzahl und patientenrelevanten Endpunkten wie Aids-definierende Erkrankungen/Tod.

Die resultierende Korrelation zwischen Behandlungseffekt auf CD4-Zellzahl und klinischem Nutzen zeigt Abbildung 26: Behandlungen, die einen größeren Anstieg der CD4-Zellzahl bedingen, zeigen eine höhere Reduktion der Rate an klinischer Progression. Mit zwei Ausnahmen zeigen alle Studien einen Anstieg der CD4-Zellzahl zusammen mit einem klinischen Vorteil; das Ausmaß des klinischen Vorteils war assoziiert mit dem Anstieg der CD4-Zellzahl.

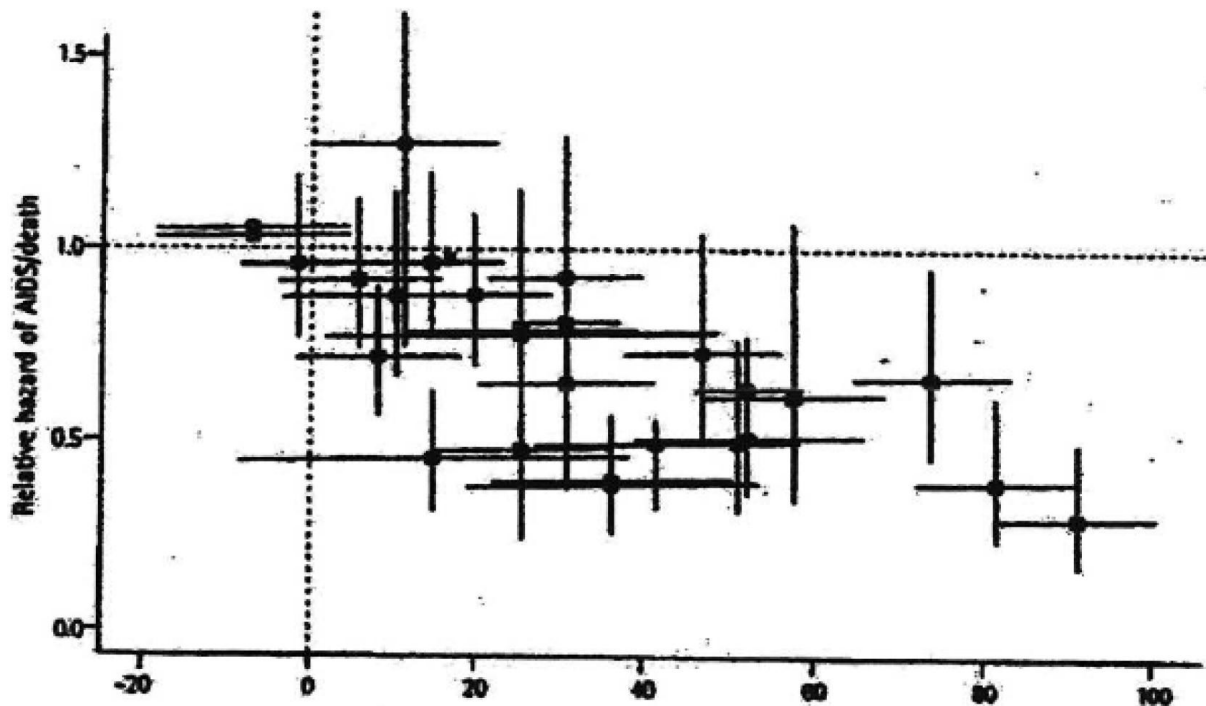


Abbildung 26: Relation zwischen Veränderung der CD4-Zellzahl und der Hazard für klinische Ereignisse (Aids/Tod) [73]

Die CD4-Zellzahl ist ein Surrogat für den immunologischen Status, das zuverlässig die Abschätzung des individuellen Risikos, an Aids zu erkranken, erlaubt und damit den primären Marker für den Beginn einer ART darstellt: Diese sollte spätestens bei einer CD4-Zellzahl von 350 CD4-Zellen/ μL begonnen werden [2]; ab einer CD4-Zellzahl von <200 CD4-Zellen/ μL muss mit zunehmender Dauer der Immunsuppression vermehrt mit dem Auftreten von Aids-definierenden Erkrankungen gerechnet werden [74]. Damit ist die Dynamik des CD4-Zellanstiegs gerade zu Beginn einer ART klinisch hochrelevant, da die Patienten bei schnellerem Anstieg der CD4-Zellzahl auch rascher oberhalb des potentiellen Risiko-Bereichs für die Entwicklung Aids-definierender Erkrankungen liegen. Die raschere Rekonstitution des Immunsystems ist somit direkt patientenrelevant.

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 GS-US-236-0102

48 Wochen

A Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Efavirenz/

Emtricitabine/ Tenofovir Disoproxil Fumarate in HIV-1 Infected, Antiretroviral Treatment-Naive Adults

Quellen:

- Studienbericht [90]
- Publikation [92]
- Registereintrag [84, 85]
- Post-hoc Analysen [119, 144]

96 Wochen

A Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Efavirenz/Emtricitabine/ Tenofovir Disoproxil Fumarate in HIV-1 Infected, Antiretroviral Treatment-Naive Adults

Quellen:

- Studienbericht [91]
- Publikation [145]
- Registereintrag [84, 85]
- Post-hoc Analysen [120, 146]

Studie GS-US-236-0103

48 Wochen

A Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Ritonavir-Boosted Atazanavir Plus Emtricitabine/Tenofovir Disoproxil Fumarate in HIV-1 Infected, Antiretroviral Treatment-Naive Adults

Quellen:

- Studienbericht [95]
- Publikation [97]
- Registereintrag [86, 87]
- Post-hoc Analysen [98, 100, 102]

96 Wochen

A Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Ritonavir-

Boosted Atazanavir Plus Emtricitabine/Tenofovir Disoproxil Fumarate in HIV-1 Infected, Antiretroviral Treatment-Naive Adults

Quellen:

- Studienbericht [96]
- Registereintrag [86, 87]
- Post-hoc Analysen [99, 101, 103]

Studie GS-US-236-0104

48 und 60 Wochen

A Phase 2, Randomized, Double-Blinded Study of the Safety and Efficacy of Elvitegravir/Emtricitabine/ Tenofovir Disoproxil Fumarate/GS-9350 Versus Atripla[®] (Efavirenz 600 mg/ Emtricitabine 200 mg/ Tenofovir Disoproxil Fumarate 300 mg) in HIV-1 Infected, Antiretroviral Treatment-Naive Adults

Quellen:

- Studienbericht [93]
- Publikation [94]
- Registereintrag [88, 89]
- Post-hoc Analysen [121, 147, 148]

Studie STARTMRK

48 Wochen

A Multicenter, Double-Blind, Randomized, Active-Controlled Study to Evaluate the Safety and Antiretroviral Activity of MK-0518 Versus Efavirenz in Treatment Naive HIV-Infected Patients, Each in Combination With TRUVADA[™]

Quellen:

- Publikation [112]
- Registereintrag [114-116]

96 Wochen

A Multicenter, Double-Blind, Randomized, Active-Controlled Study to Evaluate the Safety and Antiretroviral Activity of MK-0518 Versus Efavirenz in Treatment Naive HIV-Infected Patients, Each in Combination With TRUVADA[™]

Quellen:

- Publikation [111, 113, 117]
- Registereintrag [114-116]

GS-US-236-0123 (zur Unterstützung der Übertragbarkeit der Daten, nicht aber zur Nutzenbewertung herangezogen)

A Phase 3B Open-Label Pilot Study to Evaluate Switching from a Regimen Consisting of Raltegravir plus Emtricitabine/Tenofovir Disoproxil Fumarate Fixed-Dose Combination (FTC/TDF) to the Elvitegravir/ Cobicistat/ Emtricitabine/ Tenofovir Disoproxil Fumarate (EVG/COBI/FTC/TDF) Single-Tablet Regimen (STR) in Virologically-Suppressed, HIV-1 Infected Patients

Quellen:

- Studienprotokoll [122]
- Publikation [125, 126]
- Registereinträge [123, 124]
- Tables, Listings and Figures zu Woche 12 [127]

4.7 Referenzliste

Benennen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie in Modul 4 angegeben haben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard).

- [1] Gilead Sciences GmbH. Fachinformation Stribild 150 mg/150 mg/200 mg/245 mg Filmtabletten (Elvitegravir/Cobicistat/Emtricitabin/Tenofovirdisoproxil). Stand der Information: Mai 2013. 2013.
- [2] Deutsche AIDS-Gesellschaft (DAIG). Deutsch-Österreichische Leitlinien zur antiretroviralen Therapie der HIV-1-Infektion. AWMF-Register-Nr.: 055-001. Version 04_2012.2 vom 22.2. 2012. URL: <http://www.daignet.de/site-content/hiv-therapie/leitlinien-1/LL%20ART%20aktuell.pdf> (aufgerufen am: 26.03.2013).
- [3] European AIDS Clinical Society (EACS). Guidelines Version 6.1. November 2012. URL: <http://www.europeanaidsclinicalsociety.org/images/stories/EACS-Pdf/EacsGuidelines-v6.1-2edition.pdf> (aufgerufen am: 26.03.2013).
- [4] Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Department of Health and Human Services 2013. URL: <http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf> (aufgerufen am: 26.03.2013).
- [5] Thompson MA, Aberg JA, Hoy JF, Telenti A, Benson C, Cahn P, et al. Antiretroviral Treatment of Adult HIV Infection: 2012 Recommendations of the International Antiviral Society-USA Panel. *Jama*. 2012 Jul 25;308(4):387-402.

- [6] Williams I, Churchill D, Anderson J, Boffito M, Bower M, Cairns G, et al. British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2012. HIV Med. 2012 Sep;13 Suppl 2:1-85.
- [7] European Medicines Agency. Concept paper on extrapolation of efficacy and safety in medicine development. EMA/129698/2012. 2013.
- [8] Boehringer Ingelheim International GmbH. Fachinformation Aptivus 250 mg Weichkapseln (Tipranavir). Stand der Information: Dezember 2011. 2011.
- [9] Gilead Sciences GmbH. Fachinformation Atripla 600 mg/200 mg/245 mg Filmtabletten (Efavirenz/Emtricitabin/Tenofovirdisoproxil). Stand der Information: Januar 2013. 2013.
- [10] ViiV Healthcare GmbH. Fachinformation Celsentri 300 mg Filmtabletten (Maraviroc). Stand der Information: Februar 2013. 2013.
- [11] ViiV Healthcare GmbH. Fachinformation Combivir 150 mg/300 mg Filmtabletten (Lamivudin/Zidovudin). Stand der Information: Mai 2011. 2011.
- [12] MSD Sharp & Dohme GmbH. Fachinformation Crixivan 200 mg/400 mg Hartkapseln (Indinavir). Stand der Information: Februar 2012. 2012.
- [13] Gilead Sciences GmbH. Fachinformation Emtriva 200 mg Hartkapseln (Emtricitabin). Stand der Information: Juli 2011. 2011.
- [14] ViiV Healthcare GmbH. Fachinformation Epivir 300 mg Filmtabletten (Lamivudin). Stand der Information: Oktober 2010. 2010.
- [15] Roche Pharma AG. Fachinformation Fuzeon 90 mg/ml Pulver und Lösungsmittel zur Herstellung einer Injektionslösung (Enfuvirtid). Stand der Information: April 2010. 2010.
- [16] Janssen-Cilag International NV. Fachinformation Intelence 200 mg Tabletten (Etravirin). Stand der Information: März 2012. 2012.
- [17] Roche Pharma AG. Fachinformation Invirase 500 mg Filmtabletten (Saquinavir). Stand der Information: Januar 2013. 2013.
- [18] MSD Sharp & Dohme GmbH. Fachinformation Isentress 400 mg Filmtabletten (Raltegravir). Stand der Information: Februar 2013. 2013.
- [19] AbbVie Ltd. Fachinformation Kaletra 200 mg/50 mg Filmtabletten (Lopinavir/Ritonavir). Stand der Information: Oktober 2012. 2012.
- [20] ViiV Healthcare GmbH. Fachinformation Kivexa 600 mg/300 mg Filmtabletten (Abacavir/Lamivudin). Stand der Information: Oktober 2011. 2011.
- [21] AbbVie Ltd. Fachinformation Norvir 100 mg Filmtabletten (Ritonavir). Stand der Information: September 2012. 2012.
- [22] Janssen-Cilag International NV. Fachinformation Prezista 600 mg Filmtabletten (Darunavir). Stand der Information: Januar 2013. 2013.
- [23] Janssen-Cilag International NV. Fachinformation Prezista 800 mg Filmtabletten (Darunavir). Stand der Information: Januar 2013. 2013.
- [24] ViiV Healthcare GmbH. Fachinformation Retrovir 250 mg Hartkapseln (Zidovudin). Stand der Information: September 2012. 2012.
- [25] Bristol-Myers Squibb GmbH & Co. KGaA. Fachinformation Reyataz 150 mg/200 mg/300 mg Hartkapseln (Atazanavir). Stand der Information: August 2012. 2012.
- [26] Bristol-Myers Squibb GmbH & Co. KGaA. Fachinformation Sustiva 600 mg Filmtabletten (Efavirenz). Stand der Information: Juni 2012. 2012.
- [27] ViiV Healthcare GmbH. Fachinformation Telzir 700 mg Filmtabletten (Fosamprenavir). Stand der Information: November 2012. 2012.

- [28] ViiV Healthcare GmbH. Fachinformation Trizivir 300 mg/150 mg/300 mg Filmtabletten (Abacavir/Lamivudin/Zidovudin). Stand der Information: Mai 2012. 2012.
- [29] Gilead Sciences GmbH. Fachinformation Truvada 200 mg/245 mg Filmtabletten (Emtricitabin/Tenofovirdisoproxil). Stand der Information: Juli 2011. 2011.
- [30] Bristol-Myers Squibb GmbH & Co. KGaA. Fachinformation Videx 125 mg/200 mg/250 mg/400 mg magensaftresistente Hartkapseln (Didanosin). Stand der Information: März 2009. 2009.
- [31] Boehringer Ingelheim International GmbH. Fachinformation Viramune 200 mg Tabletten (Nevirapin). Stand der Information: Dezember 2012. 2012.
- [32] Boehringer Ingelheim International GmbH. Fachinformation Viramune 400 mg Retardtabletten (Nevirapin). Stand der Information: Dezember 2012. 2012.
- [33] Gilead Sciences GmbH. Fachinformation Viread 245 mg Filmtabletten (Tenofovirdisoproxil). Stand der Information: November 2012. 2012.
- [34] Bristol-Myers Squibb GmbH & Co. KGaA. Fachinformation Zerit 20 mg/30 mg/40 mg Hartkapseln (Stavudin). Stand der Information: April 2011. 2011.
- [35] ViiV Healthcare GmbH. Fachinformation Ziagen 300 mg Filmtabletten (Abacavir). Stand der Information: Oktober 2011. 2011.
- [36] Glass TR, Battegay M, Cavassini M, De Geest S, Furrer H, Vernazza PL, et al. Longitudinal analysis of patterns and predictors of changes in self-reported adherence to antiretroviral therapy: Swiss HIV Cohort Study. *J Acquir Immune Defic Syndr*. 2010 Jun;54(2):197-203.
- [37] Kleeberger CA, Buechner J, Palella F, Detels R, Riddler S, Godfrey R, et al. Changes in adherence to highly active antiretroviral therapy medications in the Multicenter AIDS Cohort Study. *Aids*. 2004 Mar 5;18(4):683-8.
- [38] Bangsberg DR, Perry S, Charlebois ED, Clark RA, Roberston M, Zolopa AR, et al. Non-adherence to highly active antiretroviral therapy predicts progression to AIDS. *Aids*. 2001 Jun 15;15(9):1181-3.
- [39] Kitahata MM, Reed SD, Dillingham PW, Van Rompaey SE, Young AA, Harrington RD, et al. Pharmacy-based assessment of adherence to HAART predicts virologic and immunologic treatment response and clinical progression to AIDS and death. *Int J STD AIDS*. 2004 Dec;15(12):803-10.
- [40] Mannheimer S, Friedland G, Matts J, Child C, Chesney M. The consistency of adherence to antiretroviral therapy predicts biologic outcomes for human immunodeficiency virus-infected persons in clinical trials. *Clin Infect Dis*. 2002 Apr 15;34(8):1115-21.
- [41] Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squier C, et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Annals of internal medicine*. 2000 Jul 4;133(1):21-30.
- [42] Bae JW, Guyer W, Grimm K, Altice FL. Medication persistence in the treatment of HIV infection: a review of the literature and implications for future clinical care and research. *Aids*. 2011 Jan 28;25(3):279-90.
- [43] The Antiretroviral Therapy Cohort Collaboration (ART-CC). Durability of first ART regimen and risk factors for modification, interruption or death in HIV-positive patients starting ART in Europe and N.America 2002-2009. *Aids*. 2012 Nov 28.
- [44] Elzi L, Erb S, Furrer H, Ledergerber B, Cavassini M, Hirschel B, et al. Choice of Initial Combination Antiretroviral Therapy in Individuals With HIV Infection: Determinants and Outcomes. *Archives of internal medicine*. 2012 Aug 13:1-9.

- [45] Mocroft A, Ruiz L, Reiss P, Ledergerber B, Katlama C, Lazzarin A, et al. Virological rebound after suppression on highly active antiretroviral therapy. *Aids*. 2003 Aug 15;17(12):1741-51.
- [46] Rosenblum M, Deeks SG, van der Laan M, Bangsberg DR. The risk of virologic failure decreases with duration of HIV suppression, at greater than 50% adherence to antiretroviral therapy. *PLoS One*. 2009;4(9):e7196.
- [47] Hill A, Balkin A. Risk factors for gastrointestinal adverse events in HIV treated and untreated patients. *AIDS reviews*. 2009 Jan-Mar;11(1):30-8.
- [48] Zolopa A, Rockstroh JK, Orkin C, Stellbrink HJ, Walmsley S, Cooper D, et al. Efficacy and Safety of Elvitegravir/Cobicistat/Emtricitabine/Tenofovir DF Integrated Week 96 Analyses. 20th Conference on Retroviruses and Opportunistic Infections (CROI); Atlanta, Georgia, USA.2013.
- [49] Atkinson MJ, Petrozzino JJ. An evidence-based review of treatment-related determinants of patients' nonadherence to HIV medications. *AIDS Patient Care STDS*. 2009 Nov;23(11):903-14.
- [50] Lenzi L, Wiens A, Pontarolo R. Evaluation of adverse events associated with antiretroviral therapy and the relationship to treatment adherence. *International journal of clinical pharmacology and therapeutics*. 2013 Feb;51(2):141-6.
- [51] Willig JH, Abrams S, Westfall AO, Routman J, Adusumilli S, Varshney M, et al. Increased regimen durability in the era of once-daily fixed-dose combination antiretroviral therapy. *Aids*. 2008 Oct 1;22(15):1951-60.
- [52] Parienti JJ, Bangsberg DR, Verdon R, Gardner EM. Better adherence with once-daily antiretroviral regimens: a meta-analysis. *Clin Infect Dis*. 2009 Feb 15;48(4):484-8.
- [53] Antinori A, Angeletti C, Ammassari A, Sangiorgi D, Giannetti A, Buda S, et al. Adherence in HIV-positive patients treated with single tablet regimens and multiple-pill regimens: findings from the COMPACT study. 11th International Congress on Drug Therapy in HIV Infection; Glasgow 2012. p. P14.
- [54] Murri R, Fregosi S, Bortolussi P, Marcotullio S, Maggiolo F. Among once-daily regimens, STR are associated with better adherence. 11th International Congress on Drug Therapy in HIV Infection; Glasgow 2012. p. P016.
- [55] Cohen C, Davis DL, Meyers JL. Association of partial adherence to antiretroviral therapy with hospitalizations and healthcare costs in an HIV population. 11th International Congress on Drug Therapy in HIV Infection; Glasgow 2012. p. P001.
- [56] Vera J, Aragão F, Guimarães M, Vaz Pinto I. Benefits of ART simplification on adherence, clinical and economic outcomes. 11th International Congress on Drug Therapy in HIV Infection; Glasgow 2012. p. P005.
- [57] Maggiolo F, Fregosi S, Bortolussi P, Marcotullio S, Murri R, editors. Single Tablet Regimen (STR) enhance patients' acceptability of HAART. 11th International Congress on Drug Therapy in HIV Infection; 2012; Glasgow
- [58] Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV. Beratungsanforderung 2013-B-017. Kombination Elvitegravir/Cobicistat/ Emtricitabin/ Tenofoviridisoproxil zur Behandlung von Infektionen mit dem Humanen Immundefizienzvirus 1 (HIV-1) bei Erwachsenen im Alter von 18 Jahren und darüber, bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe von Stribild® assoziiert sind. 2013 29.05.2013.
- [59] Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV. Beratungsanforderung 2012-B-063. Kombination Elvitegravir und

- Cobicistat und Emtricitabin und Tenofoviridisoproxil zur Behandlung von Infektionen mit dem Humanen Immundefizienz-Virus Typ 1 (HIV-1) bei antiretroviral nichtvorbehandelten erwachsenen Patienten. 2013 21.02.2013.
- [60] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Pixantron – Nutzenbewertung gemäß § 35a SGB V. IQWiG-Berichte – Nr. 154. . 2013.
- [61] European Medicines Agency. Guideline on the Clinical Development of Medicinal Products for the Treatment of HIV Infection. EMEA/CPMP/EWP/633/02. 2008.
- [62] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Rilpivirin – Nutzenbewertung gemäß § 35a SGB V. IQWiG-Berichte – Nr. 127. . 2012.
- [63] Hughes MD. The evaluation of surrogate endpoints in practice: experience in HIV. In: Burzykowski T, Molenberghs G, Buyse M, Hrsg. The evaluation of surrogate endpoints. New York: Springer; 2005. S. 295-321.
- [64] Mellors JW, Munoz A, Giorgi JV, Margolick JB, Tassoni CJ, Gupta P, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Annals of internal medicine*. 1997 Jun 15;126(12):946-54.
- [65] Juday T, Grimm K, Zoe-Powers A, Willig J, Kim E. A retrospective study of HIV antiretroviral treatment persistence in a commercially insured population in the United States. *AIDS Care*. 2011 Sep;23(9):1154-62.
- [66] Llibre JM, Arribas JR, Domingo P, Gatell JM, Lozano F, Santos JR, et al. Clinical implications of fixed-dose coformulations of antiretrovirals on the outcome of HIV-1 therapy. *Aids*. 2011 Sep 10;25(14):1683-90.
- [67] Strategies for Management of Antiretroviral Therapy Study Group, El-Sadr WM, Lundgren J, Neaton JD, Gordin F, Abrams D, et al. CD4+ count-guided interruption of antiretroviral treatment. *The New England journal of medicine*. 2006 Nov 30;355(22):2283-96.
- [68] Madruga JV, Cassetti I, Etzel A, Suleiman J, Zhou Y, Warren DR, et al. The 10-Year Safety and Efficacy of Tenofovir DF (TDF)-Containing HAART in Antiretroviral Naïve Patients. Poster #P86. 10th International Congress on Drug Therapy in HIV Infection (HIV-10); Glasgow, UK.2010.
- [69] Benzie AA, Bansi LK, Sabin CA, Portsmouth S, Hill T, Johnson M, et al. Increased duration of viral suppression is associated with lower viral rebound rates in patients with previous treatment failures. *Aids*. 2007 Jul 11;21(11):1423-30.
- [70] Smith CJ, Phillips AN, Dauer B, Johnson MA, Lampe FC, Youle MS, et al. Factors associated with viral rebound among highly treatment-experienced HIV-positive patients who have achieved viral suppression. *HIV Med*. 2009 Jan;10(1):19-27.
- [71] Gemeinsamer Bundesausschuss. HIV-Resistenztestung. Unterausschuss „Ärztliche Behandlung“ des Gemeinsamen Bundesausschusses, 2004.
- [72] Marschner IC, Collier AC, Coombs RW, D'Aquila RT, DeGruttola V, Fischl MA, et al. Use of changes in plasma levels of human immunodeficiency virus type 1 RNA to assess the clinical benefit of antiretroviral therapy. *J Infect Dis*. 1998 Jan;177(1):40-7.
- [73] Hill AM, DeMasi R, Dawson D. Meta-analysis of antiretroviral effects on HIV-1 RNA, CD4 cell count and progression to AIDS or death. *Antivir Ther*. 1998;3(3):139-45.
- [74] Hoffmann C. 6.11. Monitoring der ART. In: Hoffmann C, Rockstroh JK, Hrsg. HIV 2012: Medizin Fokus Verlag; 2012. S. 171-95.
- [75] Lange S, Thomas Kaiser T, Beatrice-Schüler Y, Skipka G, Volker Vervölgyi V, Wieseler B. Bewertung der klinischen Relevanz bei der Nutzenbewertung. Ein formalisiertes Vorgehen zur Bewertung der Relevanz von Gruppenunterschieden auf

