

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

*Elvitegravir/Cobicistat/Emtricitabin/Tenofoviridisoproxil  
(Stribild<sup>®</sup>)*

Gilead Sciences GmbH

**Modul 4 A**

*Zur Behandlung der Infektion mit dem Humanen  
Immundefizienzvirus 1 (HIV-1) bei Erwachsenen im  
Alter von 18 Jahren und darüber, die nicht  
antiretroviral vorbehandelt sind*

Medizinischer Nutzen und  
medizinischer Zusatznutzen,  
Patientengruppen mit therapeutisch  
bedeutsamem Zusatznutzen

# Inhaltsverzeichnis

	Seite
<b>Tabellenverzeichnis</b> .....	<b>4</b>
<b>Abbildungsverzeichnis</b> .....	<b>11</b>
<b>Abkürzungsverzeichnis</b> .....	<b>14</b>
4.1 Zusammenfassung der Inhalte von Modul 4.....	18
4.2 Methodik.....	30
4.2.1 Fragestellung.....	30
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	33
4.2.3 Informationsbeschaffung.....	36
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	37
4.2.3.2 Bibliografische Literaturrecherche.....	37
4.2.3.3 Suche in Studienregistern.....	38
4.2.3.4 Selektion relevanter Studien.....	38
4.2.4 Bewertung der Aussagekraft der Nachweise.....	39
4.2.5 Informationssynthese und -analyse.....	41
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	41
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	42
4.2.5.3 Meta-Analysen.....	50
4.2.5.4 Sensitivitätsanalysen.....	52
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	54
4.2.5.6 Indirekte Vergleiche.....	56
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	57
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	58
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	58
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	58
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	61
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	63
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	63
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	65
4.3.1.2.1 Studiendesign und Studienpopulationen.....	65
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	77
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	79
4.3.1.3.1 Endpunkte– RCT.....	80
4.3.1.3.1.1 Virologisches Ansprechen– RCT.....	81
4.3.1.3.1.2 Zeit bis Verlust des virologischen Ansprechens– RCT.....	94
4.3.1.3.1.3 Veränderung der CD4-Zellzahl– RCT.....	98
4.3.1.3.1.4 Unerwünschte Ereignisse– RCT.....	106
4.3.1.3.2 Subgruppenanalysen – RCT.....	138

4.3.1.3.2.1	Virologisches Ansprechen (Snapshot-Algorithmus; in Studie S102 präspezifizierte Subgruppenanalysen).....	161
4.3.1.3.2.2	Virologisches Ansprechen (TLOVR-Algorithmus) .....	167
4.3.1.3.2.3	Veränderung der CD4-Zellzahl .....	168
4.3.1.3.2.4	Gesamtrate unerwünschter Ereignisse .....	169
4.3.1.3.2.5	Gesamtrate schwerwiegender unerwünschter Ereignisse .....	169
4.3.1.3.2.6	Therapieabbrüche aufgrund unerwünschter Ereignisse.....	170
4.3.1.3.2.7	Unerwünschte Ereignisse Grad 3-4 .....	172
4.3.1.3.2.8	Neurologische & psychiatrische Ereignisse .....	174
4.3.1.3.2.9	Hautausschläge .....	176
4.3.1.3.2.10	Renale Ereignisse .....	180
4.3.1.3.2.11	Knochenfrakturen .....	181
4.3.1.3.2.12	Erkrankungen des Gastrointestinaltrakts (SOC).....	182
4.3.1.3.2.13	Diarrhoe .....	185
4.3.1.3.2.14	Übelkeit .....	187
4.3.1.3.2.15	Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) .....	188
4.3.1.3.2.16	Fatigue .....	190
4.3.1.3.2.17	Erkrankungen des Nervensystems (SOC) .....	191
4.3.1.3.2.18	Kopfschmerzen .....	193
4.3.1.3.2.19	Psychiatrische Erkrankungen (SOC) .....	196
4.3.1.3.2.20	Abnorme Träume .....	197
4.3.1.3.2.21	Depression .....	200
4.3.1.3.2.22	Erkrankungen der Haut und des Unterhautzellgewebes (SOC) ...	201
4.3.1.3.2.23	Hautausschlag .....	205
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	206
4.3.2	Weitere Unterlagen.....	223
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien .....	223
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche .....	223
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	223
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen .....	224
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT .....	224
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT .....	226
4.3.2.2	Nicht randomisierte vergleichende Studien .....	227
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien .....	227
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	227
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien .....	228
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	228
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien .....	229
4.3.2.3	Weitere Untersuchungen.....	229
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen .....	230
4.3.2.3.2	Charakteristika der weiteren Untersuchungen .....	230
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen .....	230
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen .....	230

4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen .....	231
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen .....	231
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	232
4.4.1	Beurteilung der Aussagekraft der Nachweise .....	232
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	232
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht .....	268
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	268
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte .....	269
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	269
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	269
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen .....	269
4.5.4	Verwendung von Surrogatendpunkten .....	270
4.6	Liste der eingeschlossenen Studien.....	275
4.7	Referenzliste.....	277
<b>Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche .....</b>		<b>283</b>
<b>Anhang 4-B : Suchstrategien – Suche in Studienregistern.....</b>		<b>285</b>
<b>Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund .....</b>		<b>287</b>
<b>Anhang 4-D : Liste der abgebrochenen Studien .....</b>		<b>288</b>
<b>Anhang 4-E : Liste der laufenden Studien.....</b>		<b>289</b>
<b>Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....</b>		<b>290</b>
<b>Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten .....</b>		<b>317</b>

**Tabellenverzeichnis**

	<b>Seite</b>
Tabelle 4-1: Gegenüberstellung von Zusatznutzen und Schaden von EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF .....	21
Tabelle 4-2: Übersicht der Kriterien für den Studieneinschluss für den Vergleich von EVG/COBI/FTC/TDF gegen EFV/FTC/TDF.....	34
Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel .....	59
Tabelle 4-4: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	61
Tabelle 4-5: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel .....	63
Tabelle 4-6: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	64
Tabelle 4-7: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel .....	66
Tabelle 4-8: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	68
Tabelle 4-9: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Demografie, Safety Analyse-Set) .....	68
Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Rasse, Safety Analyse-Set) .....	69
Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Ethnische Abstammung, Safety Analyse-Set).....	69
Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (HIV-1-RNA, Safety Analyse-Set) .....	70
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (CD4-Zellzahl, Safety Analyse-Set) .....	70
Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (HIV-Status, Safety Analyse-Set) .....	70
Tabelle 4-15: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Adhärenz, Safety Analyse-Set).....	71
Tabelle 4-16: Charakterisierung der Studienpopulationen (Patientendisposition) – RCT mit dem zu bewertenden Arzneimittel .....	71
Tabelle 4-17: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	77
Tabelle 4-18: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel .....	79
Tabelle 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel .....	79

Tabelle 4-20: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel .....	80
Tabelle 4-21: Operationalisierung von virologisches Ansprechen .....	81
Tabelle 4-22: Bewertung des Verzerrungspotenzials für virologisches Ansprechen in RCT mit dem zu bewertenden Arzneimittel .....	81
Tabelle 4-23: Ergebnisse für virologisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set) .....	86
Tabelle 4-24: Ergebnisse für virologisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set) .....	89
Tabelle 4-25: Ergebnisse für virologisches Versagen (Wirksamkeit) aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 48) .....	90
Tabelle 4-26: Ergebnisse für virologisches Versagen (Wirksamkeit) aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 96) .....	91
Tabelle 4-27: Operationalisierung von Zeit bis Verlust des virologischen Ansprechens .....	94
Tabelle 4-28: Bewertung des Verzerrungspotenzials für Zeit bis Verlust des virologischen Ansprechens in RCT mit dem zu bewertenden Arzneimittel.....	95
Tabelle 4-29: Ergebnisse für Zeit bis Verlust des virologischen Ansprechens aus RCT mit dem zu bewertenden Arzneimittel.....	96
Tabelle 4-30: Operationalisierung von Veränderung der CD4-Zellzahl.....	98
Tabelle 4-31: Bewertung des Verzerrungspotenzials für Veränderung der CD4-Zellzahl in RCT mit dem zu bewertenden Arzneimittel .....	99
Tabelle 4-32: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 48) .....	100
Tabelle 4-33: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set, Woche 48) .....	101
Tabelle 4-34: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 60) .....	102
Tabelle 4-35: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 96) .....	103
Tabelle 4-36: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set, Woche 96) .....	103
Tabelle 4-37: Operationalisierung von Unerwünschte Ereignisse .....	106
Tabelle 4-38: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel .....	108
Tabelle 4-39: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48).....	111
Tabelle 4-40: Ergebnisse für wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48) .....	112
Tabelle 4-41: Ergebnisse für unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48) .....	112

Tabelle 4-42: Ergebnisse für Erkrankungen des Gastrointestinaltrakts (SOC, Grenze: $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48) ...	113
Tabelle 4-43: Ergebnisse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC, Grenze: $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48) .....	113
Tabelle 4-44: Ergebnisse für Erkrankungen des Nervensystems (SOC, Grenze: $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48).....	114
Tabelle 4-45: Ergebnisse für psychiatrische Erkrankungen (SOC, Grenze: $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48).....	115
Tabelle 4-46: Ergebnisse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC, Grenze: $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48) .....	116
Tabelle 4-47: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 60) .....	121
Tabelle 4-48: Ergebnisse für wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 60) .....	121
Tabelle 4-49: Ergebnisse für unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 60) .....	122
Tabelle 4-50: Ergebnisse für die häufigsten unerwünschten Ereignisse (SOC, Grenze: $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 60) ...	122
Tabelle 4-51: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96).....	125
Tabelle 4-52: Ergebnisse für wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96) .....	125
Tabelle 4-53: Ergebnisse für unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96) .....	126
Tabelle 4-54: Ergebnisse für die häufigsten unerwünschten Ereignisse (SOC, Grenze: $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96) ...	126
Tabelle 4-55: Übersicht über alle präspezifizierten (fett gedruckt) und post-hoc durchgeführten Interaktionstests .....	140
Tabelle 4-56: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Alter (Woche 48) .....	161
Tabelle 4-57: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Geschlecht (Woche 48).....	162
Tabelle 4-58: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Rasse (Woche 48) .....	162
Tabelle 4-59: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Ausgangsviruslast (Woche 48) .....	163
Tabelle 4-60: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48) .....	163

Tabelle 4-61: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Therapieadhärenz (Woche 48).....	164
Tabelle 4-62: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Alter (Woche 96).....	164
Tabelle 4-63: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Geschlecht (Woche 96).....	165
Tabelle 4-64: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Rasse (Woche 96).....	165
Tabelle 4-65: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Ausgangsviruslast (Woche 96).....	166
Tabelle 4-66: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96).....	166
Tabelle 4-67: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Therapieadhärenz (Woche 96).....	167
Tabelle 4-68: Subgruppenanalyse für virologisches Ansprechen (TLOVR), stratifiziert nach Rasse (Woche 96).....	167
Tabelle 4-69: Subgruppenanalyse für virologisches Ansprechen (TLOVR), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96).....	168
Tabelle 4-70: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach Therapieadhärenz (Woche 96).....	168
Tabelle 4-71: Subgruppenanalyse für Gesamtrate unerwünschter Ereignisse, stratifiziert nach Alter (Woche 96).....	169
Tabelle 4-72: Subgruppenanalyse für Gesamtrate schwerwiegender unerwünschter Ereignisse, stratifiziert nach Rasse (Woche 96).....	169
Tabelle 4-73: Subgruppenanalyse für Therapieabbrüche aufgrund unerwünschter Ereignisse, stratifiziert nach Rasse (Woche 48).....	170
Tabelle 4-74: Subgruppenanalyse für Therapieabbrüche aufgrund unerwünschter Ereignisse, stratifiziert nach Rasse (Woche 96).....	171
Tabelle 4-75: Subgruppenanalyse für Unerwünschte Ereignisse Grad 3-4, stratifiziert nach Alter (Woche 48).....	172
Tabelle 4-76: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Rasse (Woche 48).....	173
Tabelle 4-77: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Rasse (Woche 96).....	174
Tabelle 4-78: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Therapieadhärenz (Woche 48).....	174
Tabelle 4-79: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 60).....	175
Tabelle 4-80: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Alter (Woche 96).....	176
Tabelle 4-81: Subgruppenanalyse für Hautausschläge, stratifiziert nach Alter (Woche 48).....	176



Tabelle 4-82: Subgruppenanalyse für Hautausschläge, stratifiziert nach Rasse (Woche 48)	177
Tabelle 4-83: Subgruppenanalyse für Hautausschläge, stratifiziert nach Ausgangsviruslast (Woche 60).....	178
Tabelle 4-84: Subgruppenanalyse für Hautausschläge, stratifiziert nach Therapieadhärenz (Woche 60).....	179
Tabelle 4-85: Subgruppenanalyse für Hautausschläge, stratifiziert nach Rasse (Woche 96)	179
Tabelle 4-86: Subgruppenanalyse für renale Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 48).....	180
Tabelle 4-87: Subgruppenanalyse für renale Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 96).....	181
Tabelle 4-88: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96) .....	181
Tabelle 4-89: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48) .....	182
Tabelle 4-90: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Geschlecht (Woche 60) .....	183
Tabelle 4-91: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Ausgangsviruslast (Woche 96) .....	183
Tabelle 4-92: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96) .....	184
Tabelle 4-93: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Therapieadhärenz (Woche 96) .....	184
Tabelle 4-94: Subgruppenanalyse für Diarrhoe, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48) .....	185
Tabelle 4-95: Subgruppenanalyse für Diarrhoe, stratifiziert nach Ausgangsviruslast (Woche 60).....	186
Tabelle 4-96: Subgruppenanalyse für Diarrhoe, stratifiziert nach Therapieadhärenz (Woche 60).....	186
Tabelle 4-97: Subgruppenanalyse für Diarrhoe, stratifiziert nach Therapieadhärenz (Woche 96).....	187
Tabelle 4-98: Subgruppenanalyse für Übelkeit, stratifiziert nach Ausgangsviruslast (Woche 48).....	187
Tabelle 4-99: Subgruppenanalyse für Übelkeit, stratifiziert nach Ausgangsviruslast (Woche 96).....	188
Tabelle 4-100: Subgruppenanalyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48) .....	188
Tabelle 4-101: Subgruppenanalyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), stratifiziert nach Geschlecht (Woche 96).....	189
Tabelle 4-102: Subgruppenanalyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96) .....	190
Tabelle 4-103: Subgruppenanalyse für Fatigue, stratifiziert nach Geschlecht (Woche 96)...	190