- Skalen. 2010. URL: [https://www.iqwig.de/download/IQWiG im Dialog 2010 Stefan Lange.pdf](https://www.iqwig.de/download/IQWiG_im_Dialog_2010_Stefan_Lange.pdf) (aufgerufen am: 13.03.2013).
- [76] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden Version 4.0. 2011. URL: [https://www.iqwig.de/download/IQWiG Methoden Version 4 0.pdf](https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf) (aufgerufen am: 08.05.2013).
- [77] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj*. 2003 Sep 6;327(7414):557-60.
- [78] Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. 2008.
- [79] Qaqish R, J van Wyk J, MS King MS. A comparison of the FDA TLOVR and FDA Snapshot algorithms based on studies evaluating once-daily vs. twice daily lopinavir/ritonavir (LPV/r) regimens. Tenth International Congress on Drug Therapy in HIV Infection. Glasgow, UK. *Journal of the International AIDS Society*. 2010;13 (Suppl 4):P58.
- [80] Hoffmann C. 6. ART. In: Hoffmann C, Rockstroh JK, Hrsg. *HIV 2012: Medizin Fokus Verlag*; 2012. S. 60-263.
- [81] De Wit S, Sabin CA, Weber R, Worm SW, Reiss P, Cazanova C, et al. Incidence and risk factors for new-onset diabetes in HIV-infected patients: the Data Collection on Adverse Events of Anti-HIV Drugs (D:A:D) study. *Diabetes care*. 2008 Jun;31(6):1224-9.
- [82] Tedaldi EM, Absalon J, Thomas AJ, Shlay JC, van den Berg-Wolf M. Ethnicity, race, and gender. Differences in serious adverse events among participants in an antiretroviral initiation trial: results of CPCRA 058 (FIRST Study). *J Acquir Immune Defic Syndr*. 2008 Apr 1;47(4):441-8.
- [83] Glenny AM, Altman DG, Song F, Sakarovich C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health technology assessment*. 2005 Jul;9(26):1-134, iii-iv.
- [84] ClinicalTrials.gov NCT01095796. Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Stribild Versus Atripla in Human Immunodeficiency Virus, Type 1 (HIV-1) Infected, Antiretroviral Treatment-Naive Adults. 2012 [07.05.2013]; Verfügbar unter: <http://ClinicalTrials.gov/show/NCT01095796>.
- [85] ICTRP NCT01095796. Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Stribild Versus Atripla in Human Immunodeficiency Virus, Type 1 (HIV-1) Infected, Antiretroviral Treatment-Naive Adults. 2012 [07.05.2013]; Verfügbar unter: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01095796>.
- [86] ClinicalTrials.gov NCT01106586. Study to Evaluate the Safety and Efficacy of Stribild Versus Ritonavir-Boosted Atazanavir Plus Truvada in Human Immunodeficiency Virus, Type 1 (HIV-1) Infected, Antiretroviral Treatment-Naive Adults. 2012 [07.05.2013]; Verfügbar unter: <http://www.clinicaltrials.gov/ct2/show/NCT01106586>.
- [87] ICTRP EUCTR2009-016758-42-DE. Clinical research study involving an experimental combination medication, i.e. one pill containing two experimental medications EVG and COBI and two approved medications for the treatment of HIV-1 infection, FTC and TDF, referred as EVG/COBI/FTC/TDF. This is a study that neither the patient or the investigator will know whether the patient is receiving

- EVG/COBI/FTC/TDF or the comparator drug ritonavir-boosted atazanavir with FTC/TDF. This is a randomized (by chance, like a flip of a coin) study. 2012 [07.05.2013]; Verfügbar unter: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2009-016758-42-DE>.
- [88] ClinicalTrials.gov NCT00869557. Study of the Safety and Efficacy of Stribild Versus Atripla in Human Immunodeficiency Virus, Type 1 (HIV-1) Infected, Antiretroviral Treatment-Naive Adults. 2012 [07.05.2013]; Verfügbar unter: <http://ClinicalTrials.gov/show/NCT00869557>.
- [89] ICTRP NCT00869557. Study of the Safety and Efficacy of Stribild Versus Atripla in Human Immunodeficiency Virus, Type 1 (HIV-1) Infected, Antiretroviral Treatment-Naive Adults. 2012 [07.05.2013]; Verfügbar unter: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00869557>.
- [90] Gilead Sciences Inc. GS-US-236-0102: A Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Efavirenz/Emtricitabine/Tenofovir Disoproxil Fumarate in HIV-1 Infected, Antiretroviral Treatment-Naive Adults (Interim Week 48 Clinical Study Report). 2011. Report Nr.: GS-US-236-0102.
- [91] Gilead Sciences Inc. GS-US-236-0102: A Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Efavirenz/Emtricitabine/Tenofovir Disoproxil Fumarate in HIV-1 Infected, Antiretroviral Treatment-Naive Adults (Interim Week 96 Clinical Study Report). 2012. Report Nr.: GS-US-236-0102.
- [92] Sax PE, DeJesus E, Mills A, Zolopa A, Cohen C, Wohl D, et al. Co-formulated elvitegravir, cobicistat, emtricitabine, and tenofovir versus co-formulated efavirenz, emtricitabine, and tenofovir for initial treatment of HIV-1 infection: a randomised, double-blind, phase 3 trial, analysis of results after 48 weeks. *Lancet*. 2012 Jun 30;379(9835):2439-48.
- [93] Gilead Sciences Inc. GS-US-236-0104: A Phase 2, Randomized, Double-Blinded Study of the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Atripla[®] (Efavirenz 600 mg/Emtricitabine 200 mg/Tenofovir Disoproxil Fumarate 300 mg) in HIV-1 Infected, Antiretroviral Treatment-Naive Adults (Interim Week 96 Clinical Study Report). 2011. Report Nr.: GS-US-236-0104.
- [94] Cohen C, Elion R, Ruane P, Shamblaw D, DeJesus E, Rashbaum B, et al. Randomized, phase 2 evaluation of two single-tablet regimens elvitegravir/cobicistat/emtricitabine/tenofovir disoproxil fumarate versus efavirenz/emtricitabine/tenofovir disoproxil fumarate for the initial treatment of HIV infection. *Aids*. 2011 Mar 27;25(6):F7-12.
- [95] Gilead Sciences Inc. GS-US-236-0103: A Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Ritonavir-Boosted Atazanavir Plus Tenofovir Disoproxil Fumarate in HIV-1 Infected, Antiretroviral Treatment-Naive Adults (Interim Week 48 Clinical Study Report). 2011. Report Nr.: GS-US-236-0103.
- [96] Gilead Sciences Inc. GS-US-236-0103: A Phase 3, Randomized, Double-Blind Study to Evaluate the Safety and Efficacy of Elvitegravir/Emtricitabine/Tenofovir Disoproxil Fumarate/GS-9350 Versus Ritonavir-Boosted Atazanavir Plus Tenofovir Disoproxil Fumarate in HIV-1 Infected, Antiretroviral Treatment-Naive Adults (Interim Week 96 Clinical Study Report). 2012. Report Nr.: GS-US-236-0103.

- [97] DeJesus E, Rockstroh JK, Henry K, Molina JM, Gathe J, Ramanathan S, et al. Co-formulated elvitegravir, cobicistat, emtricitabine, and tenofovir disoproxil fumarate versus ritonavir-boosted atazanavir plus co-formulated emtricitabine and tenofovir disoproxil fumarate for initial treatment of HIV-1 infection: a randomised, double-blind, phase 3, non-inferiority trial. *Lancet*. 2012 Jun 30;379(9835):2429-38.
- [98] Gilead Sciences Inc. Posthoc Analysis for Study GS-US-236-0103 Week 48. 2013.
- [99] Gilead Sciences Inc. Posthoc Analysis for Study GS-US-236-0103 Week 96. 2013.
- [100] Gilead Sciences Inc. Hazard Ratio for TLOVR Endpoint (5689) for Study GS-US-236-0103 Week 48. 2013.
- [101] Gilead Sciences Inc. Hazard Ratio for TLOVR Endpoint (5690) for Study GS-US-236-0103 Week 96. 2013.
- [102] Gilead Sciences Inc. adhoc Analysis (5258) for Study GS-US-236-0103 Week 48. 2013.
- [103] Gilead Sciences Inc. adhoc Analysis (5548.9) for Study GS-US-236-0103 Week 96. 2013.
- [104] Gilead Sciences GmbH. Subgruppenanalysen zu EVG/COBI/FTC/TDF im Vergleich zur zweckmäßigen Vergleichstherapie. 2013.
- [105] Rousseau FS, Kahn JO, Thompson M, Mildvan D, Shepp D, Sommadossi JP, et al. Prototype trial design for rapid dose selection of antiretroviral drugs: an example using emtricitabine (Coviracil). *J Antimicrob Chemother*. 2001 Oct;48(4):507-13.
- [106] Rousseau FS, Wakeford C, Mommeja-Marin H, Sanne I, Moxham C, Harris J, et al. Prospective randomized trial of emtricitabine versus lamivudine short-term monotherapy in human immunodeficiency virus-infected patients. *J Infect Dis*. 2003 Dec 1;188(11):1652-8.
- [107] Marcelin AG, Charpentier C, Wirden M, Landman R, Valantin MA, Simon A, et al. Resistance profiles of emtricitabine and lamivudine in tenofovir-containing regimens. *J Antimicrob Chemother*. 2012 Jun;67(6):1475-8.
- [108] Maserati R, De Silvestri A, Uglietti A, Colao G, Di Biagio A, Bruzzone B, et al. Emerging mutations at virological failure of HAART combinations containing tenofovir and lamivudine or emtricitabine. *Aids*. 2010 Apr 24;24(7):1013-8.
- [109] Borroto-Esoda K, Vela JE, Myrick F, Ray AS, Miller MD. In vitro evaluation of the anti-HIV activity and metabolic interactions of tenofovir and emtricitabine. *Antivir Ther*. 2006;11(3):377-84.
- [110] Mulato AS, Cherrington JM. Anti-HIV activity of adefovir (PMEA) and PMPA in combination with antiretroviral compounds: in vitro analyses. *Antiviral research*. 1997 Nov;36(2):91-7.
- [111] Lennox JL, DeJesus E, Berger DS, Lazzarin A, Pollard RB, Ramalho Madruga JV, et al. Raltegravir versus Efavirenz regimens in treatment-naïve HIV-1-infected patients: 96-week efficacy, durability, subgroup, safety, and metabolic analyses. *J Acquir Immune Defic Syndr*. 2010 Sep;55(1):39-48.
- [112] Lennox JL, DeJesus E, Lazzarin A, Pollard RB, Madruga JV, Berger DS, et al. Safety and efficacy of raltegravir-based versus efavirenz-based combination therapy in treatment-naïve patients with HIV-1 infection: a multicentre, double-blind randomised controlled trial. *Lancet*. 2009 Sep 5;374(9692):796-806.
- [113] Rockstroh JK, Lennox JL, DeJesus E, Saag MS, Lazzarin A, Wan H, et al. Long-term treatment with raltegravir or efavirenz combined with tenofovir/emtricitabine for treatment-naïve human immunodeficiency virus-1-infected patients: 156-week results from STARTMRK. *Clin Infect Dis*. 2011 Oct;53(8):807-16.

- [114] ClinicalTrials.gov NCT00369941. A Study to Evaluate the Safety and Antiretroviral Activity of MK-0518 Versus Efavirenz in Treatment Naïve HIV-Infected Patients, Each in Combination With TRUVADA (0518-021 EXT). 2012 [07.05.2013]; Verfügbar unter: <http://clinicaltrials.gov/ct2/show/NCT00369941>.
- [115] ICTRP EUCTR2006-003109-23-DE. A Multicenter, Double-Blind, Randomized, Active-Controlled Study to Evaluate the Safety and Antiretroviral Activity of MK-0518 Versus Efavirenz in Treatment Naïve HIV-Infected Patients, Each in Combination With TRUVADA™ - MK-0518 safety and efficacy study in treatment naïve patients. 2012 [08.05.2013]; Verfügbar unter: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2006-003109-23-DE>.
- [116] ICTRP NCT00369941. A Study to Evaluate the Safety and Antiretroviral Activity of MK0518 Versus Efavirenz in Treatment Naïve HIV-Infected Patients, Each in Combination With TRUVADA (0518-021). 2012 [07.05.2013]; Verfügbar unter: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00369941>.
- [117] Teppler H, Brown DD, Leavitt RY, Sklar P, Wan H, Xu X, et al. Long-term safety from the raltegravir clinical development program. Curr HIV Res. 2011 Jan;9(1):40-53.
- [118] Miettinen O, Nurminen M. Comparative analysis of two rates. Statistics in medicine. 1985 Apr-Jun;4(2):213-26.
- [119] Gilead Sciences Inc. Posthoc Analysis (5682) for Study GS-US-236-0102 Week 48. 2013.
- [120] Gilead Sciences Inc. Posthoc Analysis (5695) for Study GS-US-236-0102 Week 96. 2013.
- [121] Gilead Sciences Inc. Posthoc Analysis (5698) for Study GS-US-236-0104 Week 60. 2013.
- [122] Gilead Sciences Inc. A Phase 3B Open-Label Pilot Study to Evaluate Switching from a Regimen Consisting of Raltegravir plus Emtricitabine/Tenofovir Disoproxil Fumarate Fixed-Dose Combination (FTC/TDF) to the Elvitegravir/Cobicistat/Emtricitabine/Tenofovir Disoproxil Fumarate(EVG/COBI/FTC/TDF) Single-Tablet Regimen (STR) in Virologically-Suppressed, HIV-1 Infected Patients. 2012. Report Nr.: GS-US-236-0123.
- [123] ClinicalTrials.gov NCT01533259. A Phase 3B Open-Label Pilot Study to Evaluate Switching From a Regimen Consisting of Raltegravir Plus Emtricitabine/Tenofovir Disoproxil Fumarate Fixed-Dose Combination (FTC/TDF) to the Elvitegravir/Cobicistat/Emtricitabine/Tenofovir Disoproxil Fumarate (EVG/COBI/FTC/TDF) Single-Tablet Regimen (STR) in Virologically Suppressed, HIV-1 Infected Patients (ongoing). 2013 [16.04.2013]; Verfügbar unter: <http://clinicaltrials.gov/ct2/show/NCT01533259?term=GS-US-236-0123&rank=1>.
- [124] ICTRP NCT01533259. A Phase 3B Open-Label Pilot Study to Evaluate Switching From a Regimen Consisting of Raltegravir Plus Emtricitabine/Tenofovir Disoproxil Fumarate Fixed-Dose Combination (FTC/TDF) to the Elvitegravir/Cobicistat/Emtricitabine/Tenofovir Disoproxil Fumarate (EVG/COBI/FTC/TDF) Single-Tablet Regimen (STR) in Virologically Suppressed, HIV-1 Infected Patients. 2013 [16.04.2013]; Verfügbar unter: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01533259>.
- [125] Mills A, Crofoot G, Ortiz R, Rashbaum B, Towner W, Ward D, et al. Safety and Tolerability of Switching from Twice Daily Raltegravir plus Truvada to STRIBILD in Virologically Suppressed, HIV-1 Infected Subjects. 2012.

- [126] Mills A, Crofoot G, Ortiz R, Rashbaum B, Towner W, Ward D, et al. Phase 3b, open-label, multicenter, 48-week study of immediate switch from RAL + FTC/TDF to STB in stable, virologically controlled subjects. HIV DART 2012; San Diego, California.2012.
- [127] Gilead Sciences Inc. Tables, Figures, and Listings. 2013. Report Nr.: GS-US-236-0123.
- [128] Nicastrì E, Angeletti C, Palmisano L, Sarmati L, Chiesi A, Geraci A, et al. Gender differences in clinical progression of HIV-1-infected individuals during long-term highly active antiretroviral therapy. *Aids*. 2005 Mar 24;19(6):577-83.
- [129] Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Rilpivirin. 2012. 05.07.2012.
- [130] d'Arminio Monforte A, Testa L, Adorni F, Chiesa E, Bini T, Moscatelli GC, et al. Clinical outcome and predictive factors of failure of highly active antiretroviral therapy in antiretroviral-experienced patients in advanced stages of HIV-1 infection. *Aids*. 1998 Sep 10;12(13):1631-7.
- [131] Chene G, Sterne JA, May M, Costagliola D, Ledergerber B, Phillips AN, et al. Prognostic importance of initial response in HIV-1 infected patients starting potent antiretroviral therapy: analysis of prospective studies. *Lancet*. 2003 Aug 30;362(9385):679-86.
- [132] Courtenay-Quirk C, Wolitski RJ, Parsons JT, Gomez CA, Seropositive Urban Men's Study T. Is HIV/AIDS stigma dividing the gay community? Perceptions of HIV-positive men who have sex with men. *AIDS education and prevention : official publication of the International Society for AIDS Education*. 2006 Feb;18(1):56-67.
- [133] Moyle GJ, Sabin CA, Cartledge J, Johnson M, Wilkins E, Churchill D, et al. A randomized comparative trial of tenofovir DF or abacavir as replacement for a thymidine analogue in persons with lipodystrophy. *Aids*. 2006 Oct 24;20(16):2043-50.
- [134] UNAIDS. Stigma and Discrimination. 2003.
- [135] Cloete A, Simbayi LC, Kalichman SC, Strebel A, Henda N. Stigma and discrimination experiences of HIV-positive men who have sex with men in Cape Town, South Africa. *AIDS Care*. 2008 Oct;20(9):1105-10.
- [136] Dowshen N, Binns HJ, Garofalo R. Experiences of HIV-related stigma among young men who have sex with men. *AIDS Patient Care STDS*. 2009 May;23(5):371-6.
- [137] Emler CA. Experiences of stigma in older adults living with HIV/AIDS: a mixed-methods analysis. *AIDS Patient Care STDS*. 2007 Oct;21(10):740-52.
- [138] Rintamaki LS, Davis TC, Skripkauskas S, Bennett CL, Wolf MS. Social stigma concerns and HIV medication adherence. *AIDS Patient Care STDS*. 2006 May;20(5):359-68.
- [139] Stutterheim SE, Pryor JB, Bos AE, Hoogendijk R, Muris P, Schaalma HP. HIV-related stigma and psychological distress: the harmful effects of specific stigma manifestations in various social settings. *Aids*. 2009 Nov 13;23(17):2353-7.
- [140] Fernandes AP, Sanches RS, Mill J, Lucy D, Palha PF, Dalri MC. Lipodystrophy syndrome associated with antiretroviral therapy in HIV patients: considerations for psychosocial aspects. *Revista latino-americana de enfermagem*. 2007 Sep-Oct;15(5):1041-5.
- [141] National Centre in HIV Social Research. Living with body shape change. Positive Life NSW, Fact Sheet 5. 2003.

- [142] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Ticagrelor – Nutzenbewertung gemäß § 35a SGB V. IQWiG-Berichte – Nr. 96. . 2011.
- [143] Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *The New England journal of medicine*. 2011 Aug 11;365(6):493-505.
- [144] Gilead Sciences Inc. Hazard Ratio for TLOVR Endpoint (5682) for Study GS-US-236-0102 Week 48. 2013.
- [145] Zolopa A, Sax PE, Dejesus E, Mills A, Cohen C, Wohl D, et al. A Randomized Double-Blind Comparison of Coformulated Elvitegravir/Cobicistat/Emtricitabine/Tenofovir Disoproxil Fumarate Versus Efavirenz/Emtricitabine/Tenofovir Disoproxil Fumarate for Initial Treatment of HIV-1 Infection: Analysis of Week 96 Results. *J Acquir Immune Defic Syndr*. 2013 May 1;63(1):96-100.
- [146] Gilead Sciences Inc. Hazard Ratio for TLOVR Endpoint (5695) for Study GS-US-236-0102 Week 96. 2013.
- [147] Gilead Sciences Inc. Posthoc Analysis (5697) for Study GS-US-236-0104 Week 48. 2013.
- [148] Gilead Sciences Inc. Hazard Ratio for TLOVR Endpoint (5698) for Study GS-US-236-0104 Week 48. 2013.
- [149] Wong SS, Wilczynski NL, Haynes RB. 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*. 2006 Oct;94(4):451-5.
- [150] Haybittle JL. Repeated assessment of results in clinical trials of cancer treatment. *The British journal of radiology*. 1971 Oct;44(526):793-7.

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 ¹⁷] – 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

¹⁷ 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.

Da die erste Zulassung des zu bewertenden Arzneimittels am Tag der Einreichung des vorliegenden Nutzendossier beim G-BA nicht älter als 12 Monate ist (Datum der ersten Zulassung in den USA: 27. August 2012), ist die Durchführung einer bibliografischen Literaturrecherche für das zu bewertende Arzneimittel nicht erforderlich.

Da in der Liste der Studien des pharmazeutischen Unternehmers keine geeignete Studie für einen direkten Vergleich identifiziert werden kann, wird zur Sicherstellung/Darstellung der Datenlage und der Verfügbarkeit von RCT mit den zweckmäßigen Vergleichstherapien Efavirenz, Atazanavir (geboostert mit Ritonavir) und Raltegravir (jeweils in Kombination mit Emtricitabin/ Tenofoviridisoproxil) im Vergleich zu dem bewertenden Arzneimittel Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil eine bibliographische Literaturrecherche für einen möglichen direkten Vergleich durchgeführt.

Direkter Vergleich Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil vs. Efavirenz, Emtricitabin und Tenofoviridisoproxil

Eine Literaturrecherche zur zweckmäßigen Vergleichstherapie unter der Fragestellung, wie in Abschnitt 4.2.3.2 beschrieben, wurde durchgeführt. Die Suchstrategien sind in Tabelle 4-197, Tabelle 4-198 und Tabelle 4-199 dargestellt. Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Die Suche wurde für Abschnitt 4.3.1.1.2 durchgeführt.

Datenbankname	EMBASE
Suchoberfläche	Ovid
Datum der Suche	08.04.2013
Zeitsegment	1947 to 2013 Week 14
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- Strategy minimizing difference between sensitivity and specificity [149]

Tabelle 4-197: Ovid: Suchstrategie EMBASE 1947 to 2013 Week 14

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	259525
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	359656
3	immun* Vir*.mp.	307437
4	acquire* immun* deficienc* syndrom*.mp.	121537
5	1 or 2 or 3 or 4	389357
6	adult*.mp.	4974324

#	Suchbegriffe	Ergebnis
7	5 and 6	112461
8	Efavirenz.mp. or exp efavirenz/	12476
9	EFV.mp.	899
10	(sustiva or stocrin).mp.	1123
11	(DMP 266 or DMP-266).mp.	110
12	DMP266.mp.	1
13	154598-52-4.rn.	11425
14	8 or 9 or 10 or 11 or 12 or 13	12684
15	random*.tw.	816043
16	double-blind*.tw.	145266
17	placebo*.mp.	314957
18	15 or 16 or 17	1025005
19	7 and 14	4838
20	18 and 19	651

Datenbankname MEDLINE(R) without Revisions
Suchoberfläche Ovid
Datum der Suche 08.04.2013
Zeitsegment 1946 to March Week 4 2013
Suchfilter Filter für randomisierte kontrollierte Studien nach Wong 2006-
Strategy minimizing difference between sensitivity and specificity
[149]

Tabelle 4-198: Ovid: Suchstrategie MEDLINE(R) without Revisions to March Week 4 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	237834
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	168215
3	immun* Vir*.mp.	75993
4	mensch* immun* vir*.mp.	1
5	erworb* immun*.mp.	99

#	Suchbegriffe	Ergebnis
6	acquire* immun* deficienc* syndrom*.mp.	4874
7	1 or 2 or 3 or 4 or 5 or 6	281856
8	adult*.mp.	4134725
9	erwachs*.mp.	3029
10	8 or 9	4134816
11	7 and 10	93675
12	Efavirenz.mp. or exp efavirenz/	2471
13	EFV.mp.	478
14	(DMP 266 or DMP-266).mp.	53
15	DMP266.mp.	2
16	(sustiva or stocrin).mp.	129
17	154598-52-4.rn.	0
18	12 or 13 or 14 or 15 or 16 or 17	2564
19	randomized controlled trial.pt.	340089
20	randomized.mp.	488479
21	placebo.mp.	139670
22	19 or 20 or 21	531204
23	11 and 18	1056
24	22 and 23	305

Datenbankname Cochrane
EBM Reviews - Cochrane Central Register of Controlled Trials
March 2013

Suchoberfläche Ovid

Datum der Suche 08.04.2013

Zeitsegment 1991 to March 2013

Suchfilter kein Filter

Tabelle 4-199: Ovid: Suchstrategie Cochrane 1991 to March 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	7831
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human	4536

#	Suchbegriffe	Ergebnis
	immunodeficiency virus/	
3	immun* Vir*.mp.	2626
4	mensch* immun* vir*.mp.	0
5	erworb* immun*.mp.	3
6	acquire* immun* deficienc* syndrom*.mp.	250
7	1 or 2 or 3 or 4 or 5 or 6	8492
8	adult*.mp.	276993
9	erwachs*.mp.	61
10	8 or 9	277000
11	7 and 10	5477
12	Efavirenz.mp. or exp efavirenz/	368
13	EFV.mp.	120
14	(DMP 266 or DMP-266).mp.	0
15	DMP266.mp.	0
16	(sustiva or stocrin).mp.	2
17	[154598-52-4.rn.]	0
18	12 or 13 or 14 or 15 or 16 or 17	377
19	11 and 18	267

Direkter Vergleich Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil vs. Atazanavir, Emtricitabin und Tenofoviridisoproxil

Eine Literaturrecherche zur zweckmäßigen Vergleichstherapie unter der Fragestellung, wie in Abschnitt 4.2.3.2 beschrieben, wurde durchgeführt. Die Suchstrategien sind in Tabelle 4-200, Tabelle 4-201 und Tabelle 4-202 dargestellt. Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Die Suche wurde für Abschnitt 4.3.1.1.2 durchgeführt.