Tabelle 4-104: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Therapieadhärenz (Woche 48) .....	191
Tabelle 4-105: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Rasse (Woche 60) .....	192
Tabelle 4-106: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Therapieadhärenz (Woche 96) .....	192
Tabelle 4-107: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Geschlecht (Woche 48).....	193
Tabelle 4-108: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Rasse (Woche 48).....	194
Tabelle 4-109: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Therapieadhärenz (Woche 48) .....	195
Tabelle 4-110: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Alter (Woche 60).....	196
Tabelle 4-111: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Rasse (Woche 60) .....	196
Tabelle 4-112: Subgruppenanalyse für abnorme Träume, stratifiziert nach Rasse (Woche 48).....	197
Tabelle 4-113: Subgruppenanalyse für abnorme Träume, stratifiziert nach Alter (Woche 60).....	198
Tabelle 4-114: Subgruppenanalyse für abnorme Träume, stratifiziert nach Geschlecht (Woche 60).....	198
Tabelle 4-115: Subgruppenanalyse für abnorme Träume, stratifiziert nach Rasse (Woche 60).....	199
Tabelle 4-116: Subgruppenanalyse für abnorme Träume, stratifiziert nach Rasse (Woche 96).....	199
Tabelle 4-117: Subgruppenanalyse für Depression, stratifiziert nach Alter (Woche 48) .....	200
Tabelle 4-118: Subgruppenanalyse für Depression, stratifiziert nach Alter (Woche 96) .....	200
Tabelle 4-119: Subgruppenanalyse für Depression, stratifiziert nach Geschlecht (Woche 96).....	201
Tabelle 4-120: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Rasse (Woche 48) .....	201
Tabelle 4-121: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Therapieadhärenz (Woche 48) .....	202
Tabelle 4-122: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Geschlecht (Woche 60) .....	203
Tabelle 4-123: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Ausgangsviruslast (Woche 60) .....	204
Tabelle 4-124: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Rasse (Woche 96) .....	204

Tabelle 4-125: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Therapieadhärenz (Woche 96).....	205
Tabelle 4-126: Subgruppenanalyse für Hautausschlag, stratifiziert nach Alter (Woche 48). 205	
Tabelle 4-127: Subgruppenanalyse für Hautausschlag, stratifiziert nach Ausgangsviruslast (Woche 60).....	206
Tabelle 4-128: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche .....	224
Tabelle 4-129: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	224
Tabelle 4-130: Operationalisierung von <Endpunkt xxx>.....	225
Tabelle 4-131: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche .....	225
Tabelle 4-132: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	226
Tabelle 4-133: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien .....	228
Tabelle 4-134: Operationalisierung von <Endpunkt xxx>.....	228
Tabelle 4-135: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien .....	229
Tabelle 4-136: Operationalisierung von <Endpunkt xxx>.....	230
Tabelle 4-137: Ausmaß des Zusatznutzen von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF auf Endpunktebene .....	235
Tabelle 4-138: Gegenüberstellung von Zusatznutzen und Schaden von EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF .....	266
Tabelle 4-139: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	268
Tabelle 4-140: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens .....	269
Tabelle 4-141 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel.....	288
Tabelle 4-142 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	289
Tabelle 4-143 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0102.....	291
Tabelle 4-144 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0104.....	304
Tabelle 4-145 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GS-US-236-0102.....	318
Tabelle 4-146 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GS-US-236-0104.....	336

**Abbildungsverzeichnis**

	<b>Seite</b>
Abbildung 1: Relation zwischen Veränderung der CD4-Zellzahl und der Hazard für klinische Ereignisse (Aids/Tod) [42] .....	47
Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel .....	62
Abbildung 3: Meta-Analyse für virologisches Ansprechen (Snapshot) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, ITT-Set).....	93
Abbildung 4: Meta-Analyse für virologisches Ansprechen (Snapshot) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, PP Set).....	93
Abbildung 5: Meta-Analyse für virologisches Ansprechen (TLOVR) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, ITT-Set).....	94
Abbildung 6: Kaplan-Meier Kurven für Zeit bis zum Verlust des virologischen Ansprechens, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Studie S102).....	97
Abbildung 7: Meta-Analyse für Zeit bis Verlust des virologischen Ansprechens aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, ITT-Set).....	98
Abbildung 8: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, MWD, ITT; basierend auf den nicht-adjustierten Mittelwerten der Veränderung) .....	104
Abbildung 9: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, Hedges' g, ITT).....	104
Abbildung 10: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, MWD, PP).....	105
Abbildung 11: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 60, REM, MWD, ITT; basierend auf den nicht-adjustierten Mittelwerten der Veränderung) .....	105
Abbildung 12: Meta-Analyse für die Gesamtrate unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	129
Abbildung 13: Meta-Analyse für die Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR) .....	130
Abbildung 14: Meta-Analyse für Therapieabbrüche aufgrund unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	130
Abbildung 15: Meta-Analyse für unerwünschte Ereignisse Grad 3-4 aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	130
Abbildung 16: Meta-Analyse für neurologische & psychiatrische Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	131
Abbildung 17: Meta-Analyse für Hautausschläge aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR) .....	131
Abbildung 18: Meta-Analyse für Knochenfrakturen aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR) .....	132

Abbildung 19: Meta-Analyse für Erkrankungen des Gastrointestinaltrakts (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	132
Abbildung 20: Meta-Analyse für Diarrhoe aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	132
Abbildung 21: Meta-Analyse für Übelkeit aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	133
Abbildung 22: Meta-Analyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	133
Abbildung 23: Meta-Analyse für Fatigue aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	134
Abbildung 24: Meta-Analyse für Erkrankungen des Nervensystems (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	134
Abbildung 25: Meta-Analyse für Schwindel aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	134
Abbildung 26: Meta-Analyse für Kopfschmerzen aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	135
Abbildung 27: Meta-Analyse für psychiatrische Erkrankungen (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	135
Abbildung 28: Meta-Analyse für abnorme Träume aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	136
Abbildung 29: Meta-Analyse für Schlaflosigkeit aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	136
Abbildung 30: Meta-Analyse für Depression aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	136
Abbildung 31: Meta-Analyse für Angst aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	137
Abbildung 32: Meta-Analyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	137
Abbildung 33: Meta-Analyse für Hautausschlag aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	138
Abbildung 34: Meta-Analyse für Therapieabbrüche aufgrund unerwünschter Ereignisse, stratifiziert nach Rasse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	171
Abbildung 35: Meta-Analyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Alter, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	172
Abbildung 36: Meta-Analyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	173
Abbildung 37: Meta-Analyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Therapieadhärenz, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	175