Datenbankname	EMBASE
Suchoberfläche	Ovid
Datum der Suche	08.04.2013
Zeitsegment	1947 to 2013 Week 14
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- Strategy minimizing difference between sensitivity and specificity [149]

Tabelle 4-200: Ovid: Suchstrategie EMBASE 1947 to 2013 Week 14

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	259525
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	359656
3	immun* Vir*.mp.	307437
4	acquire* immun* deficienc* syndrom*.mp.	121537
5	1 or 2 or 3 or 4	389357
6	adult*.mp.	4974324
7	5 and 6	112461
8	atazanavir.mp. or exp atazanavir/	4961
9	(ATV or ATZ).mp.	1326
10	(BMS-232632 or BMS 232632).mp.	104
11	BMS232632.mp.	0
12	Reyataz.mp.	512
13	198904-31-3.rn.	4169
14	8 or 9 or 10 or 11 or 12 or 13	5942
15	random*.tw.	816043
16	double-blind*.tw.	145266
17	placebo*.mp.	314957
18	15 or 16 or 17	1025005
19	7 and 14	1590
20	18 and 19	245

Datenbankname MEDLINE(R) without Revisions

Suchoberfläche Ovid

Datum der Suche 08.04.2013

Zeitsegment 1946 to March Week 4 2013

Suchfilter Filter für randomisierte kontrollierte Studien nach Wong 2006-
Strategy minimizing difference between sensitivity and specificity
[149]

Tabelle 4-201: Ovid: Suchstrategie MEDLINE(R) without Revisions to March Week 4 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	237834
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	168215
3	immun* Vir*.mp.	75993
4	mensch* immun* vir*.mp.	1
5	erworb* immun*.mp.	99
6	acquire* immun* deficienc* syndrom*.mp.	4874
7	1 or 2 or 3 or 4 or 5 or 6	281856
8	adult*.mp.	4134725
9	erwachs*.mp.	3029
10	8 or 9	4134816
11	7 and 10	93675
12	atazanavir.mp. or exp atazanavir/	926
13	(ATV or ATZ).mp.	788
14	(BMS-232632 or BMS 232632).mp.	13
15	BMS232632.mp.	1
16	Reyataz.mp.	24
17	198904-31-3.rn.	0
18	12 or 13 or 14 or 15 or 16 or 17	1537
19	randomized controlled trial.pt.	340089
20	randomized.mp.	488479
21	placebo.mp.	139670
22	19 or 20 or 21	531204
23	11 and 18	372
24	22 and 23	104

Datenbankname Cochrane
EBM Reviews - Cochrane Central Register of Controlled Trials
March 2013

Suchoberfläche Ovid

Datum der Suche 08.04.2013

Zeitsegment 1991 to March 2013

Suchfilter kein Filter

Tabelle 4-202: Ovid: Suchstrategie Cochrane 1991 to March 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	7831
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	4536
3	immun* Vir*.mp.	2626
4	mensch* immun* vir*.mp.	0
5	erworb* immun*.mp.	3
6	acquire* immun* deficienc* syndrom*.mp.	250
7	1 or 2 or 3 or 4 or 5 or 6	8492
8	adult*.mp.	276993
9	erwachs*.mp.	61
10	8 or 9	277000
11	7 and 10	5477
12	atazanavir.mp. or exp atazanavir/	122
13	(ATV or ATZ).mp.	88
14	(BMS-232632 or BMS 232632).mp.	0
15	BMS232632.mp.	0
16	Reyataz.mp.	1
17	[198904-31-3.rn.]	0
18	12 or 13 or 14 or 15 or 16 or 17	156
19	11 and 18	95

Direkter Vergleich Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil vs. Raltegravir, Emtricitabin und Tenofoviridisoproxil

Eine Literaturrecherche zur zweckmäßigen Vergleichstherapie unter der Fragestellung, wie in Abschnitt 4.2.3.2 beschrieben, wurde durchgeführt. Die Suchstrategien sind in Tabelle 4-203, Tabelle 4-204 und Tabelle 4-205 dargestellt. Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Die Suche wurde für Abschnitt 4.3.1.1.2 durchgeführt.

Datenbankname EMBASE

Suchoberfläche Ovid
Datum der Suche 08.04.2013
Zeitsegment 1947 to 2013 Week 14
Suchfilter Filter für randomisierte kontrollierte Studien nach Wong 2006-
 Strategy minimizing difference between sensitivity and specificity
 [149]

Tabelle 4-203: Ovid: Suchstrategie EMBASE 1947 to 2013 Week 14

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	259525
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	359656
3	immun* Vir*.mp.	307437
4	acquire* immun* deficienc* syndrom*.mp.	121537
5	1 or 2 or 3 or 4	389357
6	adult*.mp.	4974324
7	5 and 6	112461
8	Raltegravir.mp. or exp raltegravir/	2368
9	RAL.mp.	1742
10	isentress.mp.	272
11	(mk 518 or "mk 0518").mp.	290
12	(mk518 or mk0518).mp.	7
13	871038-72-1.rn.	1952
14	8 or 9 or 10 or 11 or 12 or 13	3958
15	random*.tw.	816043
16	double-blind*.tw.	145266
17	placebo*.mp.	314957
18	15 or 16 or 17	1025005
19	7 and 14	627
20	18 and 19	109

Datenbankname MEDLINE(R) without Revisions

Suchoberfläche	Ovid
Datum der Suche	08.04.2013
Zeitsegment	1946 to March Week 4 2013
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- Strategy minimizing difference between sensitivity and specificity [149]

Tabelle 4-204: Ovid: Suchstrategie MEDLINE(R) without Revisions to March Week 4 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	237834
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	168215
3	immun* Vir*.mp.	75993
4	mensch* immun* vir*.mp.	1
5	erworb* immun*.mp.	99
6	acquire* immun* deficienc* syndrom*.mp.	4874
7	1 or 2 or 3 or 4 or 5 or 6	281856
8	adult*.mp.	4134725
9	erwachs*.mp.	3029
10	8 or 9	4134816
11	7 and 10	93675
12	Raltegravir.mp.	719
13	RAL.mp.	1098
14	(mk 518 or "mk 0518").mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	42
15	(mk518 or mk0518).mp.	3
16	isentress.mp.	20
17	871038-72-1.rn.	0
18	12 or 13 or 14 or 15 or 16 or 17	1743
19	randomized controlled trial.pt.	340089
20	randomized.mp.	488479
21	placebo.mp.	139670

#	Suchbegriffe	Ergebnis
22	19 or 20 or 21	531204
23	11 and 18	220
24	22 and 23	58

Datenbankname Cochrane
EBM Reviews - Cochrane Central Register of Controlled Trials
March 2013

Suchoberfläche Ovid

Datum der Suche 08.04.2013

Zeitsegment 1991 to March 2013

Suchfilter kein Filter

Tabelle 4-205: Ovid: Suchstrategie Cochrane 1991 to March 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	7831
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	4536
3	immun* Vir*.mp.	2626
4	mensch* immun* vir*.mp.	0
5	erworb* immun*.mp.	3
6	acquire* immun* deficienc* syndrom*.mp.	250
7	1 or 2 or 3 or 4 or 5 or 6	8492
8	adult*.mp.	276993
9	erwachs*.mp.	61
10	8 or 9	277000
11	7 and 10	5477
12	Raltegravir.mp.	60
13	RAL.mp.	63
14	(mk 518 or "mk 0518").mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]	5
15	(mk518 or mk0518).mp.	0
16	isentress.mp.	1

#	Suchbegriffe	Ergebnis
17	[871038-72-1.rn.]	0
18	12 or 13 or 14 or 15 or 16 or 17	117
19	11 and 18	41

Für die Bewertung des Zusatznutzens von Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil wurde ein indirekter Vergleich durchgeführt (Abschnitt 4.3.2.1).

Indirekter Vergleich Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil vs. Raltegravir, Emtricitabin und Tenofoviridisoproxil

Eine Literaturrecherche zur zweckmäßigen Vergleichstherapie unter der Fragestellung, wie in Abschnitt 4.2.3.2 beschrieben, wurde durchgeführt. Die Suchstrategien sind in Tabelle 4-206, Tabelle 4-207 und Tabelle 4-208 dargestellt. Studien wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien selektiert. Die Suche wurde für Abschnitt 4.3.2.1 durchgeführt.

Datenbankname	EMBASE
Suchoberfläche	Ovid
Datum der Suche	08.04.2013
Zeitsegment	1947 bis 2013 Woche 14
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006- Strategy minimizing difference between sensitivity and specificity [149]

Tabelle 4-206: Ovid: Suchstrategie EMBASE 1947 bis heute: am 08. April 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	259525
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	359656
3	immun* Vir*.mp.	307437
4	acquire* immun* deficienc* syndrom*.mp.	121537
5	1 or 2 or 3 or 4	389357
6	naiv*.mp.	62560
7	untreat*.mp.	166577
8	first-line.mp.	55940
9	initial*.mp.	851348

#	Suchbegriffe	Ergebnis
10	(therap*-naiv* or therapienaiv* or therapynaiv*).mp.	762
11	(treat*-naiv* or treatmentnaiv*).mp.	5737
12	6 or 7 or 8 or 9 or 10 or 11	1109927
13	adult*.mp.	4974324
14	5 and 12 and 13	11047
15	Raltegravir.mp. or exp raltegravir/	2368
16	RAL.mp.	1742
17	isentress.mp.	272
18	(mk 518 or "mk 0518").mp.	290
19	(mk518 or mk0518).mp.	7
20	871038-72-1.rn.	1952
21	15 or 16 or 17 or 18 or 19 or 20	3958
22	random*.tw.	816043
23	double-blind*.tw.	145266
24	placebo*.mp.	314957
25	22 or 23 or 24	1025005
26	14 and 21	137
27	25 and 26	38

Datenbankname MEDLINE(R) without Revisions
Suchoberfläche Ovid
Datum der Suche 08.04.2013
Zeitsegment 1946 bis 2013, März Woche 4
Suchfilter Filter für randomisierte kontrollierte Studien nach Wong 2006-
Strategy minimizing difference between sensitivity and specificity
[149]

Tabelle 4-207: Ovid: Suchstrategie MEDLINE(R) without Revisions bis heute: am 08. April 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	237834
2	exp acquired immune deficiency syndrome/ or exp Human	168215

#	Suchbegriffe	Ergebnis
	immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	
3	immun* Vir*.mp.	75993
4	mensch* immun* vir*.mp.	1
5	erworb* immun*.mp.	99
6	acquire* immun* deficienc* syndrom*.mp.	4874
7	1 or 2 or 3 or 4 or 5 or 6	281856
8	naiv*.mp.	42255
9	untreat*.mp.	118429
10	first-line.mp.	33361
11	initial*.mp.	587756
12	(therap*-naiv* or therapienaiv* or therapynaiv*).mp.	494
13	(treat*-naiv* or treatmentnaiv*).mp.	2903
14	unbehand*.mp.	177
15	nicht vorbehand*.mp.	6
16	einleit*.mp.	281
17	erst*.mp.	6680
18	nicht behand*.mp.	18
19	behandlungsnaiv*.mp.	0
20	nicht therap*.mp.	3
21	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	770138
22	adult*.mp.	4134725
23	erwachs*.mp.	3029
24	22 or 23	4134816
25	7 and 21 and 24	8951
26	Raltegravir.mp.	719
27	RAL.mp.	1098
28	(mk 518 or "mk 0518").mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]	42
29	(mk518 or mk0518).mp.	3
30	isentress.mp.	20

#	Suchbegriffe	Ergebnis
31	871038-72-1.rn.	0
32	26 or 27 or 28 or 29 or 30 or 31	1743
33	randomized controlled trial.pt.	340089
34	randomized.mp.	488479
35	placebo.mp.	139670
36	33 or 34 or 35	531204
37	25 and 32	51
38	36 and 37	14

Datenbankname Cochrane
EBM Reviews - Cochrane Central Register of Controlled Trials
March 2013

Suchoberfläche Ovid

Datum der Suche 08.04.2013

Zeitsegment kein Zeitsegment

Suchfilter kein Filter

Tabelle 4-208: Ovid: Suchstrategie Cochrane 1991 bis heute: am 08. März 2013

#	Suchbegriffe	Ergebnis
1	hiv*.mp.	7831
2	exp acquired immune deficiency syndrome/ or exp Human immunodeficiency virus infection/ or exp Human immunodeficiency virus 1/ or human* immun* virus*.mp. or exp Human immunodeficiency virus/	4536
3	immun* Vir*.mp.	2626
4	acquire* immun* deficienc* syndrom*.mp.	250
5	1 or 2 or 3 or 4	8492
6	naiv*.mp.	3414
7	untreat*.mp.	7774
8	first-line.mp.	3795
9	initial*.mp.	33251
10	(therap*-naiv* or therapienaiv* or therapynaiv*).mp.	115
11	(treat*-naiv* or treatmentnaiv*).mp.	716
12	6 or 7 or 8 or 9 or 10 or 11	46231

#	Suchbegriffe	Ergebnis
13	adult*.mp.	276993
14	5 and 12 and 13	974
15	Raltegravir.mp. or exp raltegravir/	60
16	RAL.mp.	63
17	isentress.mp.	1
18	(mk 518 or "mk 0518").mp.	5
19	(mk518 or mk0518).mp.	0
20	871038-72-1.rn.	0
21	15 or 16 or 17 or 18 or 19 or 20	117
22	14 and 21	10

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

Die Suche nach registrierten Studien mit Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil (EVG/COBI/FTC/TDF) bzw. Efavirenz, Atazanavir und Raltegravir zur Behandlung der HIV-1-Infektion bei Erwachsenen im Alter von 18 Jahren und darüber angewendet, die nicht antiretroviral vorbehandelt sind oder bei denen HIV-1 keine Mutationen aufweist, die bekanntermaßen mit Resistenzen gegen einen der drei antiretroviralen Wirkstoffe EVG/COBI/FTC/TDF assoziiert sind, erfolgte anhand spezifischer Stichwörter. Die Suchen wurden über die jeweilige „Advanced Search“-Funktion der unten aufgeführten Studienregister durchgeführt.

Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil (EVG/COBI/FTC/TDF)

Die Suche wurde für Abschnitt 4.3.1.1.3 durchgeführt.

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	09.04.2013
Suchstrategie	All Studies [Recruitment] AND elvitegravir/cobicistat/emtricitabine/tenofovir OR stribild OR quad [Intervention] AND hiv infection OR hiv OR acquired immunodeficiency syndrome OR aids OR human AND immunodeficiency syndrome [Condition] AND

Adult (18-65) OR Senior (65+) [Age Group]
 AND
 Phase II OR Phase III OR Phase IV [Phase]
Treffer 14

Da das Studienregister clinicalstudyresults.org zur Zeit der Erstellung dieses Dossiers nicht zur Verfügung stand, wurde hier keine Suche durchgeführt.

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 09.04.2013
Suchstrategie elvitegravir/cobicistat/emtricitabine/tenofovir OR stribild OR Quad [Intervention]
 AND
 hiv infections OR acquired immunodeficiency syndrome OR aids [Condition]
 AND
 All [Recruitment Status]
Treffer 18

Efavirenz

Die Suche wurde für Abschnitt 4.3.1.1.3 durchgeführt.

Studienregister clinicaltrials.gov
Internetadresse <http://www.clinicaltrials.gov/ct2/search/advanced>
Datum der Suche 11.04.2013
Suchstrategie All Studies [Recruitment]
 AND
 efavirenz OR DMP-266 OR DMP266 OR sustiva OR stocrin [Intervention]
 AND
 hiv infection OR hiv OR acquired immunodeficiency syndrome OR aids OR human AND immunodeficiency syndrome [Condition]
 AND
 Adult (18-65) OR Senior (65+) [Age Group]
 AND

Phase II OR Phase III OR Phase IV [Phase]
Treffer 213

Da das Studienregister clinicalstudyresults.org zur Zeit der Erstellung dieses Dossiers nicht zur Verfügung stand, wurde hier keine Suche vorgenommen.

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 09.04.2013
Suchstrategie efavirenz OR sustiva OR stocrin OR DMP 266 OR DMP-266 OR DMP266 [Intervention]
 AND
 hiv infections OR acquired immunodeficiency syndrome OR aids [Condition]
 AND
 All [Recruitment Status]
Treffer 280

Atazanavir

Die Suche wurde für Abschnitt 4.3.1.1.3 durchgeführt.

Studienregister clinicaltrials.gov
Internetadresse <http://www.clinicaltrials.gov/ct2/search/advanced>
Datum der Suche 11.04.2013
Suchstrategie All Studies [Recruitment]
 AND
 atazanavir OR BMS-232632 OR BMS232632 OR Reyataz [Intervention]
 AND
 hiv infection OR hiv OR acquired immunodeficiency syndrome OR aids OR human AND immunodeficiency syndrome [Condition]
 AND
 Adult (18-65) OR Senior (65+) [Age Group]
 AND
 Phase II OR Phase III OR Phase IV [Phase]
Treffer 109

Da das Studienregister clinicalstudyresults.org zur Zeit der Erstellung dieses Dossiers nicht zur Verfügung stand, wurde hier keine Suche vorgenommen.

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	09.04.2013
Suchstrategie	atazanavir OR ATV OR ATZ OR BMS 232632 OR BMS232632 OR BMS-232632 OR reyataz OR 198904-31-3 [Intervention] AND hiv infections OR acquired immunodeficiency syndrome OR aids [Condition] AND All [Recruitment Status]
Treffer	161

Raltegravir

Die Suche wurde für Abschnitt 4.3.1.1.3 durchgeführt.

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	11.04.2013
Suchstrategie	All Studies [Recruitment] AND Raltegravir OR mk518 OR mk0518 OR mk 518 OR mk 0518 OR Isentress [Intervention] AND hiv infection OR hiv OR acquired immunodeficiency syndrome OR aids OR human AND immunodeficiency syndrome [Condition] AND Adult (18-65) OR Senior (65+) [Age Group] AND Phase II OR Phase III OR Phase IV [Phase]
Treffer	130

Da das Studienregister clinicalstudyresults.org zur Zeit der Erstellung dieses Dossiers nicht zur Verfügung stand, wurde hier keine Suche vorgenommen.

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	09.04.2013
Suchstrategie	raltegravir OR isentress OR mk518 OR mk 518 OR mk0518 OR mk 0518 OR 871038-72-1 [Intervention] AND hiv infections OR acquired immunodeficiency syndrome OR aids [Condition] AND All [Recruitment Status]
Treffer	203

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 4-209: (Anhang): im Volltext gesichtete und ausgeschlossene Studien/Publicationen aus der Recherche zu RAL+FTC/TDF

Publikation	Ausschlussgrund
DeJesus, E., J. K. Rockstroh, et al. (2012). "Efficacy of raltegravir versus efavirenz when combined with tenofovir/emtricitabine in treatment-naïve HIV-1-infected patients: week-192 overall and subgroup analyses from STARTMRK." <i>HIV Clinical Trials</i> 13(4): 228-232	A7
Temesgen, Z. and D. S. Siraj (2008). "Raltegravir: First in class HIV integrase inhibitor." <i>Therapeutics and Clinical Risk Management</i> 4(2): 493-500	A1
Teppler, H., D. D. Brown, et al. (2011). "Long-term safety from the raltegravir clinical development program." <i>Current HIV Research</i> 9(1): 40-53*	A7

* gelistet auf der Studienregisterseite von STARTMRK und wurde aufgrund zusätzlicher Daten zu SOC/ PT für den indirekten Vergleich herangezogen.

Tabelle 4-210: (Anhang): im Volltext gesichtete und ausgeschlossene Studien/Publicationen aus dem Screening der Publikationen auf clinicaltrials.gov zu STARTMRK

Publikation	Ausschlussgrund
DeJesus, E, Rockstroh JK, Lennox JL, Saag MS, Lazzarin A, Zhao J, Wan H, Rodgers AJ, Walker ML, Miller M, DiNubile MJ, Nguyen BY, Teppler H, Leavitt R, Sklar P; STARTMRK Investigators. Efficacy of raltegravir versus efavirenz when combined with tenofovir/emtricitabine in treatment-naïve HIV-1-infected patients: week-192 overall and subgroup analyses from STARTMRK. <i>HIV Clin Trials</i> . 2012 Jul-Aug;13(4):228-32.	A7
Rockstroh J, Teppler H, Zhao J, Sklar P, Harvey C, Strohmaier K, Leavitt R, Nguyen BY. Safety and efficacy of raltegravir in patients with HIV-1 and hepatitis B and/or C virus coinfection. <i>HIV Med</i> . 2012 Feb;13(2):127-31. doi: 10.1111/j.1468-1293.2011.00933.x. Epub 2011 May 22.	A7
Rockstroh JK, Teppler H, Zhao J, Sklar P, Miller MD, Harvey CM, Strohmaier KM, Leavitt RY, Nguyen BY. Clinical efficacy of raltegravir against B and non-B subtype HIV-1 in phase III clinical studies. <i>AIDS</i> . 2011 Jul 17;25(11):1365-9. doi: 10.1097/QAD.0b013e328348065a.	A7

Nguyen BY, Isaacs RD, Teppler H, Leavitt RY, Sklar P, Iwamoto M, Wenning LA, Miller MD, Chen J, Kemp R, Xu W, Fromtling RA, Vacca JP, Young SD, Rowley M, Lower MW, Gottesdiener KM, Hazuda DJ. Raltegravir: the first HIV-1 integrase strand transfer inhibitor in the HIV armamentarium. <i>Ann N Y Acad Sci.</i> 2011 Mar;1222:83-9. doi: 10.1111/j.1749-6632.2011.05972.x. Review.	A7
--	----

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 4-211 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						

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 4-212 (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						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
GS-US-236-0102	ja	ja	nein	ja [90, 91]	ja [84, 85]	ja [92, 145]
GS-US-236-0103	ja	ja	nein	ja [95, 96]	ja [86, 87]	ja [97]
einarmig						
GS-US-236-0123 ^d	nein	ja	nein	nein (Studienprotokoll [122])	ja [123, 124]	ja [125, 126]
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.</p> <p>b: Studie, für die der Unternehmer Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>d: Die in den Studienpool aufgenommene Studie wird nicht zur Nutzenableitung herangezogen; sie dient lediglich zur Unterstützung der Übertragbarkeit der Daten aus Modul 4A.</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 4-213 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 4-213 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-213 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0102

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Untersuchung der Wirksamkeit und Sicherheit von EVG/COBI/FTC/TDF (Elvitegravir/ Cobicistat/ Emtricitabin/ Tenofoviridisoproxil) im Vergleich zu EFV/FTC/TDF (Efavirenz/ Emtricitabin/Tenofoviridisoproxil) bei HIV-1 infizierten, antiretroviral nicht-vorbehandelten Erwachsenen
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Phase 3, doppel-blind, multizentrisch, randomisiert, aktiv-kontrolliert, parallel; Zuteilungsverhältnis: 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Änderung (08. April 2010): Update der statistischen Zielkriterien, Endpunkte und Methoden, basierend auf Feedback der US FDA, welche forderte, dass der primäre Wirksamkeits-Endpunkt HIV-1-RNA < 50 Kopien/mL zu Woche 48, basierend auf der Snapshot Analyse, ist. Verwaltungsschreiben (27. Mai 2011): Überarbeitung des Reviews von Kategorie C Ereignissen: Ereignisse werden vom medizinischen Personal von Gilead gereviewed statt von einem Clinical Endpoints Committee, welches aus Prüfarzten besteht
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: <ul style="list-style-type: none"> • Fähig, eine schriftliche Einwilligungserklärung zu verstehen und zu unterschreiben, welche vor dem Start der Studienprozeduren eingeholt werden musste • Plasma HIV-1-RNA Level ≥ 5000 Kopien/mL • Kein vorheriger Gebrauch jeglicher zugelassener oder investigativer antiretroviraler Medikamente zu irgendeiner Zeit • Screening HIV-1 Genotyp Report, durch Gilead bereitgestellt, musste Sensitivität auf Emtricitabin, Tenofoviridisoproxil und Efavirenz gezeigt haben • Normales Elektrokardiogramm (EKG); oder wenn unnormal, musste der Prüfarzt bestätigen dass es nicht klinisch relevant

Item ^a	Charakteristikum	Studieninformation
		<p>ist</p> <ul style="list-style-type: none"> • Adäquate Nierenfunktion: eGFR ≥ 70 mL/min gemäß Cockcroft-Gault (CG) Formel: Männer: $eGFR \text{ (mL/min)} = \frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})}$ Frauen: $eGFR \text{ (mL/min)} = \frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})} \cdot 0,85$ • Lebertransaminasen (Aspartat-Aminotransferase (AST) und Alanin-Aminotransferase (ALT)) ≤ 5 x Obergrenze des Normbereichs (upper limit of normal, ULN) • Gesamt-Bilirubin $\leq 1,5$ mg/dL oder normales direktes Bilirubin • Adäquate hämatologische Funktion (absolute Zahl neutrophiler Granulozyten $\geq 1000/\text{mm}^3$; Thrombozyten $\geq 50.000/\text{mm}^3$; Hämoglobin $\geq 8,5$ g/dL) • Serum-Amylase ≤ 5 x ULN (Patienten mit Serum-Amylase > 5 x ULN blieben geeignet falls Serumlipase ≤ 5 x ULN war) • Gebärfähige Frauen stimmten zu, hochwirksame Verhütungsmethoden (2 separate Formen der Verhütung, von denen eine eine wirksame Barriere-Methode war, nicht-heterosexuell aktiv, sexuelle Abstinenz oder vasktomierter Partner) ab dem Screening für die Dauer der Studienbehandlung und für 12 Wochen nach der letzten Dosis der Studienbehandlung einzunehmen. <ul style="list-style-type: none"> ○ Weibliche Patienten, die hormonelle Verhütungsmethoden nutzten müssen diese Methode für mindestens 3 Monate vor Studienbehandlung verwendet haben ○ Weibliche Patienten, die vor ≥ 12 Monaten aufgehört haben zu menstruieren, die aber keine nachgewiesene ovarielle Hormoninsuffizienz aufwiesen, musste ein Serum-Follikel-stimulierendes Hormonlevel (FSH) beim Screening innerhalb des postmenopausalen Bereichs, basierend auf den

Item ^a	Charakteristikum	Studieninformation
		<p style="text-align: center;">Referenzwerten des zentralen Labors, aufweisen</p> <ul style="list-style-type: none"> • Männliche Patienten stimmten zu, eine hochwirksame Verhütungsmethode für heterosexuellen Geschlechtsverkehr während der Studie und für 12 Wochen nach Beendigung des investigativen medizinischen Produkts, zu verwenden. Eine hochwirksame Verhütungsmethode war definiert als 2 separate Verhütungsmethoden, wobei eine davon eine wirksame Barriere-Methode war, oder der Patient war nicht-heterosexuell aktiv, praktizierte sexuelle Abstinenz oder war vasktomiert • Alter ≥ 18 Jahre • Lebenserwartung ≥ 1 Jahr <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Diagnose eines neuen Aids-definierenden Zustands innerhalb von 30 Tagen vor dem Screening • Erhalt einer medikamentösen Behandlung des Hepatitis C-Virus (HCV), oder es wurde erwartet, dass sie im Laufe dieser Studie eine HCV-Behandlung erhalten würden • Erleiden einer dekompensierten Zirrhose (d.h. Aszites, Enzephalopathie, etc.) • Frauen die stillen • Positiver Serum-Schwangerschaftstest (gebärfähige Frauen) • Implantierter Defibrillator oder Schrittmacher • Aktueller Alkohol- oder Substanzgebrauch, welcher nach Meinung des Prüfarztes die Studien-Compliance des Patienten potenziell beeinträchtigt • Vorgeschichte einer Malignität innerhalb der letzten 5 Jahre (vor dem Screening) oder bestehende Malignität außer kutanes Kaposi-Sarkom (KS), Basalzellkarzinom oder reseziertes, nicht-invasives kutanes Plattenepithelkarzinom. Patienten mit kutanem KS waren geeignet, durften allerdings innerhalb 30 Tage vor Baseline keine systemische Therapie für KS erhalten, und es durfte auch nicht erwartet werden dass sie eine systemische Therapie während der Studie benötigen

Item ^a	Charakteristikum	Studieninformation																														
		<ul style="list-style-type: none"> • Aktive, ernste Infektionen (anders als HIV-1-Infektion), welche parenterale antibiotische oder antimykotische Therapie innerhalb 30 Tage vor Baseline erfordern. • Erhalt einer andauernden Therapie mit einem der folgenden Medikamente, einschließlich der Medikamente, die nicht zusammen mit Elvitegravir, Cobicistat, Emtricitabin, Efavirenz, Tenofoviridisoproxil und EFV/FTC/TDF (siehe jeweilige Fachinformation) genommen werden sollten; oder Patienten mit bekannten Allergien gegen Inhaltsstoffe von EVG/COBI/FTC/TDF oder EFV/FTC/TDF <table border="1" data-bbox="906 786 1401 2000"> <thead> <tr> <th data-bbox="906 786 1169 853">Medikamentenklasse</th> <th data-bbox="1169 786 1401 853">Verbotene Wirkstoffe*</th> </tr> </thead> <tbody> <tr> <td data-bbox="906 853 1169 920">Alpha-adrenerge Rezeptor-Antagonisten</td> <td data-bbox="1169 853 1401 920">Alfuzosin</td> </tr> <tr> <td data-bbox="906 920 1169 965">Analeptika</td> <td data-bbox="1169 920 1401 965">Modafinil</td> </tr> <tr> <td data-bbox="906 965 1169 1010">Antibiotika</td> <td data-bbox="1169 965 1401 1010">Telithromycin</td> </tr> <tr> <td data-bbox="906 1010 1169 1144">Antikonvulsiva</td> <td data-bbox="1169 1010 1401 1144">Phenobarbital, Phenytoin, Carbamazepin, Oxcarbazepin</td> </tr> <tr> <td data-bbox="906 1144 1169 1189">Antipilzmittel</td> <td data-bbox="1169 1144 1401 1189">Voriconazol</td> </tr> <tr> <td data-bbox="906 1189 1169 1256">Antihistaminika</td> <td data-bbox="1169 1189 1401 1256">Astemizol, Terfenadin</td> </tr> <tr> <td data-bbox="906 1256 1169 1357">Mittel gegen Mykobakterien</td> <td data-bbox="1169 1256 1401 1357">Rifampin, Rifapentin, Rifabutin</td> </tr> <tr> <td data-bbox="906 1357 1169 1402">Kalziumkanalblocker</td> <td data-bbox="1169 1357 1401 1402">Bepridil</td> </tr> <tr> <td data-bbox="906 1402 1169 1581">Ergotaminderivate</td> <td data-bbox="1169 1402 1401 1581">Ergotamin, Ergonovin Dihydroergotamin Methylergonovin Ergometrin</td> </tr> <tr> <td data-bbox="906 1581 1169 1704">Magen- und Darmmotilität beeinflussende Substanzen</td> <td data-bbox="1169 1581 1401 1704">Cisaprid</td> </tr> <tr> <td data-bbox="906 1704 1169 1783">Herbale / Natürliche Nahrungsergänzung</td> <td data-bbox="1169 1704 1401 1783">Johanniskraut, Echinacea</td> </tr> <tr> <td data-bbox="906 1783 1169 1883">HMG-CoA-Reduktase-Inhibitoren</td> <td data-bbox="1169 1783 1401 1883">Simvastatin, Lovastatin, Cerivastatin</td> </tr> <tr> <td data-bbox="906 1883 1169 1928">Neuroleptika</td> <td data-bbox="1169 1883 1401 1928">Pimozid</td> </tr> <tr> <td data-bbox="906 1928 1169 2000">Sedativa/ Hypnotika</td> <td data-bbox="1169 1928 1401 2000">Midazolam, Triazolam</td> </tr> </tbody> </table>	Medikamentenklasse	Verbotene Wirkstoffe*	Alpha-adrenerge Rezeptor-Antagonisten	Alfuzosin	Analeptika	Modafinil	Antibiotika	Telithromycin	Antikonvulsiva	Phenobarbital, Phenytoin, Carbamazepin, Oxcarbazepin	Antipilzmittel	Voriconazol	Antihistaminika	Astemizol, Terfenadin	Mittel gegen Mykobakterien	Rifampin, Rifapentin, Rifabutin	Kalziumkanalblocker	Bepridil	Ergotaminderivate	Ergotamin, Ergonovin Dihydroergotamin Methylergonovin Ergometrin	Magen- und Darmmotilität beeinflussende Substanzen	Cisaprid	Herbale / Natürliche Nahrungsergänzung	Johanniskraut, Echinacea	HMG-CoA-Reduktase-Inhibitoren	Simvastatin, Lovastatin, Cerivastatin	Neuroleptika	Pimozid	Sedativa/ Hypnotika	Midazolam, Triazolam
Medikamentenklasse	Verbotene Wirkstoffe*																															
Alpha-adrenerge Rezeptor-Antagonisten	Alfuzosin																															
Analeptika	Modafinil																															
Antibiotika	Telithromycin																															
Antikonvulsiva	Phenobarbital, Phenytoin, Carbamazepin, Oxcarbazepin																															
Antipilzmittel	Voriconazol																															
Antihistaminika	Astemizol, Terfenadin																															
Mittel gegen Mykobakterien	Rifampin, Rifapentin, Rifabutin																															
Kalziumkanalblocker	Bepridil																															
Ergotaminderivate	Ergotamin, Ergonovin Dihydroergotamin Methylergonovin Ergometrin																															
Magen- und Darmmotilität beeinflussende Substanzen	Cisaprid																															
Herbale / Natürliche Nahrungsergänzung	Johanniskraut, Echinacea																															
HMG-CoA-Reduktase-Inhibitoren	Simvastatin, Lovastatin, Cerivastatin																															
Neuroleptika	Pimozid																															
Sedativa/ Hypnotika	Midazolam, Triazolam																															

Item ^a	Charakteristikum	Studieninformation	
		Systemische Kortikosteroide mit Ausnahme vom kurzfristigen Gebrauch (≤ 1 Woche) von Prednison als Steroid-Burst	Alle Agenten, einschließlich Dexamethason
		* Einnahme jeglicher der genannten Medikamente wurde mindestens 21 Tage vor der Baseline Visite und für die Dauer der Studie abgebrochen	
		<ul style="list-style-type: none"> • Teilnahme in anderer klinischer Studie ohne vorheriges Einverständnis des Sponsors war während der Teilnahme an dieser Studie verboten • Jegliche andere klinische Umstände oder vorherige Therapie, die, nach Meinung des Prüfarztes, den Patienten ungeeignet für die Studie oder unfähig die Dosierungsanforderungen einzuhalten, gemacht hätten 	
4b	Studienorganisation und Ort der Studiendurchführung	102 Zentren in 2 Ländern: 97 Zentren in USA, 5 Zentren in Puerto Rico	
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>EVG/COBI/FTC/TDF-Gruppe: 1 grüne, kapselförmige Filmtablette täglich oral (150 mg EVG, 150 mg COBI, 200 mg FTC und 245 mg Tenofoviridisoproxil bzw. Placebo) mit Nahrung, um ungefähr die gleiche Tageszeit</p> <p>EFV/FTC/TDF-Gruppe: 1 rosa, modifizierte kapselförmige Filmtablette täglich oral (600 mg EFV, 200 mg FTC und 245 mg Tenofoviridisoproxil) auf leeren Magen vor dem Zubettgehen, um ungefähr die gleiche Tageszeit</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>Allgemeine Objektiv: Primär: Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF versus EFV/FTC/TDF anhand des Erreichens von HIV-1-RNA <50 Kopien/mL zu Woche 48 Sekundär: Bewertung der Wirksamkeit, Sicherheit und Verträglichkeit der beiden Substanzkombinationen nach 96 Wochen Behandlung</p> <p>Primäres Zielkriterium: Nachweis einer Nichtunterlegenheit der Behandlung mit EVG/COBI/FTC/TDF, verglichen mit der Kontrollbehandlung EFV/FTC/TDF in Bezug auf den Anteil an</p>	

Item ^a	Charakteristikum	Studieninformation
		<p>virologischen Respondern (Plasma-Viruslast < 50 Kopien/mL, gemäß Snapshot) nach 48 Wochen, bei nicht-vorbehandelten HIV-1 infizierten erwachsenen Patienten, mit einer maximal erlaubten Differenz von 12%.</p> <p>Sekundäre Zielkriterien:</p> <ul style="list-style-type: none"> • Anteil an virologischen Respondern (Erreichen und Erhalt einer Plasma-Viruslast < 50 Kopien/mL, gemäß Time-to-loss-of-virologic-response (TLOVR)) nach 48 Wochen • Zeit bis Verlust des virologischen Ansprechens nach 48 Wochen <p>Tertiäre Zielkriterien:</p> <ul style="list-style-type: none"> • Pures virologisches Versagen nach 48 Wochen • Veränderung des Plasma HIV-1-RNA von Baseline bis Woche 48 • Veränderung der CD4-Zellzahl von Baseline bis Woche 48 • Veränderung der CD4-Prozente von Baseline bis Woche 48 <p>Resistenzanalyse: Genotypische Resistenzen zu Woche 48</p> <p>Pharmakokinetik: Plasmakonzentrationen und pharmakokinetische Parameter von Elvitegravir, Cobicistat, Emtricitabin und Tenofoviridisoproxil</p> <p>Sicherheits-Zielkriterien:</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse • Schwerwiegende unerwünschte Ereignisse • Therapieabbrüche aufgrund unerwünschter Ereignisse • Unerwünschte Ereignisse von Interesse • Wichtige unerwünschte Ereignisse • Tode
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Erreichen und Erhalt des bestätigten virologischen Ansprechens (HIV-1-RNA <50 Kopien/mL) bis Woche 48 und 96 (mittels TLOVR-Algorithmus) wurden als sekundäre Wirksamkeitsparameter

Item ^a	Charakteristikum	Studieninformation
		<p>hinzugefügt</p> <p>Änderung (08. April 2010):</p> <p>Update der statistischen Zielkriterien, Endpunkte und Methoden, basierend auf Feedback der US FDA, welche forderte, dass der primäre Wirksamkeits-Endpunkt HIV-1-RNA < 50 Kopien/mL zu Woche 48, basierend auf der Snapshot Analyse, ist.</p>
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Eine Gesamtzahl an 700 Patienten, 1:1 randomisiert auf 2 Gruppen, hatten mindestens 95% Power um eine Nicht-Unterlegenheit hinsichtlich der Responderaten von HIV-1-RNA <50 Kopien/mL zu Woche 48 zwischen den beiden Behandlungsgruppen, gemäß FDA Snapshot-Analyse, zu zeigen. Es wurde angenommen, dass beide Behandlungsgruppen eine Responderate von 0,795 (basierend auf Studie GS-01-934), eine Nichtunterlegenheitsgrenze von 0,12 und ein 1-seitiges 0,025-Signifikanzlevel hatten.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<ul style="list-style-type: none"> • Erste Interim-Analyse: nachdem 350 Patienten die Woche 12-Visiten beendet haben oder die Studienbehandlung vorzeitig beendet haben • Woche 24 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 24 Visite oder die Studienbehandlung vorzeitig beendet hat • Woche 48 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 48 Visite oder die Studienbehandlung vorzeitig beendet hat • Woche 96 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 96 Visite oder die Studienbehandlung vorzeitig beendet hat • Finale Analyse, wenn der letzte Patient die Studie beendet hat <p>Es war seitens Gileads nicht beabsichtigt, das IDMC zu ersuchen, eine vorzeitige Beendigung der Studie zu erwägen, auch wenn frühzeitig Evidenz für bessere Wirksamkeit vorlag. Da es nicht beabsichtigt war, die Studie vorzeitig zu beenden, wurde die konservative Haybittle Prozedur als Stoppregel angewandt [150] . Für jede</p>

Item ^a	Charakteristikum	Studieninformation
		Interim-Analyse durch das IDMC wurde eine alpha-Strafe von 0,001 angewandt.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	IVRS/IWRS (Randomisierung durchgeführt von Bracket)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratifizierung nach HIV-1-RNA (≤ 100.000 Kopien/mL oder > 100.000 Kopien/mL) beim Screening
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 einer individuellen Patientenummer nach der Einwilligungserklärung. Anschließend Zuteilung per IVRS/IWRS.
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 durch den Prüfarzt bzw. den von ihm Bevollmächtigten
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Ja b) Ja c) Ja
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Double-dummy-Design Keine beabsichtigte Ähnlichkeit der Interventionen. Um die Verblindung zu gewährleisten wurde in jeder Gruppe eine aktive (EVG/COBI/FTC/TDF oder EFV/FTC/TDF) und eine Placebo-Tablette (EFV/FTC/TDF oder EVG/COBI/FTC/TDF) gegeben
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Populationen <u>Randomisiertes Analyse-Set (RAS)</u> Alle Patienten die in die Studie randomisiert wurden <u>Intention-to-treat-Analyse-Set (ITT)</u> Alle Patienten die in die Studie randomisiert wurden und mindestens eine Dosis der Studienbehandlung erhalten haben (primäres Analyse-Set für Wirksamkeits-Analysen) <u>Per Protokoll Set (PPS)</u> Alle Patienten die in die Studie randomisiert wurden, mindestens eine Dosis der Studienbehandlung erhalten haben und keine schweren Protokollverletzungen begangen

Item ^a	Charakteristikum	Studieninformation
		<p>haben</p> <p><u>Safety Analyse-Set (SAS)</u> Alle Patienten die randomisiert wurden und mindestens eine Dosis der Studienbehandlung erhalten haben (primäres Analyse-Set für Sicherheits-Analysen). Daten wurden bis 30 Tage nach Medikament Einnahme gesammelt.</p> <p><u>Pharmakokinetik Analyse-Set</u> Alle Patienten die randomisiert wurden, mindestens eine Dosis der Studienbehandlung erhalten haben und für die mindestens eine Konzentrationsbeobachtung eines Analyten (EVG, COBI, FTC oder TFV) vorliegt.</p> <p>Demografie und Baseline Charakteristika: Darstellung von</p> <ul style="list-style-type: none"> • Stichprobengröße • Mittelwert • Standardabweichung • Median • 1. Quartil • 3. Quartil • Minimum • Maximum <p>für kontinuierliche Daten, und Anzahl und Prozent für kategoriale Daten. Für kategoriale Daten wurde zum Vergleich der Behandlungsgruppen ein Cochran-Mantel-Haenszel Test verwendet. Bei kontinuierlichen Daten wurde dazu ein 2-seitiger Wilcoxon Rangsummen Test verwendet.</p> <p>Dauer der Medikament Einnahme Allgemein: Deskriptive Beschreibung.</p> <ul style="list-style-type: none"> • Zeit bis zum vorzeitigen Abbruch der Studienbehandlung: Kaplan-Meier Analyse • Unterschiede in Dauer der Medikament Einnahme (stratifiziert nach Baseline HIV-1-RNA Level): Log-Rank Test <p>Primärer Endpunkt Analysefenster für Woche 48: Tag 309 bis inklusive Tag 378</p> <p><u>Snapshot-Algorithmus:</u></p> <ul style="list-style-type: none"> • <i>Virologisches Ansprechen:</i> Patienten deren letzter verfügbarer HIV-1-RNA Wert <50 Kopien/mL im Woche 48-

Item ^a	Charakteristikum	Studieninformation
		<p>Analysefenster während der randomisierten Behandlung lag</p> <ul style="list-style-type: none"> • <i>Virologisches Versagen:</i> Patienten deren letzter verfügbarer HIV-1-RNA Wert ≥ 50 Kopien/mL im Woche 48-Analysefenster während der randomisierten Behandlung lag; oder die keinen HIV-1-RNAHIV-1-RNA-Wert im Woche 48-Analysefenster während der Behandlung hatten, aufgrund: <ul style="list-style-type: none"> ○ Behandlungsabbruch aufgrund fehlender Wirksamkeit ○ Behandlungsabbruch aus anderen Gründen als unerwünschte Ereignisse, Tod oder fehlende Wirksamkeit und letzter verfügbarer HIV-1-RNAHIV-1-RNA-Wert während Behandlung war ≥ 50 Kopien/mL • <i>Keine virologischen Daten im Woche 48-Analysefenster:</i> Patienten die keine HIV-1-RNAHIV-1-RNA-Werte während Behandlung im Woche 48-Analysefenster hatten, aufgrund: <ul style="list-style-type: none"> ○ Studienbehandlung wurde aufgrund unerwünschtes Ereignis oder Tod abgebrochen (unabhängig vom letzten HIV-1-RNAHIV-1-RNA-Wert) ○ Behandlungsabbruch aus anderen Gründen als unerwünschte Ereignisse, Tod und fehlende Wirksamkeit (zum Beispiel Einverständnis zurückgezogen, lost-to-follow-up, umgezogen) und letzter verfügbarer HIV-1-RNAHIV-1-RNA-Wert während Behandlung war < 50 Kopien/mL ○ Fehlende Werte während dem Analysefenster, aber in Studie verbleibend <p>Primäre Analyse basierend auf dem ITT-Set, sekundäre Analyse basierend auf dem PP-Set Prüfen auf Nicht-Unterlegenheit zur 12% Grenze mittels konventionellem 95%-Konfidenzintervall-Ansatz. Bei Eintreten der Nicht-Unterlegenheit wurde anschließend auf Überlegenheit getestet. Da wegen den beiden Interim-Analysen je eine alpha-Strafe von 0,001 angewandt wurde, ergibt sich ein</p>

Item ^a	Charakteristikum	Studieninformation
		<p>95,2%-KI, welches als 95%-KI bezeichnet wird</p> <p>Die Baseline HIV-1-RNA Schicht (≤ 100.000 oder > 100.000 Kopien/mL)-gewichtete Differenz in den Ansprechraten (und 95%-KI) wurde mittels Schicht-adjustiertem Mantel-Haenszel Anteil berechnet.</p> <ol style="list-style-type: none"> 1. Sensitivitätsanalyse: Ausschluss der Patienten die die Studie aus anderen Gründen als unerwünschte Ereignisse, Tod und fehlende Wirksamkeit (zum Beispiel Einverständnis zurückgezogen, lost-to-follow-up, umgezogen) abbrachen und deswegen keine HIV-1-RNA Daten im Analysefenster hatten und deren letzter verfügbarer HIV-1-RNA HIV-1-RNA-Wert während Behandlung < 50 Kopien/mL war 2. Sensitivitätsanalyse: Patienten die die Studie aus anderen Gründen als unerwünschte Ereignisse, Tod und fehlende Wirksamkeit (zum Beispiel Einverständnis zurückgezogen, lost-to-follow-up, umgezogen) abbrachen und deswegen keine HIV-1-RNA Daten im Analysefenster hatten und deren letzter verfügbarer HIV-1-RNA HIV-1-RNA-Wert während Behandlung < 50 Kopien/mL war, wurden als virologische Ansprecher gewertet 3. Sensitivitätsanalyse: Suche nach Confoundern: <ol style="list-style-type: none"> a. Stratifizierung nach Region b. Ohne jegliche Stratifizierung <p>Sekundäre und tertiäre Endpunkte</p> <p>Allgemein: gleiche Analysemethoden wie primärer Endpunkt</p> <ul style="list-style-type: none"> • Time to loss of virologic response: Kaplan-Meier und Log Rank Test stratifiziert nach Baseline HIV-1-RNA Level ($\leq / > 100.000$ Kopien/mL) • Veränderung \log_{10} HIV-1-RNA, CD4-Zellzahl und CD4%: deskriptiv bzw. mittels ANOVA mit Baseline HIV-1-RNA Level ($\leq / > 100.000$ Kopien/mL) als fester Effekt • Anteil Patienten mit HIV-1-RNA < 50 Kopien/mL mittels M=F und M=E Methoden.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Für den primären Endpunkt wurden Subgruppenanalysen basierend auf der ITT-

Item ^a	Charakteristikum	Studieninformation
		<p>Set berechnet. Stratifiziert wurde dabei nach:</p> <ul style="list-style-type: none"> • Ausgangs-HIV-1-RNA Level (\leq/$>$100.000 Kopien/mL) • Alter ($<$/\geq40 Jahre) • Geschlecht (männlich/weiblich) • Rasse (weiß/nicht-weiß) • CD4-Zellzahl zu Baseline- (\leq/$>$350 Zellen/μL) • Adhärenz zur Studienmedikation ($<$/\geq95%) <p>Alle Subgruppenanalysen wurden mittels Snapshot-Algorithmus berechnet und bezüglich HIV-1-RNA Level (\leq/$>$100.000 Kopien/mL) adjustiert (falls dies nicht der die Subgruppen-definierende Faktor war). Es wurden prozentuale Unterschiede basierend auf den adjustierten Mantel-Haenszel-Anteilen berechnet, sowie jeweils eine logistische Regression mit Baseline HIV-1-RNA Level, Subgruppen-Faktor, Behandlung und Interaktion zwischen dem Subgruppen-Faktor und der Behandlung. Die Homogenität des Behandlungseffektes in den Subgruppen wurde mittels Wald-Test bewertet.</p>
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flow-Chart unten
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p>a) EVG/COBI/FTC/TDF-Gruppe: n=353 EFV/FTC/TDF-Gruppe: n=354</p> <p>b) EVG/COBI/FTC/TDF-Gruppe: n=348 EFV/FTC/TDF-Gruppe: n=352</p> <p>c) EVG/COBI/FTC/TDF-Gruppe: n=348 EFV/FTC/TDF-Gruppe: n=352</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart unten
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	<p>Screening des ersten Patienten: 16. März 2010</p> <p>Letzte Beobachtung des letzten Patienten: 10. August 2011</p> <p>Behandlungsdauer: 96 Wochen</p> <p>Nachbeobachtungsdauer : 30 Tage</p>
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete regulär.
a: nach CONSORT 2010.		