Abbildung 38: Meta-Analyse für Hautausschläge, stratifiziert nach Alter, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	177
Abbildung 39: Meta-Analyse für Hautausschläge, stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	178
Abbildung 40: Meta-Analyse für renale Ereignisse, stratifiziert nach Ausgangsviruslast, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	180
Abbildung 41: Meta-Analyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	182
Abbildung 42: Meta-Analyse für Diarrhoe, stratifiziert nach CD4-Zellzahl zu Baseline, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	185
Abbildung 43: Meta-Analyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), stratifiziert nach CD4-Zellzahl zu Baseline, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	189
Abbildung 44: Meta-Analyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Therapieadhärenz, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR) .....	191
Abbildung 45: Meta-Analyse für Kopfschmerzen, stratifiziert nach Geschlecht, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	193
Abbildung 46: Meta-Analyse für Kopfschmerzen, stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	194
Abbildung 47: Meta-Analyse für Kopfschmerzen, stratifiziert nach Therapieadhärenz, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	195
Abbildung 48: Meta-Analyse für abnorme Träume, stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	197
Abbildung 49: Meta-Analyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	202
Abbildung 50: Meta-Analyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Therapieadhärenz, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	203
Abbildung 51: Meta-Analyse für Hautausschlag, stratifiziert nach Alter, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR).....	206
Abbildung 52: Relation zwischen Veränderung der CD4-Zellzahl und der Hazard für klinische Ereignisse (Aids/Tod) [42] .....	275
Abbildung 53: Flow-Chart für Studie GS-US-236-0102 .....	303
Abbildung 54: Flow-Chart für Studie GS-US-236-0104 .....	316

**Abkürzungsverzeichnis**

<b>Abkürzung</b>	<b>Bedeutung</b>
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/ $\mu$ L oder Zellen/ $\text{mm}^3$ als Äquivalent verwendet)
CDC	Centers for Disease Control and Prevention
CG	Cockcroft-Gault
COBI	Cobicistat
CONSORT	Consolidated Standards of Reporting Trials
DIMDI	Deutsches Institut für Medizinische Dokumentation
EFV	Efavirenz
EG	Europäische Gemeinschaft
eGFR	geschätzte glomeruläre Filtrationsrate (estimated glomerular filtration rate)
EKG	Elektrokardiogramm
EMA	European Medicines Agency
EVG	Elvitegravir
FDA	US Food and Drug Administration
FEM	Modell mit festen Effekten (Fixed effect model)
FSH	Serum-Follikel-stimulierendes Hormonlevel
FTC	Emtricitabin
G-BA	Gemeinsamer Bundesausschuss
GCP	Good clinical practice
HAART	Highly Active Antiretroviral Therapy
HCV	Hepatitis C-Virus
HIV-1	Humanes Immundefizienz-Virus (Typ 1) (human immunodeficiency virus (type 1))

HIVAN	HIV-assozierte Nephropathie
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 trial registry platform
IDMC	Independent Data-Monitoring Committee
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
IVRS	Interactive Voice Response System
IWRS	Interactive web response system
k.A.	keine Angaben
KI	Konfidenzintervall
KS	Kaposi-Sarkom
M/S=F	Missing/Art-Switch=Failure
M=E	Missing=Excluded
M=F	Missing=Failure
MedDRA	Medical Dictionary for Regulatory Activities
mL	Milliliter
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-nukleosidale Reverse Transkriptase Inhibitor
NRTI	Nukleosidale Reverse Transkriptase Inhibitor
OR	Odds Ratio
POR	Peto Odds Ratio
PP	Per Protokoll
PT	Preferred term
RAS	Randomisiertes Analyse-Set



RCT	Randomized Controlled Trial
RD	Risikodifferenz
REM	Modell mit zufälligen Effekten (Random effects model)
RNA	Ribonukleinsäure (ribonucleic acid)
s.u.	Siehe unten
SAS	Safety Analyse-Set
SD	Standardabweichung
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
TDF	Tenofoviridisoproxilfumarat (300 mg Tenofoviridisoproxilfumarat entsprechen 245 mg Tenofoviridisoproxil)
TLOVR	Time to loss of virological 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
ZNS	Zentrales Nervensystem

#### 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 Substanzkombination aus Efavirenz, Emtricitabin und Tenofoviridisoproxil (im Folgenden: EFV/FTC/TDF).

Die gemäß Indikation zugelassene Patientenpopulation umfasst zum einen nicht-vorbehandelte, erwachsene Patienten ( $\geq 18$  Jahre), und zum anderen erwachsene Patienten ( $\geq 18$  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.

Im vorliegenden Modul 4A wird die Patientenpopulation, definiert als erwachsene Patienten ( $\geq 18$  Jahre) mit einer HIV-1 Infektion, die antiretroviral nicht-vorbehandelt sind, untersucht.

Die Patientenpopulation der nicht-vorbehandelten Patienten ( $\geq 18$  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 ( $\geq 18$  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.

### Datenquellen

Die Bewertung des Zusatznutzens von EVG/COBI/FTC/TDF basiert 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. Es werden Daten zu Woche 48, 60 und 96 präsentiert. Zusätzlich zu den Informationen aus dem Studienbericht werden post-hoc durchgeführte Subgruppenanalysen dargestellt.

### Ein-/Ausschlusskriterien für Studien

Die Einschlusskriterien für die einzuschließenden Studien wurden wie folgt definiert:

- Population: Antiretroviral nicht-vorbehandelte erwachsene Patienten mit einer HIV-1 Infektion

- 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)
- 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])
- Studientyp: RCT
- Studiendauer:  $\geq 48$  Wochen
- Publikationstyp: Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht

Die Ausschlusskriterien wurden wie folgt festgelegt:

- 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
- 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.

### **Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen**

Alle hier eingeschlossenen Studien waren randomisierte kontrollierte Studien und entsprechen demnach der Evidenzstufe Ib.

Im Folgenden wird zusammenfassend 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 EFV/FTC/TDF

<b>Zusatznutzen durch EVG/COBI/FTC/TDF</b> <u>Format:</u> <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des ZN Behandlungseffekt (Zeitpunkt)	<b>Schaden durch EVG/COBI/FTC/TDF</b> <u>Format:</u> <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des Schadens Behandlungseffekt (Zeitpunkt)
<ul style="list-style-type: none"> <li>• <b>Veränderung der CD4-Zellzahl</b>  <u>Gesamt: Beleg für geringen ZN</u>            MWD: 36,26 [12,42;60,10], p=0,003 (Woche 48)</li> <li>• <b>Therapieabbrüche aufgrund UEs</b>  <u>Nicht-kaucasisch: Beleg für beträchtlichen ZN<sup>b</sup></u>            OR: 0,16 [0,04;0,64], p=0,010 (Woche 48)            OR: 0,24 [0,06;0,87], p=0,0303 (Woche 96)</li> <li>• <b>Allg. Erkrankungen und Beschwerden am Verabreichungsort SOC</b>  <u>Gesamt: Hinweis auf geringen ZN</u>            OR: 0,64 [0,46;0,89], p=0,0009 (Woche 48)  <u>≤350 Zellen/μL: Beleg für geringen ZN</u>            OR: 0,46 [0,27;0,75], p=0,002 (Woche 48)            OR: 0,49 [0,29;0,81], p=0,0052 (Woche 96)  <u>&gt;350 Zellen/μL: Hinweis auf geringen ZN<sup>c</sup></u>            OR: 0,83 [0,53;1,31], p=0,43 (Woche 48)            OR: 0,98 [0,62;1,53], p=0,9214 (Woche 96)</li> <li>• <b>Erkrankungen des Nervensystems SOC</b>  <u>Gesamt: Beleg für beträchtlichen ZN</u>            OR: 0,55 [0,41;0,75], p=0,0001 (Woche 48)            OR: 0,58 [0,42;0,78], p=0,0004 (Woche 96)  <b>Schwindel</b>  <u>Gesamt: Beleg für beträchtlichen ZN</u>            OR: 0,21 [0,13;0,34], p&lt;0,00001 (Woche 48)            OR: 0,22 [0,14;0,36], p&lt;0,00001 (Woche 96)  <b>Neurologische &amp; psychiatrische Ereignisse (Gruppe von PTs)</b>  <u>Gesamt: Beleg für beträchtlichen ZN</u>            OR: 0,46 [0,34;0,61], p&lt;0,00001 (Woche 48)            OR: 0,45 [0,33;0,61], p&lt;0,001 (Woche 96)</li> <li>• <b>Psychiatrische Erkrankungen SOC</b>  <u>Gesamt: Beleg für beträchtlichen ZN</u>            OR: 0,59 [0,44;0,79], p=0,0003 (Woche 48)</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Gesamtrate SUEs<sup>a</sup></b>  <u>Gesamt: Beleg für geringen Schaden</u>            OR: 1,77 [1,06;2,97], p=0,03 (Woche 48)            OR: 1,85 [1,17;2,93], p=0,0083 (Woche 96)  <u>Kaukasisch: Hinweis auf geringen Schaden</u>            OR: 2,49 [1,33;4,66], p=0,0043 (Woche 96)  <u>Nicht-kaucasisch: Anhaltspunkt für geringen Schaden<sup>b</sup></u>            OR: 1,25 [0,63;2,48], p=0,5268 (Woche 96)</li> <li>• <b>Renale Ereignisse</b>  <u>Gesamt: Hinweis auf geringen Schaden</u>            POR: 4,60 [1,14;18,54], p=0,037 (Woche 96)</li> <li>• <b>Erkrankungen des Gastrointestinaltrakts SOC</b>  <u>Gesamt: Hinweis auf geringen Schaden</u>            OR: 1,44 [1,08;1,91], p=0,01 (Woche 48)  <u>≤350 Zellen/μL: Beleg für geringen Schaden</u>            OR: 1,87 [1,20;2,93], p=0,006 (Woche 48)            OR: 1,74 [1,10;2,75], p=0,0171 (Woche 96)  <u>&gt;350 Zellen/μL: Hinweis auf geringen Schaden<sup>c</sup></u>            OR: 1,22 [0,83;1,77], p=0,31 (Woche 48)            OR: 1,11 [0,75;1,66], p=0,5966 (Woche 96)  <b>Übelkeit</b>  <u>Gesamt: Beleg für geringen Schaden</u>            OR: 1,61 [1,09;2,38], p=0,02 (Woche 48)            OR: 1,58 [1,07;2,32], p=0,0213 (Woche 96)  <u>&gt;100.000 Kopien/mL: Beleg für geringen Schaden</u>            OR: 2,46 [1,20;5,04], p=0,0135 (Woche 48)            OR: 2,48 [1,24;4,98], p=0,0104 (Woche 96)  <u>≤100.000 Kopien/mL: Hinweis auf geringen Schaden<sup>c</sup></u>            OR: 1,36 [0,84;2,21], p=0,2172 (Woche 48)            OR: 1,26 [0,79;2,02], p=0,3321 (Woche 96)</li> </ul>

<b>Zusatznutzen durch EVG/COBI/FTC/TDF</b> <u>Format:</u> <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des ZN Behandlungseffekt (Zeitpunkt)	<b>Schaden durch EVG/COBI/FTC/TDF</b> <u>Format:</u> <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des Schadens Behandlungseffekt (Zeitpunkt)
<p>OR: 0,64 [0,47;0,86], p=0,0030 (Woche 96)</p> <p><b><i>Abnorme Träume</i></b></p> <p><u>Gesamt: Beleg für beträchtlichen ZN</u></p> <p>OR: 0,41 [0,22;0,78], p=0,006 (Woche 48)</p> <p>OR: 0,22 [0,06;0,77], p=0,0180 (Woche 60)</p> <p>OR: 0,47 [0,32;0,69], p=0,0001 (Woche 96)</p> <p><u>Kaukasisch: Beleg für beträchtlichen ZN</u></p> <p>OR: 0,26 [0,08;0,87], p=0,03 (Woche 48)</p> <p>OR: 0,10 [0,02;0,56], p=0,0089 (Woche 60)</p> <p>OR: 0,38 [0,24;0,62], p=0,0001 (Woche 96)</p> <p><u>Nicht-kaukasisch: Hinweis auf beträchtlichen ZN<sup>c</sup></u></p> <p>OR: 0,73 [0,40;1,33], p=0,30 (Woche 48)</p> <p>OR: 1,00 [0,08;12,56], p=1,0000 (Woche 60)</p> <p>OR: 0,69 [0,37;1,26], p=0,2270 (Woche 96)</p> <p><b><i>Schlaflosigkeit</i></b></p> <p><u>Gesamt: Beleg für geringen ZN</u></p> <p>OR: 0,62 [0,39;0,98], p=0,04 (Woche 48)</p> <p>OR: 0,63 [0,41;0,99], p=0,0428 (Woche 96)</p> <p>• <b>Erkrankungen der Haut und des Unterhautzellgewebes SOC</b></p> <p><u>Kaukasisch: Beleg für beträchtlichen ZN</u></p> <p>OR: 0,47 [0,32;0,69], p=0,0001 (Woche 48)</p> <p>OR: 0,53 [0,36;0,78], p=0,0013 (Woche 96)</p> <p><u>Nicht-kaukasisch: Hinweis auf beträchtlichen ZN<sup>c</sup></u></p> <p>OR: 0,94 [0,55;1,60], p=0,83 (Woche 48)</p> <p>OR: 0,96 [0,58;1,61], p=0,8896 (Woche 96)</p> <p><b><i>Hautausschlag</i></b></p> <p><u>Gesamt: Beleg für geringen ZN</u></p> <p>OR: 0,49 [0,30;0,81], p=0,006 (Woche 48)</p> <p>OR: 0,50 [0,30;0,82], p=0,0066 (Woche 96)</p> <p><u>&lt;40 Jahre: Hinweis auf beträchtlichen ZN</u></p> <p>OR: 0,65 [0,35;1,24], p=0,19 (Woche 48)</p> <p><u>≥40 Jahre: Anhaltspunkt für beträchtlichen ZN<sup>c</sup></u></p> <p>OR: 0,31 [0,13;0,72], p=0,006 (Woche 48)</p> <p><b>Hautausschläge (Gruppe von PTs)</b></p> <p><u>Kaukasisch: Beleg für beträchtlichen ZN</u></p> <p>OR: 0,42 [0,28;0,65], p&lt;0,0001 (Woche 48)</p>	