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

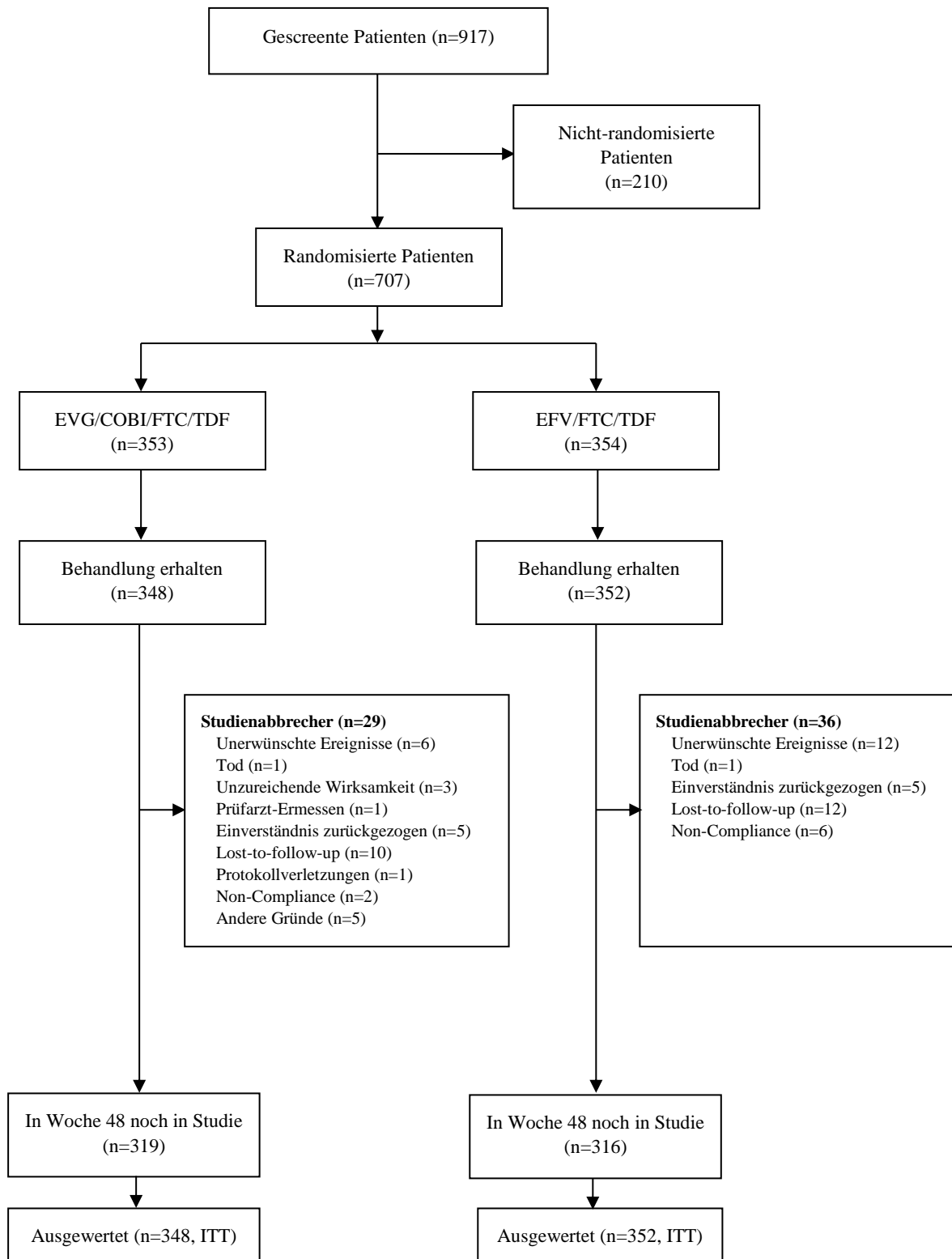


Abbildung 27: Flow-Chart für Studie GS-US-236-0102

Tabelle 4-214 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0103

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Untersuchung der Wirksamkeit und Sicherheit von EVG/COBI/FTC/TDF (Elvitegravir (EVG)/Cobicistat (COBI)/Emtricitabin (FTC)/ Tenofoviridisoproxil (TDF)) im Vergleich zu Ritonavir-geboostertem Atazanavir plus Emtricitabin/Tenofoviridisoproxil (ATV/r+FTC/TDF) bei HIV-1 infizierten, antiretroviral nicht-vorbehandelten Erwachsenen
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Phase 3, doppel-blind, multizentrisch, randomisiert, double-dummy, aktiv-kontrolliert, parallel; Zuteilungsverhältnis: 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Änderung (15. März 2010): Verwendung von RTV-Tabletten, statt RTV-Kapseln Update der statistischen Zielkriterien, Endpunkte und Methoden basierend auf Feedback der US FDA Hinzufügen einer Dual Energy X-ray Absorptiometry (DEXA)-Substudie, die an einem Teil der Patienten durchgeführt werden sollte basierend auf Kommentaren der schwedischen Medical Products Agency Überarbeitung der Einleitung um die aktuelleren Sicherheits-, Wirksamkeits- und klinische Pharmakologie-Daten beendeter und laufender Studien (gesponsort von Gilead Sciences) widerzugeben 14. Juni 2010: Hinzufügen von Text, dass Patienten nach 24 Wochen Behandlung von der Studie ausgeschlossen werden, wenn HIV-1-RNA ≥ 50 Kopien/mL aufgrund suboptimalen virologischen Ansprechens (Forderung des Französischen Ethik Komitee)</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> • Fähig eine schriftliche Einwilligungserklärung zu verstehen und zu unterschreiben, welche vor dem Start der Studienprozeduren eingeholt werden musste • Plasma HIV-1-RNA Level ≥ 5000 Kopien/mL • Kein vorheriger Gebrauch jeglicher zugelassener oder investigativer antiretroviraler Medikamente für jegliche Zeit • Screening HIV-1 Genotyp Report, durch Gilead bereitgestellt, musste Sensitivität auf Emtricitabin, Tenofoviridisoproxil und Atazanavir gezeigt haben • Normales Elektrokardiogramm (EKG); oder wenn unnormal, musste der Prüfarzt bestätigen dass es

Item a	Charakteristikum	Studieninformation
		<p>nicht klinisch relevant ist</p> <ul style="list-style-type: none"> • Adäquate Nierenfunktion: eGFR ≥ 70 mL/min gemäß Cockcroft-Gault (CG) Formel: Männer: eGFR (mL/min) = $\frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})}$ Frauen: eGFR (mL/min) = $\frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})} \cdot 0,85$ • Lebertransaminasen (Aspartat-Aminotransferase (AST) und Alanin-Aminotransferase (ALT) ≤ 5 x Obergrenze des Normbereichs (upper limit of normal, ULN) • Gesamt-Bilirubin $\leq 1,5$ mg/dL oder normales direktes Bilirubin • Adäquate hämatologische Funktion (absolute Zahl neutrophiler Granulozyten $\geq 1000/\text{mm}^3$; Thrombozyten $\geq 50.000/\text{mm}^3$; Hämoglobin $\geq 8,5$ g/dL) • Serum-Amylase ≤ 5 x ULN (Patienten mit Serum-Amylase > 5 x ULN blieben geeignet falls Serumlipase ≤ 5 x ULN war) • Gebärfähige Frauen stimmten zu, hochwirksame Verhütungsmethoden (2 separate Formen der Verhütung, von denen eine eine wirksame Barriere-Methode war, nicht-heterosexuell aktiv, sexuelle Abstinenz oder vasktomierter Partner) ab dem Screening für die Dauer der Studienbehandlung und für 30 Tage nach der letzten Dosis der Studienbehandlung anzuwenden <ul style="list-style-type: none"> ○ Weibliche Patienten, die hormonelle Verhütungsmethoden nutzten müssen diese Methode für mindestens 3 Monate vor Studienbehandlung verwendet haben ○ Weibliche Patienten, die vor ≥ 12 Monaten aufgehört haben zu menstruieren, die aber keine nachgewiesene ovarielle Hormoninsuffizienz aufwiesen, musste ein Serum-Follikel-stimulierendes Hormonlevel (FSH) beim Screening innerhalb des postmenopausalen Bereichs, basierend auf den Referenzwerten des zentralen Labors, aufweisen • Männliche Patienten stimmten zu, eine hochwirksame Verhütungsmethode für heterosexuellen Geschlechtsverkehr während der Studie und für 30 Tage nach Beendigung des investigativen medizinischen Produkts, zu verwenden. Eine hochwirksame Verhütungsmethode war definiert als 2 separate Verhütungsmethoden, wobei eine davon eine wirksame Barriere-Methode war, oder der Patient

Item a	Charakteristikum	Studieninformation				
		<p>war nicht-heterosexuell aktiv, praktizierte sexuelle Abstinenz oder war vasektomiert</p> <ul style="list-style-type: none"> • Alter ≥ 18 Jahre • Lebenserwartung ≥ 1 Jahr <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Diagnose einer neuen Aids-definierenden Erkrankung innerhalb von 30 Tagen vor dem Screening • Erhalt einer Hepatitis C-Virus (HCV)-Therapie, oder Erwartung, dass im Verlauf dieser Studie eine HCV-Behandlung notwendig werden würde • Erleiden einer dekompensierten Zirrhose (zum Beispiel Aszites, Enzephalopathie, etc.) • Frauen die stillen • Positiver Serum-Schwangerschaftstest bei gebärfähigen Frauen • Implantierter Defibrillator oder Schrittmacher • EKG PR-Intervall von ≥ 220ms • Aktueller Alkohol- oder Substanzgebrauch, welcher nach Meinung des Prüfarztes die Studien-Compliance des Patienten potenziell beeinträchtigt • Vorgeschichte einer Malignität innerhalb der letzten 5 Jahre (vor dem Screening) oder bestehende Malignität außer kutanes Kaposi-Sarkom (KS), Basalzellkarzinom oder reseziertes, nicht-invasives kutanes Plattenepithelkarzinom. Patienten mit kutanem KS waren geeignet, durften allerdings innerhalb 30 Tage vor Baseline keine systemische Therapie für KS erhalten, und es durfte auch nicht erwartet werden dass sie eine systemische Therapie während der Studie benötigen • Aktive, ernste Infektionen (andere als HIV-1-Infektion), die parenterale antibiotische oder antimykotische Therapie innerhalb 30 Tage vor Baseline erfordern. • Erhalt einer andauernden Therapie mit einem der folgenden Medikamente, einschließlich der Medikamente die nicht zusammen mit Elvitegravir, Cobicistat, Emtricitabin, Tenofoviridisoproxil, Atazanavir und Ritonavir (siehe jeweilige Fachinformation) genommen werden sollten; oder Patienten mit bekannten Allergien gegen Inhaltsstoffe von EVG/COBI/FTC/TDF, Emtricitabin/Tenofoviridisoproxil-Tabletten, Atazanavir-Kapseln oder Ritonavir-Tabletten. <table border="1" data-bbox="794 1865 1401 1977"> <thead> <tr> <th data-bbox="794 1865 1114 1910">Medikamentenklasse</th> <th data-bbox="1114 1865 1401 1910">Verbotene Wirkstoffe*</th> </tr> </thead> <tbody> <tr> <td data-bbox="794 1910 1114 1977">Alpha-adrenerge Rezeptor-Antagonisten</td> <td data-bbox="1114 1910 1401 1977">Alfuzosin</td> </tr> </tbody> </table>	Medikamentenklasse	Verbotene Wirkstoffe*	Alpha-adrenerge Rezeptor-Antagonisten	Alfuzosin
Medikamentenklasse	Verbotene Wirkstoffe*					
Alpha-adrenerge Rezeptor-Antagonisten	Alfuzosin					

Item a	Charakteristikum	Studieninformation	
		Analeptika	Modafinil
		Antibiotika	Telithromycin
		Antikonvulsiva	Phenobarbital, Phenytoin, Carbamazepin, Oxcarbazepin
		Antimykotika	Voriconazol
		Antihistaminika	Astemizol, Terfenadin
		Mittel gegen Mykobakterien	Rifampin, Rifapentin, Rifabutin
		Kalziumkanalblocker	Bepridil
		Ergotaminderivaten	Ergotamin, Ergonovin Dihydroergotamin Methylegonovin Ergometrin
		Magen- und Darmmotilität beeinflussende Substanzen	Cisaprid
		Pflanzliche / natürliche Nahrungsergänzung	Johanniskraut, Echinacea
		HMG-CoA-Reduktase-Inhibitoren	Simvastatin, Lovastatin, Cerivastatin
		Neuroleptika	Pimozid
		Sedativa/ Hypnotika	Midazolam, Triazolam
		Systemische Kortikosteroide mit Ausnahme vom kurzfristigen Gebrauch (≤ 1 Woche) von Prednison als Steroid-Burst	Alle Wirkstoffe einschließlich Dexamethason
		* Einnahme jeglicher der genannten Medikamente wurde mindestens 21 Tage vor der Baseline Visite und für die Dauer der Studie abgebrochen	
		<ul style="list-style-type: none"> • Teilnahme in anderer klinischer Studie ohne vorheriges Einverständnis des Sponsors war während der Teilnahme an dieser Studie verboten • Jegliche andere klinische Umstände oder vorherige Therapie, die, nach Meinung des Prüfarztes, den Patienten ungeeignet für die Studie oder unfähig die Dosierungsanforderungen einzuhalten, gemacht hätten 	
4b	Studienorganisation und Ort der Studiendurchführung	146 Zentren in 16 Ländern: 88 Zentren in den USA und Puerto Rico, 11 in Frankreich, 8 in Deutschland, 8 in Australien, 7 in Kanada, 6 im Vereinigten Königreich, 3 in Belgien, 3 in Italien, 3 in Österreich, 2 in Thailand, 2 in den Niederlanden, 1 in Portugal, 1 in Mexiko, 1 in der Schweiz und 1 in Schweden.	
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur	EVG/COBI/FTC/TDF-Gruppe: 1 grüne, Kapsel-förmige Filmtablette EVG/COBI/FTC/TDF (150 mg EVG, 150 mg COBI,	

Item a	Charakteristikum	Studieninformation
	Administration etc.	<p>200 mg FTC, 300 mg TDF) einmal täglich, oral während des Essens ungefähr zur gleichen Zeit plus Placebo für ATV/r+FTC/TDF einmal täglich</p> <p>ATV/r+FTC/TDF-Gruppe: 1 undurchsichtige rot/blau, harte Gelatine-Kapsel Atazanavir (ATV), 1 weiße, ovale Filmtablette Ritonavir (RTV) mit eingepprägtem Abbott ‚A‘-Logo und Abo-Code NK und 1 blaue, Kapsel-förmige Filmtablette Emtricitabin/Tenofovirdisoproxil (FTC/TDF) jeweils einmal täglich, oral während des Essens ungefähr zur gleichen Zeit plus Placebo für EVG/COBI/FTC/TDF einmal täglich Dosierung ATV/r+FTC/TDF: 300 mg ATV, 100 mg RTV, 200 mg FTC/300 mg TDF (=FTC/TDF)</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>Allgemeine Zielkriterien: Primär: Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF versus ATV/r+FTC/TDF in HIV-1 infizierten, antiretroviral nicht-vorbehandelten Erwachsenen, gemessen anhand des Erreichens von HIV-1-RNA <50 Kopien/mL zu Woche 48 Sekundär: Bewertung der Wirksamkeit, Sicherheit und Verträglichkeit der beiden Behandlungen nach 96 Wochen</p> <p>Primäres Zielkriterium: Nachweis einer Nichtunterlegenheit von EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF in Bezug auf den Anteil der virologischen Ansprecher (HIV-1-RNA < 50 Kopien/mL, gemäß Snapshot) nach 48 Wochen, bei nicht-vorbehandelten HIV-1 infizierten Erwachsenen</p> <p>Sekundäre Zielkriterien:</p> <ul style="list-style-type: none"> • % Patienten mit virologischem Ansprechen (Erreichen und Erhalt einer Plasma-Viruslast < 50 Kopien/mL, gemäß Time-to-loss-of-virologic-response (TLOVR)) nach 48 Wochen <p>Tertiäre Zielkriterien:</p> <ul style="list-style-type: none"> • Pures virologisches Versagen (HIV-1 ≥ 50 Kopien/mL) nach 48 Wochen • Prozentualer Anteil an virologischen Respondern (HIV-1-RNA < 50 Kopien/mL) unter Verwendung von Missing=Failure- und Missing=Excluded-Methoden • Veränderung des Plasma HIV-1-RNA von Baseline bis Woche 48

Item a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Veränderung der CD4-Zellzahl von Baseline bis Woche 48 • Veränderung der CD4-Prozentzahlen von Baseline bis Woche 48 <p>Pharmakokinetik: Zusammenhang zwischen EVG-Talkonzentration und virologischem Response</p> <p>Sicherheits-Zielkriterien:</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse • Begleitmedikation • Klinische Labortests • 12-Kanal EKG • Größe, Gewicht • Veränderung der Knochenmineraldichte zwischen Baseline und Woche 24/48 in ausgewählten Patienten
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Änderung (15. März 2010): Erreichen und Beibehalten von bestätigtem HIV-1-RNA Level <50 Kopien bis Woche 48 und 96 (mittels TLOVR-Algorithmus) wurde als sekundärer Endpunkt hinzugefügt
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Eine Gesamtzahl an 700 Patienten, 1:1 randomisiert auf 2 Gruppen, hatten mindestens 95% Power, um eine Nichtunterlegenheit hinsichtlich der Responderaten von HIV-1-RNA <50 Kopien/mL zu Woche 48 zwischen den beiden Behandlungsgruppen, gemäß FDA Snapshot-Analyse, zu zeigen. Es wurde angenommen, dass beide Behandlungsgruppen eine Responderate von 0,795 (basierend auf Studie GS-01-934), eine Nichtunterlegenheitsgrenze von 0,12 und ein 1-seitiges 0,025-Signifikanzlevel hatten.
7b	Falls notwendig, Beschreibung von Interim-Analysen und Kriterien für einen vorzeitigen Studienabbruch	<ul style="list-style-type: none"> • Erste Interim-Analyse: nachdem 350 Patienten die Woche 12-Visiten beendet haben oder die Studienbehandlung vorzeitig beendet haben • Woche 24 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 24 Visite oder die Studienbehandlung vorzeitig beendet hat • Woche 48 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 48 Visite oder die Studienbehandlung vorzeitig beendet hat • Woche 96 Analyse: nachdem der letzte Patient die Woche 96 Visite oder die Studienbehandlung vorzeitig beendet hat

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Finale Analyse wenn der letzte Patient die Studie beendet hat <p>Es war seitens Gileads nicht beabsichtigt, das IDMC zu ersuchen, eine vorzeitige Beendigung der Studie zu erwägen, auch wenn frühzeitig Evidenz für bessere Wirksamkeit vorlag. Da es nicht beabsichtigt war, die Studie vorzeitig zu beenden, wurde die konservative Haybittle Prozedur als Stoppregel angewandt [150]. Für jede Interim-Analyse durch das IDMC wurde eine alpha-Strafe von 0,001 angewandt.</p>
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	IVRS/IWRS (Randomisierung durchgeführt von Bracket)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratifizierung nach HIV-1-RNA (≤ 100.000 Kopien/mL oder >100.000 Kopien/mL) beim Screening
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 einer individuellen Patientenummer nach der Einwilligungserklärung. Anschließend Zuteilung per IVRS/IWRS.
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 durch den Prüfarzt bzw. den von ihm Bevollmächtigten
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Ja b) Ja c) Ja
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Double-dummy-Design: Keine beabsichtigte Ähnlichkeit der Interventionen. Um die Verblindung zu gewährleisten wurde in jeder Gruppe eine aktive (EVG/COBI/FTC/TDF oder ATV/r+FTC/TDF) und eine Placebo-Tablette (ATV/r+FTC/TDF oder EVG/COBI/FTC/TDF) gegeben. Das Placebo hatte dieselbe äußere Erscheinung wie das passende Medikament.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Populationen <u>Randomisiertes Analyse-Set (RAS)</u>

Item a	Charakteristikum	Studieninformation
		<p>Alle Patienten die in die Studie randomisiert wurden <u>Intention-to-treat-Analyse-Set (ITT)</u></p> <p>Alle Patienten, die in die Studie randomisiert wurden und mindestens eine Dosis der Studienbehandlung erhalten haben (primäres Analyse-Set für Wirksamkeits-Analysen)</p> <p><u>Per Protokoll Set (PPS)</u></p> <p>Alle Patienten, die in die Studie randomisiert wurden, mindestens eine Dosis der Studienbehandlung erhalten haben und keine schweren Protokollverletzungen begangen haben</p> <p><u>Safety Analyse-Set (SAS)</u></p> <p>Alle Patienten, die randomisiert wurden und mindestens eine Dosis der Studienbehandlung erhalten haben (primäres Analyse-Set für Sicherheits-Analysen). Daten wurden bis 30 Tage nach Medikamenten-Einnahme gesammelt.</p> <p><u>Pharmakokinetik Substudien Analyse-Set</u></p> <p>Alle Patienten, die randomisiert wurden, mindestens eine Dosis der Studienbehandlung EVG/COBI/FTC/TDF erhalten haben, in der Pharmakokinetik-Substudie teilnahmen und für die pharmakokinetische Parameter des Steady State (stationären Zustands) berechenbar waren (primäres Analyse-Set für Pharmakokinetik-Analysen).</p> <p><u>Pharmakokinetik Analyse-Set</u></p> <p>Alle Patienten, die randomisiert wurden, mindestens eine Dosis der Studienbehandlung EVG/COBI/FTC/TDF erhalten haben und für die mindestens eine Konzentrationsbeobachtung eines Analyten (EVG, COBI, FTC oder TFV) vorliegt (verwendet für die Evaluierung der Exposition während der Studie).</p> <p><u>Pharmakokinetik/Pharmakodynamik Analyse-Set</u></p> <p>Alle Patienten, die in die EVG/COBI/FTC/TDF-Gruppe randomisiert wurden, Teil des ITT-Set waren und die mindestens eine Probe nicht-fehlender EVG-Konzentration haben, wobei die Probe innerhalb von 20 bis 24 Stunden nach Dosierzeit gesammelt werden musste.</p> <p><u>DEXA Substudien Analyse-Set</u></p> <p>Alle Patienten, die an der DEXA-Substudie teilnahmen und für die nicht-fehlende Hüft- oder Wirbelsäulen-Knochenmineraldichte-Messungen zur Baseline und zu mindestens einer post-Baseline Visite vorliegen.</p> <p>Demografie und Baseline Charakteristika: Darstellung von</p> <ul style="list-style-type: none"> • Stichprobengröße • Mittelwert