<b>Zusatznutzen durch EVG/COBI/FTC/TDF</b> <u>Format:</u> <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des ZN Behandlungseffekt (Zeitpunkt)	<b>Schaden durch EVG/COBI/FTC/TDF</b> <u>Format:</u> <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des Schadens Behandlungseffekt (Zeitpunkt)
OR: 0,51 [0,33;0,77], p=0,0016 (Woche 96) <u>Nicht-kaukasisch: Hinweis auf beträchtlichen ZN<sup>c</sup></u> OR: 0,97 [0,52;1,84], p=0,94 (Woche 48) OR: 0,92 [0,50;1,67], p=0,7740 (Woche 96)	
ZN = Zusatznutzen <sup>a</sup> Rein formal wird hier ein Beleg für einen geringen Schaden in der Gesamtpopulation, bzw. ein Hinweis auf einen geringen Schaden in der Gruppe der kaukasischen Patienten und ein Anhaltspunkt auf einen geringen Schaden in der Gruppe der nicht-kaukasischen Patienten abgeleitet - medizinisch kann durch den fehlenden Zusammenhang zwischen Intervention und Nebenwirkung jedoch kein Schaden abgeleitet werden (Details siehe Gesamtfazit zum Zusatznutzen in Abschnitt 4.4.2). <sup>b</sup> Der in der nicht-kaukasischen Subgruppe gefundene Unterschied bei den Therapieabbrüchen aufgrund von UEs 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. <sup>c</sup> 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 Nutzenaussage für die Subgruppe mit dem nicht signifikanten Effekt.	

### Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Zusammenfassend wurde in der vorliegenden Nutzenbewertung anhand von patientenrelevanten und validen Endpunkten gezeigt, dass EVG/COBI/FTC/TDF, bei gleichwertiger Wirksamkeit wie EFV/FTC/TDF (Nichtunterlegenheit zur 12%-Grenze wurde gezeigt), eine vom Nebenwirkungsprofil her verbesserte Behandlungsoption darstellt als die zweckmäßige Vergleichstherapie EFV/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 EFV/FTC/TDF besitzt. Dies zeigt sich anhand der Endpunkte virologisches Ansprechen, Zeit bis zum Verlust des virologischen Ansprechens und Veränderung der CD4-Zellzahl. Hinsichtlich der Veränderung der CD4-Zellzahl zeigte sich zu Woche 48 sogar ein Vorteil für EVG/COBI/FTC/TDF, was insbesondere zu Beginn einer antiretroviralen Therapie hoch patientenrelevant ist, weil durch eine raschere Dynamik der Entwicklung der CD4-Zellzahl das Immunsystem schneller rekonstituiert und so der für die Entwicklung Aids-definierender Erkrankungen relevante, niedrige CD4-Zellbereich rascher verlassen wird. Aus diesem Grund wird für diesen Endpunkt ein Beleg für einen geringen Zusatznutzen abgeleitet.



Bei sicherheitsbezogenen Endpunkten zeigten sich unter EVG/COBI/FTC/TDF deutliche patientenrelevante Vorteile hinsichtlich der Vermeidung von Therapieabbrüchen aufgrund unerwünschter Ereignisse, der Vermeidung neurologischer & psychiatrischer Ereignisse, von Hautausschlägen, allgemeiner Erkrankungen und Beschwerden am Verabreichungsort (darunter Fatigue), Erkrankungen des Nervensystems (darunter Schwindel), psychiatrischer Erkrankungen (darunter abnorme Träume und Schlaflosigkeit) und Erkrankungen der Haut und des Unterhautzellgewebes (darunter Hautausschlag). Bei der Ableitung eines Gesamtfazits zum Zusatznutzen werden die Gruppen von PT (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 Gesamtzusatznutzens bei.

Erkrankungen der Haut und des Unterhautzellgewebes werden von den Betroffenen nicht nur als sehr unangenehm empfunden, sondern stellen auch einen 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 Hautausschlag im Kreis der HIV-infizierten Bevölkerung somit 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 [1, 2].

Auch weitere direkt durch die Behandlung verursachte Erkrankungen des Nervensystems und psychiatrische Erkrankungen können die Adhärenz mit den beschriebenen klinischen Folgen negativ beeinflussen. Zudem stellen sie für Patienten äußerst unangenehme Belastungen dar, die sie in ihrem Alltag und Berufsleben stark beeinträchtigen und zu einer stark reduzierten Lebensqualität führen. Zu diesen zählen Nebenwirkungen wie Schwindel, Schlaflosigkeit, abnorme Träume und Alpträume. Die hohe und persistierende Inzidenz von Schwindel stellt ein besonderes Risiko bei Personen dar, die beruflich Maschinen bedienen, Kraftfahrzeuge lenken und/oder hochkonzentriert arbeiten müssen. Die hohe Patientenrelevanz des zentralnervösen Nebenwirkungsspektrums zeigen Daten aus verschiedenen Kohorten (Real-Life-Setting): In einer englischen Kohorte mussten etwa 20% aller Patienten, die eine Therapie mit EFV/FTC/TDF beginnen, innerhalb der ersten 42 Therapiewochen (Median) umgestellt werden – und zwar primär aufgrund der ZNS-Toxizität von Efavirenz [3]. In einer skandinavischen Kohorte zeigte sich, dass etwa 50% der Patienten unter einer EFV-haltigen Initialtherapie innerhalb des ersten Behandlungsjahres umgestellt werden müssen [4]. Zudem besteht das Risiko, dass die Patienten ihre Therapie aufgrund der subjektiv belastenden Nebenwirkungen unkontrolliert abbrechen, was zu Resistenzentwicklung, Verlust an

Therapieoptionen und Krankheitsprogression führen kann. Auch der Großteil der Therapieabbrüche in den hier vorliegenden Studien ist psychiatrischer Natur.