Item a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Standardabweichung • Median • 1. Quartil • 3. Quartil • Minimum • Maximum <p>für kontinuierliche Daten, und Anzahl und Prozent für kategoriale Daten. Für kategoriale Daten wurde zum Vergleich der Behandlungsgruppen ein Cochran-Mantel-Haenszel Test verwendet. Bei kontinuierlichen Daten wurde dazu ein 2-seitiger Wilcoxon Rangsummen Test verwendet.</p> <p>Dauer der Medikament Einnahme Allgemein: Deskriptive Beschreibung.</p> <ul style="list-style-type: none"> • Zeit bis zum vorzeitigen Abbruch der Studienbehandlung: Kaplan-Meier Analyse • Unterschiede in Exposition mit den beiden Medikamenten (stratifiziert nach Baseline HIV-1-RNA Level): Log-Rank Test • Medikamenten-Adhärenz: deskriptive Zusammenfassung basierend auf dem SAS <p>Primärer Endpunkt Analysefenster für Woche 48: Tag 309 bis inklusive Tag 378</p> <p><u>Snapshot-Algorithmus:</u></p> <ul style="list-style-type: none"> • <i>Virologisches Ansprechen:</i> Patienten deren letzter verfügbarer HIV-1-RNA Wert <50 Kopien/mL im Woche 48-Analysefenster während der randomisierten Behandlung lag • <i>Virologisches Versagen:</i> Patienten deren letzter verfügbarer HIV-1-RNA Wert \geq50 Kopien/mL im Woche 48-Analysefenster während der randomisierten Behandlung lag; oder die keinen HIV-1-RNA-Wert im Woche 48-Analysefenster während der Behandlung hatten, aufgrund: <ul style="list-style-type: none"> ○ Behandlungsabbruch aufgrund fehlender Wirksamkeit ○ Behandlungsabbruch aus anderen Gründen als unerwünschte Ereignisse, Tod oder fehlende Wirksamkeit und letzter verfügbarer HIV-1-RNA-Wert während Behandlung war \geq50 Kopien/mL • <i>Keine virologischen Daten im Woche 48 Analysefenster:</i> Patienten die keine HIV-1-RNA-Werte während Behandlung im Woche 48-Analysefenster hatten, aufgrund: <ul style="list-style-type: none"> ○ Studienbehandlung wurde aufgrund

Item a	Charakteristikum	Studieninformation
		<p>unerwünschtes Ereignis oder Tod abgebrochen (unabhängig vom letzten HIV-1-RNA-Wert)</p> <ul style="list-style-type: none"> ○ Behandlungsabbruch aus anderen Gründen als unerwünschte Ereignisse, Tod und fehlende Wirksamkeit (zum Beispiel Einverständnis zurückgezogen, lost-to-follow-up, umgezogen) und letzter verfügbarer HIV-1-RNA-Wert während Behandlung war <50 Kopien/mL ○ Fehlende Werte während dem Analysefenster, aber in Studie verbleibend <p>Primäre Analyse basierend auf dem ITT-Set, sekundäre Analyse basierend auf dem PP-Set</p> <p>Prüfen auf Nichtunterlegenheit zur 12%-Grenze mittels konventionellem 95%-Konfidenzintervall-Ansatz. Bei Eintreten der Nichtunterlegenheit wurde anschließend auf Überlegenheit getestet. Da wegen den beiden Interim-Analysen je eine alpha-Strafe von 0,001 angewandt wurde, ergibt sich ein 95,2%-KI, welches als 95%-KI bezeichnet wird</p> <p>Die Baseline HIV-1-RNA Schicht (≤ 100.000 oder > 100.000 Kopien/mL)-gewichtete Differenz in den Ansprechraten (und 95%-KI) wurde mittels Schicht-adjustiertem Mantel-Haenszel Anteil berechnet.</p> <ol style="list-style-type: none"> 1. Sensitivitätsanalyse: Ausschluss der Patienten die die Studie aus anderen Gründen als unerwünschte Ereignisse, Tod und fehlende Wirksamkeit (zum Beispiel Einverständnis zurückgezogen, lost-to-follow-up, umgezogen) abbrechen und deswegen keine HIV-1-RNA Daten im Analysefenster hatten und deren letzter verfügbarer HIV-1-RNA-Wert während Behandlung <50 Kopien/mL war 2. Sensitivitätsanalyse: Patienten die die Medikamenteneinnahme aus anderen Gründen als unerwünschte Ereignisse, Tod und fehlende Wirksamkeit (zum Beispiel Einverständnis zurückgezogen, lost-to-follow-up, umgezogen) abbrechen und deswegen keine HIV-1-RNA Daten im Analysefenster hatten und deren letzter verfügbarer HIV-1-RNA-Wert während Behandlung <50 Kopien/mL war, wurden als virologische Ansprecher gewertet 3. Sensitivitätsanalyse: Suche nach Confoundern: <ol style="list-style-type: none"> a. Stratifizierung nach Region b. Ohne jegliche Stratifizierung <p>Sekundäre und tertiäre Endpunkte</p> <p>Allgemein: gleiche Analysemethoden wie primärer Endpunkt</p>

Item a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Time to loss of virologic response: Kaplan-Meier und Log Rank Test stratifiziert nach Baseline HIV-1-RNA Level (\leq/$>$100.000 Kopien/mL) • Veränderung \log_{10} HIV-1-RNA, CD4-Zellzahl und CD4%: deskriptiv bzw. mittels ANOVA mit Baseline HIV-1-RNA Level (\leq/$>$100.000 Kopien/mL) als fester Effekt • Anteil Patienten mit HIV-1-RNA $<$50 Kopien/mL mittels Missing=Failure und Missing=Excluded Methoden.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Für den primären Endpunkt wurden Subgruppenanalysen basierend auf der ITT-Set berechnet. Stratifiziert wurde dabei nach:</p> <ul style="list-style-type: none"> • Ausgangs-HIV-1-RNA Level (\leq/$>$100.000 Kopien/mL) • Alter ($<$/\geq40 Jahre) • Geschlecht (männlich/weiblich) • Rasse (weiß/nicht-weiß) • CD4-Zellzahl zu Baseline (\leq/$>$350 Zellen/μL) • Therapieadhärenz ($<$/\geq95%) <p>Alle Subgruppenanalysen wurden mittels Snapshot-Algorithmus berechnet und bezüglich HIV-1-RNA Level (\leq/$>$100.000 Kopien/mL) adjustiert (falls dies nicht der die Subgruppen-definierende Faktor war). Es wurden prozentuale Unterschiede basierend auf den adjustierten Mantel-Haenszel-Anteilen berechnet, sowie jeweils eine logistische Regression mit Baseline HIV-1-RNA Level, Subgruppen-Faktor, Behandlung und Interaktion zwischen dem Subgruppen-Faktor und der Behandlung. Die Homogenität des Behandlungseffektes in den Subgruppen wurde mittels Wald-Test bewertet.</p>
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flow-Chart unten
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) EVG/COBI/FTC/TDF-Gruppe: n=357 ATV/r+FTC/TDF-Gruppe: n=358 b) EVG/COBI/FTC/TDF-Gruppe: n=353 ATV/r+FTC/TDF-Gruppe: n=355 c) EVG/COBI/FTC/TDF-Gruppe: n=353 ATV/r+FTC/TDF-Gruppe: n=355
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart unten
14	Aufnahme / Rekrutierung	

Item ^a	Charakteristikum	Studieninformation
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Screening des ersten Patienten: 06. April 2010 Letzte Beobachtung des letzten Patienten: 13. September 2011 Behandlungsdauer: 96 Wochen Nachbeobachtungsdauer : 30 Tage
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete regulär.

a: nach CONSORT 2010.

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

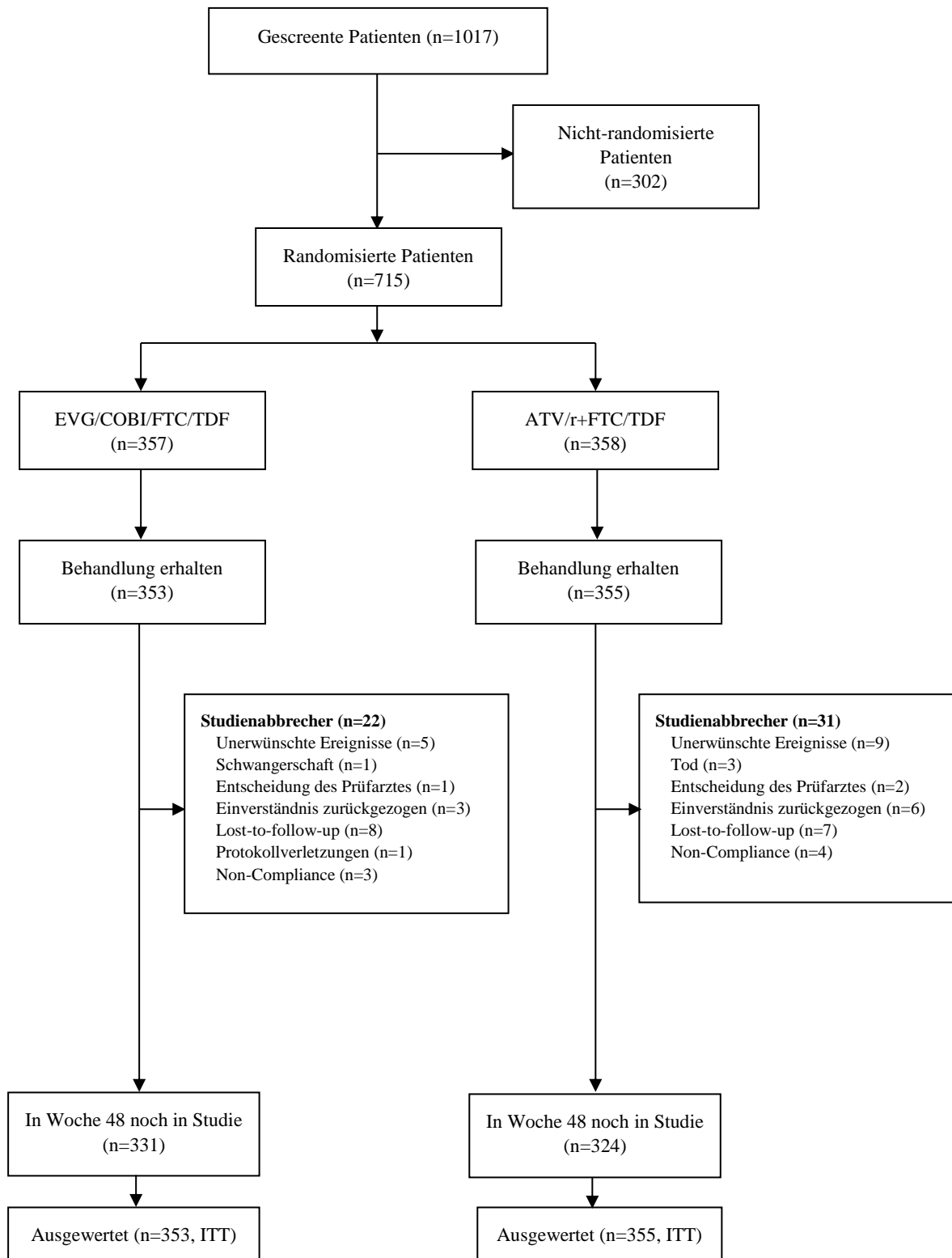


Abbildung 28: Flow-Chart für Studie GS-US-236-0103

Tabelle 4-215 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0104

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Untersuchung der Wirksamkeit und Sicherheit von EVG/COBI/FTC/TDF(Elvitegravir (EVG)/ Cobicistat (COBI)/ Emtricitabin (FTC)/ Tenofoviridisoproxil (TDF)) im Vergleich zu EFV/FTC/TDF (Efavirenz (EFV)/ Emtricitabin (FTC)/ Tenofoviridisoproxil (TDF)) bei HIV-1 infizierten, antiretroviral nicht-vorbehandelten Erwachsenen
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Phase 2, doppel-blind, double-dummy, multizentrisch, randomisiert, aktiv-kontrolliert, parallel; Zuteilungsverhältnis: 2:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Änderung (29 Juli 2009): COBAS AmpliPrep/COBAS Taqman HIV-1 Test wurde durch den COBAS Amplicor HIV-1 Monitor Test zur Evaluierung des virologischen Response ersetzt Update der Definition von HIV-1 bezogenen Krankheiten Änderung (16. März 2010): Der Abschnitt zu den statistischen Analysen wurde überarbeitet um die Evaluierung der Nicht-Unterlegenheit der Wirksamkeitsendpunkte einzuschließen Verwaltungsschreiben (26. Mai 2011): Überarbeitung des Reviews von Kategorie C Ereignissen: Das Clinical Endpoints Committee beinhaltet keine Prüfarzte mehr
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: <ul style="list-style-type: none"> • Fähig eine schriftliche Einwilligungserklärung zu verstehen und zu unterschreiben, welche vor dem Start der Studienprozeduren eingeholt werden musste • Plasma HIV-1-RNA Level ≥ 5000 Kopien/mL • Keine vorherige Einnahme eines zugelassenen oder experimentellen HIV-Medikaments • Normales Elektrokardiogramm (EKG); oder wenn unnormal, musste der Prüfarzt bestätigen dass es nicht klinisch relevant ist

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Adäquate Nierenfunktion: eGFR ≥ 80 mL/min gemäß Cockcroft-Gault (CG) Formel: Männer: $eGFR \text{ (mL/min)} = \frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})}$ Frauen: $eGFR \text{ (mL/min)} = \frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})} \cdot 0,85$ • Lebertransaminasen (Aspartat-Aminotransferase (AST) und Alanin-Aminotransferase (ALT)) $\leq 2,5$ x Obergrenze des Normbereichs (upper limit of normal, ULN) • Gesamt-Bilirubin $\leq 1,5$ mg/dL oder normales direktes Bilirubin • Adäquate hämatologische Funktion (absolute Zahl neutrophiler Granulozyten $\geq 1000/\text{mm}^3$; Thrombozyten $\geq 50.000/\text{mm}^3$; Hämoglobin $\geq 8,5$ g/dL) • Zahl CD4-Zellen > 50 Zellen/μL • Serum-Amylase $\leq 1,5$ x ULN (Patienten mit Serum-Amylase $> 1,5$ x ULN blieben geeignet falls Serumlipase $\leq 1,5$ x ULN war) • Normales Thyroid-stimulierende Hormon • Für gebärfähige Frauen: negativer Serum-Schwangerschaftstest • Männer und gebärfähige Frauen (nicht-menopausale Frauen oder Frauen, die < 2 Jahre in der Menopause sind und follikelstimulierendes Hormon $< 40\text{mIU/mL}$ aufweisen, und die keine Hysterektomie, bilaterale Ovarektomie oder medizinisch dokumentierte Gebärmutterinsuffizienz hatten; junge Frauen, die noch nicht menstruieren sind hier eingeschlossen) stimmten zu, hochwirksame Verhütungsmethoden (2 separate Formen der Verhütung, von denen eine eine wirksame Barriere-Methode war, nicht-heterosexuell aktiv, sexuelle Abstinenz vasktomierter Partner) ab dem Screening für die Dauer der Studienbehandlung und für 12 Wochen nach der letzten Dosis der Studienbehandlung anzuwenden. • Weibliche Patienten, die hormonelle Verhütungsmethoden nutzten, müssen diese Methode für mindestens 3 Monate vor Studienbehandlung verwendet haben

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Bei postmenopausalen Frauen musste ein Serum-Follikel-stimulierendes Hormonlevel (FSH) > 40mIU/mL vorliegen. Falls das FSH ≤ 40mIU/mL lag, musste die Patientin einer Verwendung einer hochwirksamen Verhütungsmethode zustimmen. • Alter ≥18 Jahre • Lebenserwartung ≥1 Jahr <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Diagnose einer neuen Aids-definierenden Erkrankung innerhalb von 30 Tagen vor dem Screening • Dokumentierte Medikamentenresistenz Resistenz-assoziierte Mutationen gegen NRTI oder NNRTI oder primäre PI-Mutationen, nach den aktuellen International Antiviral Society-United States of America (IAS-USA)-Richtlinien • Hepatitis B Oberflächenantigen (HBsAg)- positiv • Hepatitis C Virus-Antikörper (HCVAb) – positiv • Vorhandensein einer Zirrhose, Aszites oder Enzephalopathie • Stillende Frauen • Positiver Serum-Schwangerschaftstest (gebärfähige Frauen) • Impfung innerhalb 90 Tage der Studiendosierung • Persönliche oder familiäre Vorgeschichte bezüglich des Long-QT-Syndroms (Wolfe-Parkinson-White-Syndrom) oder familiäre Vorgeschichte bezüglich plötzlichem Herztod oder unerklärtem Tod bei einer sonst gesunden Person < 30 Jahre • Vorliegen von oder Vorgeschichte bezüglich kardiovaskulärer Erkrankungen, Kardiomyopathie und/oder kardialer Leitungsstörungen • Ausgedehntes, für die Herzfrequenz nach Fridericia's Formel korrigiertes QT-Intervall (QTcF) beim Screening (QTcF > 450ms bei Männern, QtcF >470ms bei Frauen) • PR-Intervall ≥ 200ms oder ≤ 120ms im EKG bei der Screening-Visite • Implantierter Defibrillator oder

Item ^a	Charakteristikum	Studieninformation																										
		<p>Schrittmacher</p> <ul style="list-style-type: none"> • Erhalt einer andauernden Therapie mit einem der folgenden Medikamente, einschließlich der Medikamente die nicht zusammen mit Elvitegravir, Cobicistat, Emtricitabin, Efavirenz und Tenofoviridisoproxil (siehe jeweilige Fachinformation) genommen werden sollten. <table border="1" data-bbox="906 568 1401 1957"> <thead> <tr> <th data-bbox="906 568 1169 636">Medikamentenklasse</th> <th data-bbox="1169 568 1401 636">Verbotene Wirkstoffe*</th> </tr> </thead> <tbody> <tr> <td data-bbox="906 636 1169 703">Alpha-adrenerge Rezeptor-Antagonisten</td> <td data-bbox="1169 636 1401 703">Alfuzosin</td> </tr> <tr> <td data-bbox="906 703 1169 893">Antiarrhythmika</td> <td data-bbox="1169 703 1401 893">Amiodaron, Flecainid, Quinidin, Propafenon, systemisches Lidocain, Mexilitin, Disopyramid</td> </tr> <tr> <td data-bbox="906 893 1169 938">Antibiotika</td> <td data-bbox="1169 893 1401 938">Telithromycin</td> </tr> <tr> <td data-bbox="906 938 1169 1039">Antikonvulsiva</td> <td data-bbox="1169 938 1401 1039">Phenobarbital, Phenytoin, Carbamazepin</td> </tr> <tr> <td data-bbox="906 1039 1169 1140">Antimykotika</td> <td data-bbox="1169 1039 1401 1140">Voriconazol, Ketoconazol, Itraconazol</td> </tr> <tr> <td data-bbox="906 1140 1169 1218">Antihistaminika</td> <td data-bbox="1169 1140 1401 1218">Astemizol, Terfenadin</td> </tr> <tr> <td data-bbox="906 1218 1169 1319">Mittel gegen Mykobakterien</td> <td data-bbox="1169 1218 1401 1319">Rifampin, Rifapentin, Rifabutin</td> </tr> <tr> <td data-bbox="906 1319 1169 1487">Kalziumkanalblocker</td> <td data-bbox="1169 1319 1401 1487">Felodipin, Bepridil, Nifedipin, Nicardipin, Verapamil, Diltiazem, Amlodipin</td> </tr> <tr> <td data-bbox="906 1487 1169 1677">Ergotaminderivaten</td> <td data-bbox="1169 1487 1401 1677">Ergotamin, Ergonovin, Dihydroergotamin, Methylegonovin, Ergometrin</td> </tr> <tr> <td data-bbox="906 1677 1169 1800">Magen- und Darmmotilität beeinflussende Substanzen</td> <td data-bbox="1169 1677 1401 1800">Cisaprid</td> </tr> <tr> <td data-bbox="906 1800 1169 1879">Pflanzliche Nahrungsergänzung</td> <td data-bbox="1169 1800 1401 1879">Johanniskraut, Echinacea</td> </tr> <tr> <td data-bbox="906 1879 1169 1957">HMG-CoA-Reduktase-Inhibitoren</td> <td data-bbox="1169 1879 1401 1957">Simvastatin, Lovastatin, Cerivastatin</td> </tr> </tbody> </table>	Medikamentenklasse	Verbotene Wirkstoffe*	Alpha-adrenerge Rezeptor-Antagonisten	Alfuzosin	Antiarrhythmika	Amiodaron, Flecainid, Quinidin, Propafenon, systemisches Lidocain, Mexilitin, Disopyramid	Antibiotika	Telithromycin	Antikonvulsiva	Phenobarbital, Phenytoin, Carbamazepin	Antimykotika	Voriconazol, Ketoconazol, Itraconazol	Antihistaminika	Astemizol, Terfenadin	Mittel gegen Mykobakterien	Rifampin, Rifapentin, Rifabutin	Kalziumkanalblocker	Felodipin, Bepridil, Nifedipin, Nicardipin, Verapamil, Diltiazem, Amlodipin	Ergotaminderivaten	Ergotamin, Ergonovin, Dihydroergotamin, Methylegonovin, Ergometrin	Magen- und Darmmotilität beeinflussende Substanzen	Cisaprid	Pflanzliche Nahrungsergänzung	Johanniskraut, Echinacea	HMG-CoA-Reduktase-Inhibitoren	Simvastatin, Lovastatin, Cerivastatin
Medikamentenklasse	Verbotene Wirkstoffe*																											
Alpha-adrenerge Rezeptor-Antagonisten	Alfuzosin																											
Antiarrhythmika	Amiodaron, Flecainid, Quinidin, Propafenon, systemisches Lidocain, Mexilitin, Disopyramid																											
Antibiotika	Telithromycin																											
Antikonvulsiva	Phenobarbital, Phenytoin, Carbamazepin																											
Antimykotika	Voriconazol, Ketoconazol, Itraconazol																											
Antihistaminika	Astemizol, Terfenadin																											
Mittel gegen Mykobakterien	Rifampin, Rifapentin, Rifabutin																											
Kalziumkanalblocker	Felodipin, Bepridil, Nifedipin, Nicardipin, Verapamil, Diltiazem, Amlodipin																											
Ergotaminderivaten	Ergotamin, Ergonovin, Dihydroergotamin, Methylegonovin, Ergometrin																											
Magen- und Darmmotilität beeinflussende Substanzen	Cisaprid																											
Pflanzliche Nahrungsergänzung	Johanniskraut, Echinacea																											
HMG-CoA-Reduktase-Inhibitoren	Simvastatin, Lovastatin, Cerivastatin																											