Die SOC allgemeine Erkrankungen und Beschwerden am Verabreichungsort beschreibt eine eher heterogene Gruppe von Ereignissen; sie umfasst einerseits allgemeine Erkrankungen wie Fatigue, Fieber und Schmerzen, und berücksichtigt andererseits auch Beschwerden am Verabreichungsort. Während die Beschwerden am Verabreichungsort bei einer oral einzunehmenden Therapie von untergeordneter Bedeutung sind, umfassen die allgemeinen Erkrankungen eine ganze Reihe von Ereignissen mit hoher Patientenrelevanz. Die unter dieser SOC beobachteten Ereignisse in Studie S102 zu Woche 48 umfassen Ereignisse wie unter anderem Fatigue, Pyrexie und Brustschmerzen, die alle direkten Einfluss auf das Wohlbefinden des Patienten haben und somit patientenrelevant sind.

Durch die bessere Verträglichkeit sinkt 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 antiretrovirale Therapie (ART) wechseln, ein erstrebenswertes Ziel [5, 6]. 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 ebenfalls zu bedenken ist, dass eine ART durch die Übertragung resistenter Viren von Anfang an komplizierter und im Zweifel auch teurer wird [7]. 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 [8].

Insgesamt zeigen verschiedene Daten, dass beeinträchtigende Nebenwirkungen generell einen negativen Einfluss auf die Therapieadhärenz und damit Persistenz haben, was das Risiko der Bildung von Resistenzen und virologischem Versagen mit 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 [7, 9, 10].

Neben den positiven sicherheitsbezogenen Aspekten wurden auch einige Nachteile von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF beobachtet. Dies zeigte sich bei der Gesamtrate schwerwiegender unerwünschter Ereignisse, renaler Ereignisse und Erkrankungen des Gastrointestinaltrakts (darunter Übelkeit).

Die Detail-Analyse zeigt, dass der gefundene Unterschied bei den beobachteten schwerwiegenden unerwünschten Ereignissen primär getrieben wurde durch einen Unterschied zwischen beiden Behandlungsarmen in der SOC Infektionen und parasitäre

Erkrankungen in Studie S102 – dieser Unterschied zeigte sich weder in Studie S104 noch in Studie S103 [11]. Dies ist insofern überraschend, als dass ein erhöhter Anteil an Infektionen unter einer bestimmten Behandlung im Gegensatz zur Vergleichstherapie im HIV-Bereich zumeist ein unter dieser Behandlung nicht ausreichend gestärktes Immunsystem nahelegen würde (siehe Modul 3) - die Erhöhung der CD4-Zellzahlen unter EVG/COBI/FTC/TDF war jedoch vergleichbar bzw. zu Woche 48 sogar signifikant stärker als unter EFV/FTC/TDF. Damit kann ein Versagen der Behandlung ebenso wie ein durch die HIV-1-Infektion geschwächtes Immunsystem als Grund ausgeschlossen werden. Darüber hinaus fällt auf, dass der Unterschied zwischen den Armen insbesondere auf die Krankheitsbilder Cellulitis und Appendicitis zurückzuführen ist: Hier zeigte sich in Studie S102 ein deutlich größerer Anteil an EVG/COBI/FTC/TDF-Patienten als EFV/FTC/TDF-Patienten. Bei der Cellulitis handelt es sich um eine häufige, bakterielle Infektion von Dermis und Unterhautgewebe, primär verursacht durch Streptokokken und Staphylokokkus Aureus. Die Risikofaktoren für eine Cellulitis sind klar umschrieben und umfassen neben Hautläsionen (Risse, Schnitte oder Bisse) und kutanen Pilzinfektionen vor allem Diabetes mellitus, ein geschwächtes Immunsystem (beispielsweise durch HIV oder nach Chemotherapie), Lymphödeme und intravenösen Drogengebrauch. Laut der individuellen Fallberichte („patient narratives“) der EVG/COBI/FTC/TDF-Patienten mit schwerwiegender Cellulitis wies jeder dieser Patienten Risikofaktoren für Cellulitis auf; entsprechend wurden alle aufgetretenen Fälle vom jeweiligen Prüfarzt als nicht in Zusammenhang mit der Studienmedikation stehend beurteilt [12-17]. Ein Zusammenhang mit EVG/COBI/FTC/TDF ist für die Cellulitis damit ebenso wenig nachvollziehbar wie für die Appendicitis; auch die Appendicitis-Fälle wurden vom jeweiligen Prüfarzt als nicht assoziiert mit der Studienmedikation bewertet [18-20]. Zudem wurden keine Untersuchungen zu Unbalanziertheiten bezüglich solcher Risikofaktoren durchgeführt und somit kann eine Unbalanziertheit der Risikofaktoren zuungunsten von EVG/COBI/FTC/TDF nicht ausgeschlossen werden. Des Weiteren ist hier anzumerken, dass in keiner anderen EVG/COBI/FTC/TDF-Studie (S103 und S104) ein Unterschied zuungunsten EVG/COBI/FTC/TDF gefunden wurde. Ein Zusammenhang zwischen EVG/COBI/FTC/TDF und schweren Infektionen ist medizinisch nicht plausibel und somit kann das ungewöhnlich häufige Auftreten dieser Infektionen in Studie S102 hier nicht in einen kausalen Zusammenhang mit der Medikation gebracht werden. Korrigiert man um die Cellulitis- und Appendicitis-Fälle, geht der statistisch signifikante Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF verloren.

Bei den renalen Ereignissen zeigt sich ein statistisch signifikanter Unterschied zwischen den beiden Armen zuungunsten von EVG/COBI/FTC/TDF; renale Ereignisse waren dabei selten und traten in Studie S102 bei 5 Patienten unter EVG/COBI/FTC/TDF und 1 Patienten unter EFV/FTC/TDF auf. Unter denselben zu vergleichenden Therapien trat in Studie S104 bei keinem Patienten ein renales Ereignis auf. Potentielle renale Nebenwirkungen im Sinne einer Schädigung des proximalen Tubulus unter Tenofovir sind seit langem bekannt, und werden durch den „Risk Management Plan“ durch die Europäische Zulassungsbehörde adäquat überwacht. Der hier vorliegende statistisch signifikante Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF hinsichtlich renaler Ereignisse ist insofern überraschend, als dass Tenofovir als potentiell verursachendes Agens in gleicher Dosierung in

beiden Armen enthalten ist – entsprechend findet sich in Studie S104 ebenso wie im Vergleich zwischen EVG/COBI/FTC/TDF und ATV/r+FTC/TDF in Studie S103 kein Unterschied hinsichtlich des Auftretens renaler Ereignisse [11].