Item ^a	Charakteristikum	Studieninformation	
		Neuroleptika	Pimozid
		Sedativa/ Hypnotika	Midazolam, Triazolam
		Systemische Kortikosteroide	Alle Mittel, einschließlich Dexamethason
		Mittel zur systemischen antineoplastischen Chemotherapie	Alle Mittel
		* Einnahme jeglicher der genannten Medikamente wurde mindestens 30 Tage vor der Baseline Visite und für die Dauer der Studie abgebrochen	
		<ul style="list-style-type: none"> • Aktueller Alkohol- oder Substanzgebrauch, welcher nach Meinung des Prüfarztes die Studien-Compliance des Patienten potenziell beeinträchtigt • Vorgeschichte oder Vorliegen einer Malignität (inklusive unbehandelter in situ Karzinome), außer kutanes Kaposi-Sarkom (KS), Basalzellkarzinom oder reseziertes, nicht-invasives kutanes Plattenepithelkarzinom. Patienten mit Biopsie-bestaätigtem kutanem KS waren geeignet, durften allerdings innerhalb 30 Tage vor Baseline keine systemische Therapie für KS erhalten, und es durfte auch nicht erwartet werden dass sie eine systemische Therapie während der Studie benötigen • Aktive, ernste Infektionen (anders als HIV-1-Infektion), welche parenterale antibiotische oder antimykotische Therapie innerhalb 30 Tage vor Baseline erfordern. • Teilnahme in einer anderen klinischen Studie ohne vorheriges Einverständnis des Sponsors war während der Teilnahme an dieser Studie verboten • Kontraindizierte Medikamente hinsichtlich FTC, EFV oder TDF • Bekannte Allergien gegen sonstige Bestandteile der EVG/COBI/FTC/TDF- oder EFV/FTC/TDF-Tabletten • Jegliche andere klinische Umstände oder vorherige Therapie, die, nach Meinung des Prüfarztes, den Patienten ungeeignet für die Studie oder unfähig die Dosierungsanforderungen einzuhalten, gemacht hätten 	
4b	Studienorganisation und Ort der Studiendurchführung	30 Zentren in den USA	

Item ^a	Charakteristikum	Studieninformation
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p>EVG/COBI/FTC/TDF-Gruppe:</p> <p>Verblindete Phase: 1 weiße kapselförmige Filmtablette täglich oral (150 mg EVG, 150 mg COBI, 200 mg FTC und 245 mg Tenofovirdisoproxil bzw. Placebo)</p> <p>Unverblindete Phase: 1 grüne kapselförmige Filmtablette mit aufgeprägtem „GILEAD“ auf einer Seite (Inhalte wie in verblindeter Phase)</p> <p>EFV/FTC/TDF-Gruppe:</p> <p>Verblindete Phase: 1 rosa, modifizierte kapselförmige Filmtablette mit aufgeprägtem „123“ auf einer Seite abends oral (600 mg EFV, 200 mg FTC und 245 mg Tenofovirdisoproxil)</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>Allgemeine Objektiv:</p> <p>Primär: Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF versus EFV/FTC/TDF anhand des Erreichens von HIV-1-RNA <50 Kopien/mL zu Woche 24</p> <p>Sekundär: Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF versus EFV/FTC/TDF, anhand des Erreichens von HIV-1-RNA <50 Kopien/mL zu Woche 48</p> <p>Primäres Zielkriterium:</p> <p>Vergleich des Anteils der virologischen Ansprecher (Plasma-Viruslast < 50 Kopien/mL gemäß COBAS Amplicor HIV-1 Monitor Test) unter EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF zu Woche 24 bei nicht-vorbehandelten HIV-1 infizierten erwachsenen Patienten.</p> <p>Sekundäre Zielkriterien:</p> <ul style="list-style-type: none"> • % Patienten mit virologischem Ansprechen (Plasma-Viruslast <50 Kopien/mL gemäß COBAS Amplicor HIV-1 Monitor Test) zu Woche 48 mit Missing=Failure, Missing=Excluded und Missing/Switch=Failure Methoden • Virologisches Ansprechen (Plasma-Viruslast <50 Kopien/mL) zu Woche 24 und 48 gemäß FDA-definierter Snapshot Analyse • Veränderung der HIV-1-RNA (log₁₀ Kopien/mL) zwischen Baseline und Woche 24 bzw. 48 • Veränderung von Anzahl und Prozent der

Item ^a	Charakteristikum	Studieninformation
		<p>CD4-Zellzahl zwischen Baseline und Woche 24 bzw. 48</p> <p>Pharmakokinetik: Pharmakokinetische Blutprobe, die 20-24 Stunden nach einer beobachteten Dosis EVG/COBI/FTC/TDF (bzw. Placebo) genommen wurde</p> <p>Sicherheits-Zielkriterien:</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse • Begleitmedikation • Klinische Labortests • Rückenlägriges 12-Kanal-EKG • Größe, Gewicht
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Änderung (29 Juli 2009): Test zur Evaluierung des virologischen Responses wurde geändert: COBAS AmpliPrep/COBAS Taqman HIV-1 Test wurde durch den COBAS Amplicor HIV-1 Monitor Test ersetzt</p> <p>Änderung (16. März 2010): Der Abschnitt zu den statistischen Analysen wurde überarbeitet, um die Evaluierung der Nicht-Unterlegenheit der Wirksamkeitsendpunkte einzuschließen</p>
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Eine Stichprobengröße von 50 in der EVG/COBI/FTC/TDF-Gruppe wurde festgesetzt, um die Responderaten von HIV-1-RNA <50 Kopien/mL zu Woche 24 für die Planung der Phase 3-Studie zu schätzen. 75 Patienten, randomisiert in einem Verhältnis von 2:1, lieferten 26% Power um eine Nicht-Unterlegenheit hinsichtlich der Responderaten von HIV-1-RNA <50 Kopien/mL zu Woche 24 zu zeigen, falls eine Responderate von 84% in beiden Behandlungsgruppen und einer Nicht-Unterlegenheitsgrenze von 0,12 angenommen wurde.</p> <p>Insgesamt wurden nur 71 Patienten in die Studie aufgenommen. Da es sich um eine Studie zur Planung einer Phase 3-Studie handelt, hat diese Verringerung der Stichprobengröße nur begrenzten Einfluss auf die Schätzung der Responderaten.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<ul style="list-style-type: none"> • Woche 16 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die 16. Studienwoche beendet hat

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Woche 16 Zwischen-Pharmakokinetik-Analyse • Woche 24 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 24 Visite oder die Studienbehandlung vorzeitig beendet hat • Woche 48 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 48 Visite oder die Studienbehandlung vorzeitig beendet hat • Woche 48+ Interim-Analyse: nachdem der letzte Patient die Studie beendet oder die Studienbehandlung vorzeitig beendet hat, inklusive der 30 Tage Follow-up-Visite • Woche 96 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 96 Visite oder die Studienbehandlung vorzeitig beendet hat • Finale Analyse wenn der letzte Patient die Studie beendet hat <p>Es war seitens Gileads nicht beabsichtigt, das IDMC zu ersuchen, eine vorzeitige Beendigung der Studie zu erwägen, auch wenn frühzeitig Evidenz für bessere Wirksamkeit vorlag. Da es nicht beabsichtigt war, die Studie vorzeitig zu beenden, wurde die konservative Haybittle Prozedur als Stoppregele angewandt [150]. Für jede Interim-Analyse durch das IDMC wurde eine alpha-Strafe von 0,001 angewandt.</p>
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	IVRS/IWRS
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratifizierung nach HIV-1-RNA (≤ 100.000 Kopien/mL oder > 100.000 Kopien/mL) beim Screening Blockrandomisierung mit einer Blockgröße von 6
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 einer individuellen Patientenummer nach der Einwilligungserklärung. Anschließend Zuteilung per IVRS/IWRS.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die	Randomisierung durch den Prüfarzt bzw. den von ihm Bevollmächtigten

Item ^a	Charakteristikum	Studieninformation
	Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Ja b) Ja c) ja
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Double-dummy-Design Keine beabsichtigte Ähnlichkeit der Interventionen. Um die Verblindung zu gewährleisten wurde in jeder Gruppe eine aktive (EVG/COBI/FTC/TDF oder EFV/FTC/TDF) und eine Placebo-Tablette (EFV/FTC/TDF oder EVG/COBI/FTC/TDF) gegeben
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Populationen</p> <p><u>Randomisiertes Analyse-Set (RAS)</u> Alle Patienten, die in die Studie randomisiert wurden</p> <p><u>Intention-to-treat-Analyse-Set (ITT)</u> Alle Patienten, die in die Studie randomisiert wurden und mindestens eine Dosis der Studienbehandlung erhalten haben (primäres Analyse-Set für Wirksamkeits-Analysen)</p> <p><u>Per Protokoll Set (PPS)</u> Alle Patienten, die in die Studie randomisiert wurden, mindestens eine Dosis der Studienbehandlung erhalten haben und keine schweren Protokollverletzungen begangen haben</p> <p><u>Safety Analyse-Set (SAS)</u> Alle Patienten, die randomisiert wurden und mindestens eine Dosis der Studienbehandlung erhalten haben (primäres Analyse-Set für Sicherheits-Analysen). Daten wurden bis 30 Tage nach Medikament Einnahme gesammelt.</p> <p><u>All EVG/COBI/FTC/TDF Enrolled</u> Alle Patienten, die mindestens eine Dosis EVG/COBI/FTC/TDF in der verblindeten oder unverblindeten Phase erhalten haben</p> <p><u>All EVG/COBI/FTC/TDF Efficacy</u> Alle Patienten, die mindestens eine Dosis EVG/COBI/FTC/TDF (verblindet/unverblindet) erhalten haben und mindestens eine post-Baseline HIV-1-RNA Wert aufweisen (Analyse-Set für Wirksamkeit bzgl. EVG/COBI/FTC/TDF).</p>

Item ^a	Charakteristikum	Studieninformation
		<p><u>All EVG/COBI/FTC/TDF Safety</u> Alle Patienten, die mindestens eine Dosis EVG/COBI/FTC/TDF erhalten haben (verblindet/unverblindet). Analyse-Set für Sicherheitsanalysen bzgl. EVG/COBI/FTC/TDF</p> <p><u>Pharmakokinetik-Substudien Analyse-Set</u> Alle Patienten, die randomisiert wurden, mindestens eine Dosis der Studienbehandlung EVG/COBI/FTC/TDF erhalten haben und für die steady-state pharmakokinetische Parameter berechenbar waren (primäres Analyse-Set für Pharmakokinetik-Analysen).</p> <p><u>Pharmakokinetik Analyse-Set</u> Alle Patienten, die randomisiert wurden, mindestens eine Dosis der Studienbehandlung EVG/COBI/FTC/TDF erhalten haben und für die mindestens eine Konzentrationsbeobachtung eines Analyten (EVG, COBI, FTC oder TFV) vorliegt (verwendet für die Evaluierung der Exposition während der Studie).</p> <p>Demografie und Baseline Charakteristika: Darstellung von</p> <ul style="list-style-type: none"> • Stichprobengröße • Mittelwert • Standardabweichung • Median • 1. Quartil • 3. Quartil • Minimum • Maximum <p>für kontinuierliche Daten, und Anzahl und Prozent für kategoriale Daten. Für kategoriale Daten wurde zum Vergleich der Behandlungsgruppen ein exakter Fisher-Test (nicht-ordinal) bzw. ein Cochran-Mantel-Haenszel Test (ordinal) verwendet. Bei kontinuierlichen Daten wurde dazu ein 2-seitiger Wilcoxon Rangsummen Test verwendet.</p> <p>Dauer der Medikamenten-Einnahme Allgemein: Deskriptive Beschreibung.</p> <ul style="list-style-type: none"> • Zeit bis zum vorzeitigen Abbruch der Studienbehandlung: Kaplan-Meier Analyse • Unterschiede in Dauer der Medikamenten-Einnahme (stratifiziert)

Item ^a	Charakteristikum	Studieninformation
		<p>nach Baseline HIV-1-RNA Level): Log-Rank Test</p> <p>Primärer Endpunkt Primäre Analyse basierend auf dem ITT-Set, sekundäre Analyse basierend auf dem PP Set Primäre Analyse des primären Endpunkts basierend auf Missing=Failure Methode; sekundäre Analyse basierend auf Missing/Switch=Failure Methode Prüfen auf Nicht-Unterlegenheit gegenüber EFV/FTC/TDF zur 12% Grenze mittels konventionellem 95%-Konfidenzintervall-Ansatz. Hierzu wurden Baseline HIV-1-RNA Kategorien (≤ 100.000 oder > 100.000 Kopien/mL)-gewichtete MH-Anteile und Normalverteilungsapproximation verwendet.</p> <p>Sekundäre und tertiäre Endpunkte Allgemein: gleiche Analysemethoden wie primärer Endpunkt</p> <ul style="list-style-type: none"> • HIV-1-RNA < 50 Kopien/mL wurde zu Woche 24 und 48 gemäß FDA-Snapshot-Methode evaluiert • Veränderung \log_{10} HIV-1-RNA, CD4-Zellzahl und CD4%: deskriptiv bzw. mittels ANOVA mit Baseline HIV-1-RNA Level ($\leq / > 100.000$ Kopien/mL) als fester Effekt • Jeweils für Missing=Failure, Missing=Excluded und Missing/Switch=Failure Methoden
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Nicht zutreffend
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flow-Chart unten
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) EVG/COBI/FTC/TDF-Gruppe: n=48 EFV/FTC/TDF-Gruppe: n=23 b) EVG/COBI/FTC/TDF-Gruppe: n=48 EFV/FTC/TDF-Gruppe: n=23 c) EVG/COBI/FTC/TDF-Gruppe: n=48 EFV/FTC/TDF-Gruppe: n=23
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart unten

Item ^a	Charakteristikum	Studieninformation
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Screening des ersten Patienten: 30. März 2009 Letzte Beobachtung des letzten Patienten: 17. Mai 2011 Behandlungsdauer: 96 Wochen Nachbeobachtungsdauer : 30 Tage
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete regulär.
a: nach CONSORT 2010.		

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

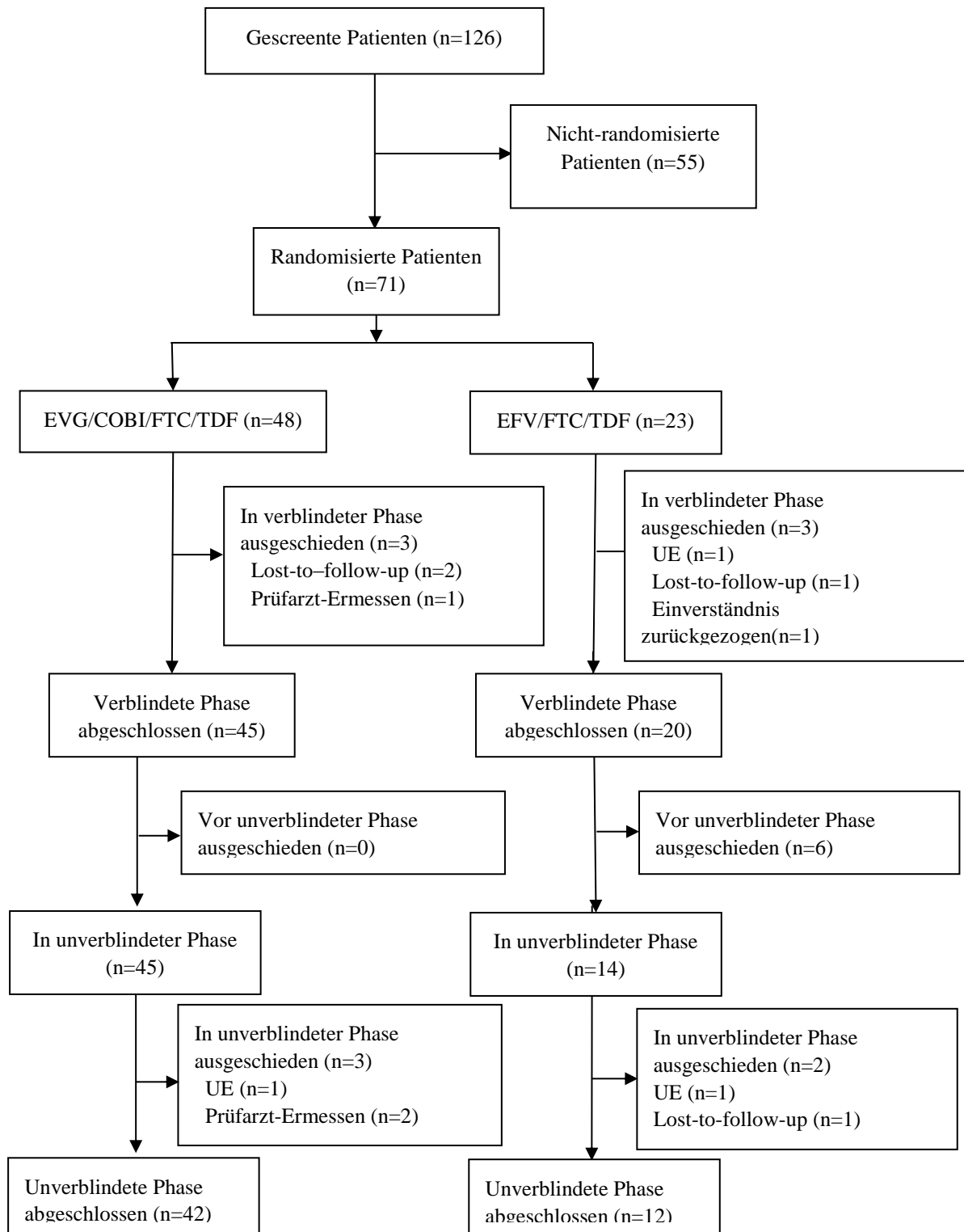


Abbildung 29: Flow-Chart für Studie GS-US-236-0104

Tabelle 4-216 (Anhang): Studiendesign und -methodik für STARTMRK-Studie

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Untersuchung der Sicherheit und Wirksamkeit einer Raltegravir- versus einer Efavirenz-basierten Kombinationstherapie bei antiretroviral nicht-vorbehandelten erwachsenen Patienten mit HIV-1 Infektion
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Phase 3, multizentrisch, doppelblind, randomisiert, aktiv-kontrolliert, parallel; Zuteilungsverhältnis: 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Angaben
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien:</u> <ul style="list-style-type: none"> • HIV-1 Infektion • Bisher noch keine antiretrovirale Therapie erhalten • ≥ 18 Jahre • HIV-RNA Konzentration > 5.000 Kopien/mL <u>Ausschlusskriterien:</u> <ul style="list-style-type: none"> • Akute oder dekompensierte chronische Hepatitis (Patienten mit chronischer Hepatitis, deren Serum Aminotransferase Konzentrationen geringer als 5fach die Obergrenze des Normbereichs waren, wurden eingeschlossen) • Niereninsuffizienz (Patienten brauchten Dialyse, hatten Serum-Kreatinin-Clearance Konzentration von mehr als 2 mal die Obergrenze des Normbereichs oder geschätzte Kreatinin-Clearance beim Screening geschätzte Kreatinin-Clearance beim Screening ≤ 30 mL/min nach der Formel von Cockcroft-Gault berechnet) • Jegliche medizinische Erkrankung, die die Ausführung oder Interpretation der Studie beeinflussen könnte • Resistenz gegen Tenofovir, Emtricitabin oder Efavirenz • Schwanger oder stillend
4b	Studienorganisation und Ort der Studiendurchführung	67 Zentren auf 5 Kontinenten: Australien, Brasilien, Chile, Deutschland, Frankreich, Indien, Italien, Kanada, Kolumbien, Mexiko,

Item ^a	Charakteristikum	Studieninformation
		Peru, Spanien, Thailand und USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Raltegravir-Gruppe: 2 mal täglich 400 mg Raltegravir oral unabhängig von den Mahlzeiten + Placebo-Tablette für Efavirenz + 300 mg Tenofoviridisoproxilfumarat (entspricht 245 mg Tenofoviridisoproxil) und 200 mg Emtricitabin als Fixkombination morgens zu den Mahlzeiten</p> <p>Efavirenz-Gruppe: 1 mal täglich 600 mg Efavirenz oral auf leeren Magen vor dem Zubettgehen + Placebo-Tablette für Raltegravir + 300 mg Tenofoviridisoproxilfumarat (entspricht 245 mg Tenofoviridisoproxil) und 200 mg Emtricitabin als Fixkombination morgens zu den Mahlzeiten</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><u>Primärer Endpunkt:</u> Absinken der Viruslast auf <50 Kopien/mL zu Woche 48 (Nichtunterlegenheit zur 12%-Grenze)</p> <p><u>Sekundäre Endpunkte:</u></p> <ul style="list-style-type: none"> • Absinken der Viruslast auf <400 Kopien/mL zu Woche 48 • Veränderung der CD4-Zellzahl ab Baseline bis Woche 48 <p><u>Weitere Endpunkte:</u></p> <ul style="list-style-type: none"> • Gesamtrate unerwünschter Ereignisse • Schwerwiegende unerwünschte Ereignisse • Therapieabbruch aufgrund unerwünschter Ereignisse
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Angaben
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Basierend auf 75% der 275 Patienten in jeder Behandlungsgruppe, die eine HIV-RNA Konzentration von weniger als 50 Kopien/mL erreichen, hätte die Studie 90% Power um die Nichtunterlegenheit von Raltegravir zu Efavirenz zu zeigen.</p> <p>Mit 275 Patienten in jeder Gruppe hat die Studie 90% Power um einen Gruppenunterschied hinsichtlich unerwünschter Ereignisse, die bei 20% der Patienten in jeglicher Behandlungsgruppe auftreten, von bis zu 11,1% zu zeigen</p>

Item ^a	Charakteristikum	Studieninformation
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Ein data safety monitoring board (DSMB, bestehend aus fünf Akademikern und einem Merck-Statistiker, der die Behandlungszuteilung kannte) reviewte alle 3 Monate maskierte Sicherheits- und Wirksamkeitsergebnisse. Es konnten durch das DSMB auch unmaskierte Daten angefordert werden, um auf dieser Basis Empfehlungen zur Beendigung der Studie allgemein oder für einzelne Patienten zu empfehlen. Die Empfehlungen wurden an ein Aufsichtskomitee (Mitglieder des Merck Senior Managements) weitergeleitet, welches für die allgemeine wissenschaftliche Richtung des Programms verantwortlich war und darüber entschied ob die Studie vorzeitig beendet werden sollte.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Computer-generierter Plan der randomisierten Zuteilung
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratifizierung nach Baseline HIV-RNA Konzentration (>50.000 vs. ≤50.000 Kopien/mL)
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	Behandlungszuteilung wurde geheimgehalten vor Prüfarzten, Studienzentrums-Personal, Patienten, Monitoren und Zentrallabor-Personal indem verblindete Zugangscodes verwendet wurden. Die Zuteilung der Behandlung wurde mittels IVRS durchgeführt.
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?	Studienzentrums-Personal griff auf das IVRS zu und teilte die Patienten gemäß computer-generiertem Plan der randomisierten Zuteilung zufällig zu.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) ja b) ja c) ja
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die beiden Interventionen, Raltegravir und Efavirenz, gehören unterschiedlichen Substanzklassen an und hatten keine beabsichtigte Ähnlichkeit miteinander.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<u>Primärer Endpunkt:</u> <ul style="list-style-type: none"> • Primäre Analyse: per Protokoll • Nichtunterlegenheitshypothese zur 12%-

Item ^a	Charakteristikum	Studieninformation
		<p>Grenze (95%-KIs wurden nach Mietتين und Nurminen [108] berechnet)</p> <ul style="list-style-type: none"> • Fehlende Werte wurden als Versagen gewertet (außer direkte vor und nach dem fehlenden Wert war vRNA <50 Kopien/mL, dann wurde der Wert als fehlend gewertet) • Observed-Failure-Methode als sekundäre Analysemethode • Im Falle eines Abbruchs aufgrund fehlender Wirksamkeit wurden Baseline-Werte fortgeschrieben <p><u>Sekundäre Endpunkte:</u></p> <ul style="list-style-type: none"> • CD4-Zellzahl: observed-failure Methode, post hoc t-Test • Time to virologic response: Kaplan-Meier, Log-Rank Test • Unerwünschte Ereignisse: Fisher's exakter Test:
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Subgruppenanalysen wurden basierend auf der observed-failure Methode durchgeführt
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe unten
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) Raltegravir: n=282, Efavirenz: n=284 b) Raltegravir: n=281, Efavirenz: n=282 c) Raltegravir: n=280, Efavirenz: n=281
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Raltegravir (n=24):</p> <ul style="list-style-type: none"> • Fehlende Wirksamkeit: n=4 • Unerwünschte Ereignisse: n=8 • Lost-to-follow-up: n=3 • Andere Gründe: n=9 <p>Efavirenz (n=35):</p> <ul style="list-style-type: none"> • Fehlende Wirksamkeit: n=2 • Unerwünschte Ereignisse: n=17 • Lost-to-follow-up: n=7 • Andere Gründe: n=9
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	14. September 2006 - 05. Juni 2008

Item ^a	Charakteristikum	Studieninformation
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete regulär
a: nach CONSORT 2010.		