Erkrankungen des Gastrointestinaltrakts gehören zu den häufigsten Nebenwirkungen; sie können insbesondere in der Anfangsphase bei fast allen antiretroviralen Arzneimitteln auftreten, wobei NNRTI ein günstiges gastrointestinales Profil aufweisen, geboosterte Protease-Inhibitoren hingegen häufig mit gastrointestinalen Nebenwirkungen assoziiert sind [21]. Hill et al. zeigen, dass es einen komplexen, kombinierten Effekt aus HIV-Infektion und ART in Bezug auf die Inzidenz gastrointestinaler Symptome gibt – was bedeutet, dass die Mehrheit der gastrointestinalen Nebenwirkungen in einigen Studien nicht in Zusammenhang mit der ART stehen könnten. Entsprechend gibt es klinische Daten, die zeigen, dass sich die gastrointestinalen Nebenwirkungen durch die Therapie einer fortgeschrittenen HIV-Infektion als Ergebnis der immunologischen Rekonstitution verbessern [22]. Dies steht im Einklang mit den gefundenen Ergebnissen: Während der Unterschied in den kränkeren Subgruppen mit Ausgangsviruslast  $>100.000$  Kopien/mL, resp. CD4-Zellzahl  $\leq 350$  CD4-Zellen/ $\mu$ L signifikant zuungunsten von EVG/COBI/FTC/TDF ist, zeigt sich bei den Patienten mit Ausgangsviruslast  $\leq 100.000$  Kopien/mL, bzw. CD4-Zellzahl  $>350$  CD4-Zellen/ $\mu$ L kein Unterschied zwischen den Behandlungen. Möglicherweise führt der zusätzliche gastrointestinale Einfluss der HIV-Infektion zu einem additiven Effekt von Infektion und Medikation, der bei den kränkeren Patienten unter EVG/COBI/FTC/TDF ausgeprägter ist als unter dem gastrointestinal gut verträglichen EFV-basierten Regime. Verbessert sich also der Immunstatus durch die Therapie, ist davon auszugehen, dass der Unterschied wieder verschwindet. Die Transienz der gastrointestinalen Nebenwirkungen bestätigt auch eine kürzlich veröffentlichte Analyse der Studien S102 und S103, in der gezeigt wird, dass die Inzidenz gastrointestinaler Nebenwirkungen unter EVG/COBI/FTC/TDF zu Beginn einer Therapie am höchsten ist und über die Zeit deutlich abnimmt [23]. Insbesondere Diarrhoe lässt sich auch nicht-medikamentös sehr gut behandeln, desgleichen Übelkeit - insbesondere wenn diese, wie im Fall von EVG/COBI/FTC/TDF, nicht in Verbindung mit Erbrechen steht. Die überschaubare klinische Relevanz dieser Nebenwirkung wird zudem deutlich in der Tatsache, dass lediglich ein Patient aufgrund von gastrointestinalen Nebenwirkungen (hier Übelkeit) abgebrochen hat. Des Weiteren wurde der Großteil der gastrointestinalen Nebenwirkungen in Studie S102 zu Woche 48 als Grad 1-2 eingestuft.

Die Studiendaten zeigen außerdem deutliche Unterschiede in der Patientenrelevanz von neurologischen & psychiatrischen Nebenwirkungen einerseits und gastrointestinalen Ereignissen andererseits: Während in der gepoolten Analyse der Studien S102 und S104 lediglich ein Patient die Studie aufgrund von gastrointestinalen Nebenwirkungen abbrach, führten in Studie S102 neurologische (bis Woche 96: 4 Patienten) und psychiatrische (bis Woche 96: 13 Patienten) Ereignisse (mittels SOC) zu deutlich mehr Studien-Abbrüchen.

### **Ableitung des Gesamt-Zusatznutzens**

Zur Abwägung des Gesamt-Zusatznutzens lässt sich ableiten, dass der Hinweis für einen geringen Schaden durch EVG/COBI/FTC/TDF bei Erkrankungen des Gastrointestinaltrakts

durch den Hinweis auf einen geringen Zusatznutzen hinsichtlich allgemeiner Erkrankungen und Beschwerden am Verabreichungsort ausgeglichen wird. Ebenfalls ausgeglichen ist die Nutzen-Schaden-Abwägung in der Subgruppe der Patienten mit CD4-Zellzahl zu Baseline  $\leq 350$  Zellen/ $\mu\text{L}$ , in der einem Beleg für einen geringen Schaden hinsichtlich Erkrankungen des Gastrointestinaltrakts unter EVG/COBI/FTC/TDF ein Beleg für einen geringen Zusatznutzen hinsichtlich allgemeiner Erkrankungen unter EVG/COBI/FTC/TDF gegenübersteht.

Die tatsächliche Patientenrelevanz des Schadens hinsichtlich schwerwiegender unerwünschter Ereignisse wird als gering eingeschätzt, da er hauptsächlich auf die unter EVG/COBI/FTC/TDF ungewöhnlich häufig aufgetretenen schweren Infektionen in Studie S102 (vor allem Cellulitis und Appendicitis) zurückzuführen ist, die medizinisch nicht in Zusammenhang mit dem zu bewertenden Arzneimittel gebracht werden können, und entsprechend auch vom Prüfarzt als nicht in Zusammenhang mit der Studienmedikation stehend beurteilt wurden. Medizinisch plausibel kann für schwerwiegende unerwünschte Ereignisse also kein Schaden abgeleitet werden.

Damit bleiben die Belege für einen beträchtlichen Zusatznutzen von EVG/COBI/FTC/TDF sowohl bei Hautausschlag als auch bei den Erkrankungen des Nervensystems (SOC), darunter Schwindel, und bei neurologischen & psychiatrischen Ereignissen (Gruppe von PT). Des Weiteren zeigt sich ein Beleg für einen beträchtlichen Zusatznutzen von EVG/COBI/FTC/TDF bei den psychiatrischen Ereignissen, darunter ein Beleg für einen beträchtlichen Zusatznutzen hinsichtlich der Vermeidung abnormer Träume, sowie ein Beleg für einen geringen Zusatznutzen von EVG/COBI/FTC/TDF im Hinblick auf Schlaflosigkeit.

Es zeigt sich außerdem ein Beleg für einen geringen Zusatznutzen durch EVG/COBI/FTC/TDF hinsichtlich der Veränderung der CD4-Zellzahl. Da durch eine raschere Dynamik der Entwicklung der CD4-Zellzahl das Immunsystem schneller rekonstituiert, und so der für die Entwicklung Aids-definierender Erkrankungen relevante, niedrige CD4-Zellbereich rascher verlassen wird, ist dieser Vorteil von besonderer Patientenrelevanz. Diesem Zusatznutzen steht ein Hinweis auf einen geringen Schaden hinsichtlich renaler Ereignisse gegenüber. Diese sind insgesamt selten und unter Tenofovir seit langem bekannt; ein entsprechender Risk-Management-Plan wurde implementiert.

Insgesamt sind somit unter einer Behandlung mit EVG/COBI/FTC/TDF deutlich weniger