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

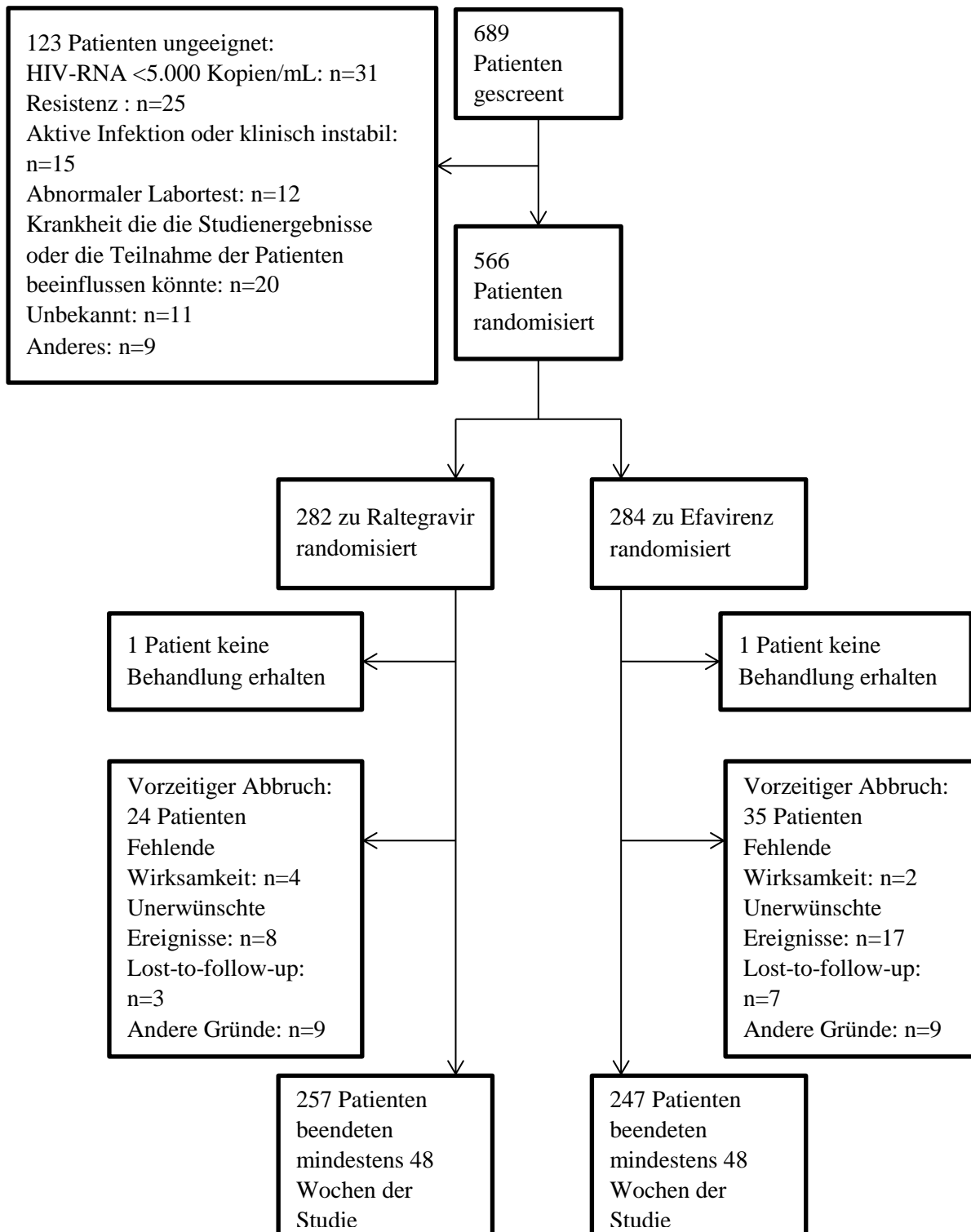


Abbildung 30: Flow Chart der Studie Lennox (2009)

Tabelle 4-217 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0123 (nach TREND-Statement)

Item ^a	Charakteristikum	Studieninformation
-	Titel und Zusammenfassung	
1	<p>Titel</p> <p>Zusammenfassung</p>	<p><i>A Phase 3B Open-Label Pilot Study to Evaluate Switching from a Regimen Consisting of Raltegravir plus Emtricitabine/Tenofovir Disoproxil Fumarate Fixed-Dose Combination (FTC/TDF) to the Elvitegravir/Cobicistat/Emtricitabine/Tenofovir Disoproxil Fumarate (EVG/COBI/FTC/TDF) Single-Tablet Regimen (STR) in Virologically-Suppressed, HIV-1 Infected Patients</i></p> <p>Einarmige, offene, multizentrische Phase 3B-Studie zur Untersuchung der Umstellung eines Regimes von Raltegravir in Kombination mit einer Fixkombination aus Emtricitabin/Tenofovirdisoproxil (FTC/TDF) auf ein Single-Tablet-Regimen (STR) bestehend aus Elvitegravir/Cobicistat/Emtricitabin/Tenofovirdisoproxil (EVG/COBI/FTC/TDF) bei virussupprimierten HIV-1-infizierten Patienten</p> <p>Studiensponsor ist Gilead Sciences, Inc.</p>
	Einleitung	
2	Hintergrund	<p>Die Ziele einer antiretroviralen Therapie sind die Unterdrückung der viralen Replikation, die Erhöhung der CD4-Zellzahl (immunologische Rekonstitution) und die Verzögerung der Krankheitsprogression. Trotz der insgesamt guten Verträglichkeit weisen antiretrovirale Substanzen Nebenwirkungen, beispielsweise im Bereich des zentralen Nervensystems, auf. Je einfacher die Einnahme der Medikation ist, desto günstiger ist dies für die Adhärenz.</p> <p>EVG/COBI/FTC/TDF wurde als zusätzliche Behandlungsoption für nicht-vorbehandelte Patienten entwickelt. EVG/COBI/FTC/TDF zeichnet sich neben sehr guter Wirksamkeit durch ein günstiges Nebenwirkungsprofil und die Formulierung als Single-Tablet-Regimen aus – beide Faktoren sollen die Adhärenz positiv unterstützen.</p> <p>Primärer Endpunkt: Untersuchung der virologischen Wirksamkeit des EVG/COBI/FTC/TDF Single-Tablet-Regimens bei Patienten nach Umstellung von RAL+FTC/TDF anhand der Beibehaltung der virologischen Suppression (HIV-1-RNA<50 Kopien/mL) zu Woche 12</p>
	Methoden	
3	Patienten (zum Beispiel. Einschlusskriterien, sampling Methode, Rekrutierungsplan)	<p><u>Diagnose und Kriterien für den Einschluss in die Studie</u></p> <p>Virologisch supprimierte Patienten mit HIV-1 unter einer antiretroviralen Therapie bestehend aus RAL+ FTC/TDF .</p> <p>Die multizentrische Studie wurde an 7 Zentren in den USA durchgeführt. Die Fallzahl war n=48</p>

Item ^a	Charakteristikum	Studieninformation
		<p><u>Einschlusskriterien:</u></p> <ul style="list-style-type: none"> • Schriftliche Patienteneinwilligung • Dauerhafte virologische Suppression unter RAL+FTC/TDF für ≥6 Monate <ul style="list-style-type: none"> ○ HIV-1 RNA <50 Kopien/mL ○ Keine HIV-1 RNA >50 Kopien/mL an zwei aufeinander folgenden Messungen, nachdem <50 Kopien/mL erreicht wurde • HIV-1 RNA <50 Kopien/mL zum Zeitpunkt des Screenings • Keine bekannten Resistenzen gegen eine der Substanzen der Studienmedikation (Genotyp) • Normales EKG • Hepatische Transaminasen (AST und ALT) ≤5* ULN (über der oberen Grenze des Normalbereichs) • Gesamtbilirubin ≤1,5 mg/dL • Adäquate hämatologische Funktion (absolute Zahl der neutrophilen Granulozyten ≥1.000/mm³, Thrombozyten ≥50.000/mm³, Hämoglobin ≥8,5 g/dL) • Serumamylase ≤5* ULN • Adäquate Nierenfunktion: eGFR ≥90 mL/min gemäß Cockcroft-Gault (CG) Formel: Männer: eGFR (mL/min) = $\frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})}$ Frauen: eGFR (mL/min) = $\frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})} \cdot 0,85$ • Frauen in gebärfähigem Alter müssen eine adäquate Verhütungsmethode verwenden • Männer müssen für die Dauer der Studie ebenfalls ein adäquates Verhütungsmittel verwenden • Patienten müssen mindestens 18 Jahre alt sein <p><u>Ausschlusskriterien</u></p> <ul style="list-style-type: none"> • Auftreten einer neuen Aids-assoziierten Diagnose binnen der letzten 21 Tage vor dem Screening • Stillende Frauen • Positiver Schwangerschaftstest • Patienten mit akuter oder chronischer Hepatitis B oder C • Patienten mit einer dekompensierten Zirrhose • Implantierter Defibrillator oder Herzschrittmacher • Alkoholmissbrauch oder der Missbrauch von Substanzen, die nach Einschätzung des Prüfarztes möglicherweise mit dem Studienpräparat interferieren • Vorgeschichte eine Malignität binnen der letzten 5 Jahre • Aktive, schwerwiegende Infektion, neben HIV-1, die die Gabe

Item ^a	Charakteristikum	Studieninformation
		<p>eines parenteralen Antibiotikums oder eine antimykotische Therapie binnen der letzten 21 Tage vor Screening nötig macht</p> <ul style="list-style-type: none"> • Erhalt eines investigativen Medikaments • Teilnahme an einer weiteren klinischen Studie, die nicht vom Sponsor genehmigt wurde • Behandlung mit Immunsuppressiva oder Erhalt einer Chemotherapie binnen 3 Monaten vor dem Screening • Irgendein klinischer Zustand oder vorangegangene Behandlung, die den Probanden nach Meinung des Prüfarztes ungeeignet für die klinische Studie macht • Einnahme von Medikamenten, die eine Kontraindikation mit dem Studienprodukt haben oder bei der Behandlung nicht empfohlen werden <p>Die Studie hat eine geplante Beobachtungsdauer von 48 Wochen und läuft derzeit noch. Der erste Patient erhielt die erste Dosis am 10. Februar 2012.</p> <p><u>Gründe für Studienabbruch</u> Es gab keine Studienabbrüche.</p>
4	Intervention	<p><u>Studienmedikation</u> Grüne Tabletten, die 150 mg EVG, 150 mg COBI, 200 mg FTC und 300 mg TDF enthalten. Die Gabe erfolgt einmal täglich zur selben Uhrzeit während einer Mahlzeit.</p>
5	Ziele	<p>Einarmige, offene, multizentrische Phase 3B-tudie zur Untersuchung der Umstellung eines Regimes von Raltegravir in Kombination mit einer Fixkombination aus Emtricitabin/Tenofovirdisoproxil (FTC/TDF) auf ein Single-Tablet-Regimen (STR) bestehend aus Elvitegravir/Cobicistat/Emtricitabin/Tenofovirdisoproxil (EVG/COBI/FTC/TDF) bei virussupprimierten HIV-1-infizierten Patienten</p>
6	Zielkriterien (z.B. Methoden zur Datengewinnung)	<p><u>Primäres Ziel:</u> Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF als STR gemessen anhand HIV-1 RNA-Viruslast <50 Kopien/mL zu Woche 12</p> <p><u>Sekundäre Ziele:</u></p> <ul style="list-style-type: none"> • Bewertung der Sicherheit und Verträglichkeit von EVG/COBI/FTC/TDF STR über 24 und 48 Wochen Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF STR anhand der HIV-1 RNA Viruslast <50 Kopien/mL zu Woche 24 und 48
7	Fallzahlbestimmung	<p>Auswahl der Fallzahl anhand der Realisierbarkeit. Es sollen 50 Patienten in die Studie aufgenommen werden.</p>
8	Zuordnungsmethode	<p>Es fand keine Zuordnung zu bestimmten Gruppen statt, da es sich um eine einarmige Studie handelt. Alle Patienten erhielten EVG/COBI/FTC/TDF.</p>

Item ^a	Charakteristikum	Studieninformation
9	Verblindung	Es fand keine Verblindung statt (<i>open-label</i>)
10	Analyseeinheit (z.B. Beschreibung der kleinsten Analyseeinheit, analytische Methode)	<u>Kleinste Analyse-Einheit</u> : Patient
11	Statistische Methoden	<p><u>Volles Analyseset</u>: Enthält alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben.</p> <p><u>Sicherheits-Analyse-Set</u>: Enthält alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben. Hier werden alle Daten bis 30 Tage nach Gabe der letzten Studienmedikation erfasst.</p> <p><u>Verwendete Methoden</u>: Zur Bewertung der Baseline-Charakteristika, der Wirksamkeit und der Sicherheit werden deskriptive Methoden verwendet Für den primären Endpunkt wird die Veränderung vom Ausgangswert zu im Protokoll vordefinierten Zeitpunkten angegeben. Für sekundäre Endpunkte werden die Patientenanteile mit 95%-Konfidenzintervall zusammengefasst. Hier wird jeweils das volle Analyse-Set verwendet. Für die Bewertung der Sicherheit wird das Sicherheits-Analyse-Set verwendet.</p>
Resultate		
12	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	<p>Für Studieneinschluss evaluierte Patienten: 48 Gescreente Patienten: 48 Anzahl der Patienten, die mindestens eine Dosis von EVG/COBI/FTC/TDF erhalten haben: 48 Anzahl der Patienten, die die Behandlung bis Woche 12 beendet haben: 48</p> <p>Kein Patient hat die Behandlung abgebrochen.</p>
13	Rekrutierung	Die Studie wurde zwischen Februar 2012 bis Oktober 2013 durchgeführt.
a: TREND Statement (http://www.cdc.gov/trendstatement/)		

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

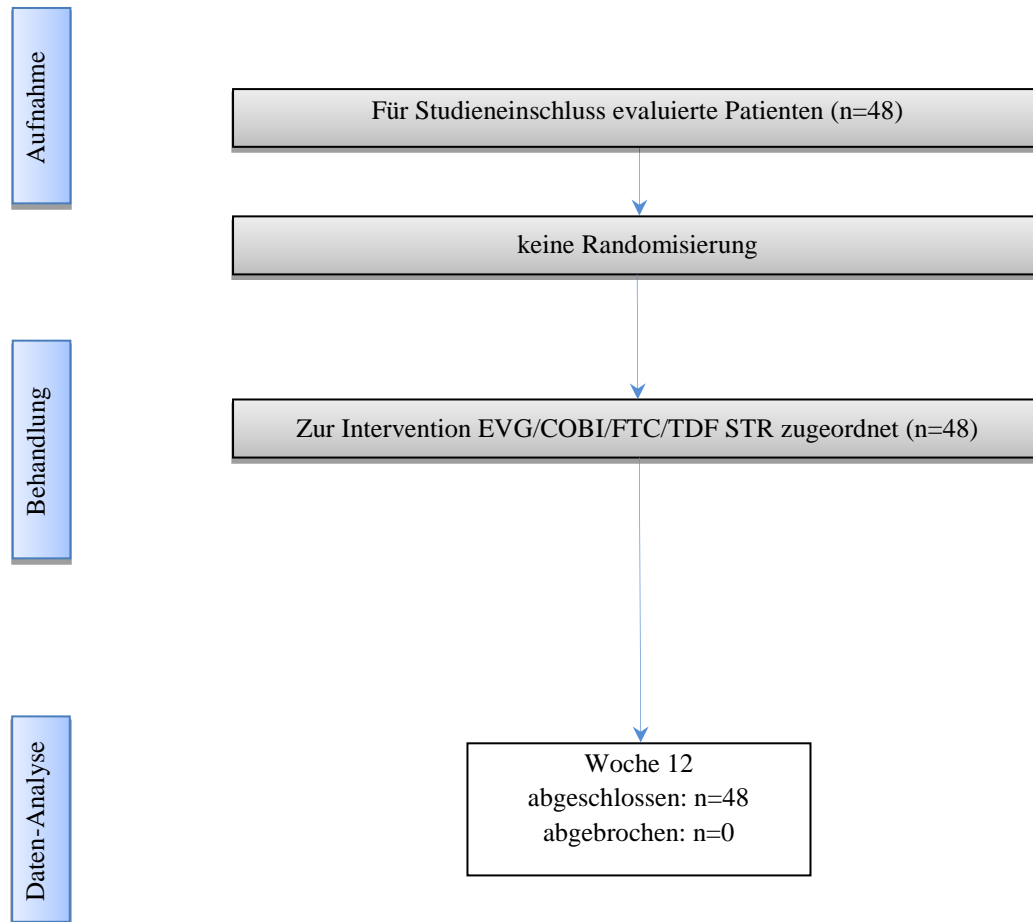


Abbildung 31: Flow-Chart für Studie GS-US-236-0123

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 4-218 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GS-US-236-0102

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: GS-US-236-0102

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
GS-US-236-0102 Clinical trial report zu Woche 48 und zu Woche 96	S102
GS-US-236-0102 Clinical trial protocol	S102 CTP
GS-US-236-0102 Statistical analysis plan zu Woche 48 und zu Woche 96	S102 SAP

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

In Studie S102 werden keine Aussagen zur Erzeugung der Randomisierungsliste gemacht. Da jedoch ein IVRS/IWRS zur Randomisierung verwendet wurde, wird davon ausgegangen dass die Erzeugung der Randomisierungsliste adäquat durchgeführt wurde.

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:

In Studie S102 werden keine Aussagen zur Verdeckung der Gruppenzuteilung gemacht. Da jedoch ein IVRS/IWRS verwendet wurde, wird davon ausgegangen dass die Gruppenzuteilung verdeckt war.

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:

Patienten waren gegenüber der Zuteilung verblindet

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:

Behandler bzw. Prüfarzte waren gegenüber der Zuteilung verblindet

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: Virologisches Ansprechen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Endpunkterheber waren verblindet.

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 Analyse-Set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien,

statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung der CD4-Zellzahl**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Endpunkterheber waren verblindet.

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 Analyse-Set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.
- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

(hier werden folgende Endpunkte gemeinsam bewertet: Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse, unerwünschte Ereignisse von Interesse und häufigste unerwünschte Ereignisse (SOC). Es ist möglich diese Endpunkte zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und die gleiche Anzahl an Patienten in die Analyse einging.)

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

Patienten und Prüfarzte waren jederzeit gegenüber der Gruppenzuteilung verblindet

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 Analyse-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
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-219 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GS-US-236-0103

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: GS-US-236-0103

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
GS-US-236-0103 Clinical trial report zu Woche 48 und zu Woche 96	S103
GS-US-236-0103 Clinical trial protocol	S103 CTP
GS-US-236-0103 Statistical analysis plan zu Woche 48	S103 SAP

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

In Studie S103 werden keine Aussagen zur Erzeugung der Randomisierungsliste gemacht. Da jedoch ein IVRS/IWRS zur Randomisierung verwendet wurde, wird davon ausgegangen dass die Erzeugung der Randomisierungsliste adäquat durchgeführt wurde.

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:

In Studie S103 werden keine Aussagen zur Verdeckung der Gruppenzuteilung gemacht. Da jedoch ein IVRS/IWRS verwendet wurde, wird davon ausgegangen dass die Gruppenzuteilung verdeckt war.

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:

Patienten waren gegenüber der Zuteilung verblindet

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:

Behandler bzw. Prüfarzte waren gegenüber der Zuteilung verblindet

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: Virologisches Ansprechen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Endpunkterheber waren verblindet.

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 Analyse-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: Zeit bis zum Verlust des virologischen Ansprechens**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Endpunkterheber waren verblindet.

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 Analyse-Set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- unplausible Angaben*
- Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung der CD4-Zellzahl

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Endpunkterheber waren verblindet.

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 Analyse-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

(hier werden folgende Endpunkte gemeinsam bewertet: Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse, unerwünschte Ereignisse von Interesse und häufigste unerwünschte Ereignisse (SOC). Es ist möglich diese Endpunkte zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und die gleiche Anzahl an Patienten in die Analyse einging.)

1. Verblindung der Endpunkterheber

*Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.
In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.*

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Prüfarzte waren jederzeit gegenüber der Gruppenzuteilung verblindet

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 Analyse-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 4-220 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GS-US-236-0104

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: GS-US-236-0104

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
GS-US-236-0104 Clinical trial report zu Woche 96	S104
GS-US-236-0104 Clinical trial protocol	S104 CTP
GS-US-236-0102 Statistical analysis plan zu Woche 48, Woche 48+ und Woche 96	S104 SAP

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

In Studie S104 werden keine Aussagen zur Erzeugung der Randomisierungsliste gemacht. Da jedoch ein IVRS/IWRS zur Randomisierung verwendet wurde, wird davon ausgegangen dass die Erzeugung der Randomisierungsliste adäquat durchgeführt wurde.

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:

In Studie S104 werden keine Aussagen zur Verdeckung der Gruppenzuteilung gemacht. Da jedoch ein IVRS/IWRS verwendet wurde, wird davon ausgegangen dass die Gruppenzuteilung verdeckt war.

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:

Patienten waren gegenüber der Zuteilung verblindet

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:

Behandler bzw. Prüfarzte waren gegenüber der Zuteilung verblindet

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: Virologisches Ansprechen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Endpunkterheber waren verblindet.

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 Analyse-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:

Der Endpunkt virologisches Ansprechen (mittels Imputationsmethode TLOVR) wurde für Studie S104 nachberechnet. Da dies aber, unabhängig von den Daten, lediglich aus Vergleichsgründen mit Studie S102 geschah, wird das Verzerrungspotenzial für den Endpunkt dennoch als niedrig eingestuft.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung der CD4-Zellzahl

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Endpunkterheber waren verblindet.

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 Analyse-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

(hier werden folgende Endpunkte gemeinsam bewertet: Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, unerwünschte Ereignisse Grad 3-4, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse, unerwünschte Ereignisse von Interesse und häufigste unerwünschte Ereignisse (SOC). Es ist möglich diese Endpunkte zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und die gleiche Anzahl an Patienten in die Analyse einging.)

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

Patienten und Prüfärzte waren jederzeit gegenüber der Gruppenzuteilung verblindet

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 Analyse-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:

Wichtige unerwünschte Ereignisse (neurologische & psychiatrische Ereignisse und Hautausschläge) wurden in Studie S104 basierend auf verblindeten Daten ebendieser Studie S104 und Studie S102 zu der Analyse hinzugefügt. Angaben zum Zeitpunkt der Aufnahme der Endpunkte wurden dabei nicht gemacht. Unerwünschte Ereignisse von Interesse wurden weder im Protokoll noch im statistischen Analyseplan präspezifiziert. Eine ergebnisgesteuerte Berichterstattung kann somit nicht ausgeschlossen werden. Das Verzerrungspotenzial wird bei den entsprechenden Endpunkten separat berücksichtigt.

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 4-221 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie STARTMRK

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: STARTMRK

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Lennox et al. (2009)	
Lennox et al. (2010)	

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Patienten waren gegenüber der Zuteilung verblindet

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:

Behandler bzw. Prüfärzte waren gegenüber der Zuteilung verblindet

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b)

oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die

ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Da kein Studienprotokoll oder statistischer Analyseplan vorliegt können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden.

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: Virologisches Ansprechen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Viruslast wurde in einem Zentrallabor untersucht, welches gegenüber der Gruppenzuteilung verblindet war

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 Analyse-Set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Veränderung der CD4-Zellzahl

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die CD4-Zellzahl wurde in einem Zentrallabor untersucht, welches gegenüber der Gruppenzuteilung verblindet war

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 Analyse-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:

Da kein Studienprotokoll oder statistischer Analyseplan vorliegt können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden.

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

(hier werden folgende Endpunkte gemeinsam bewertet: Mortalität, Gesamtrate unerwünschter Ereignisse, schwerwiegende unerwünschte Ereignisse, Therapieabbrüche aufgrund unerwünschter Ereignisse, wichtige unerwünschte Ereignisse und häufigste unerwünschte Ereignisse (SOC). Es ist möglich diese Endpunkte zusammengefasst zu bewerten, da sie alle im Rahmen der Safety-Analyse analysiert wurden und die gleiche Anzahl an Patienten in die Analyse einging.)

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

Patienten und Prüfärzte waren jederzeit gegenüber der Gruppenzuteilung verblindet

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 Analyse-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:

Da kein Studienprotokoll oder statistischer Analyseplan vorliegt können keine sicheren Aussagen zur ergebnisgesteuerten Berichterstattung gemacht werden.

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: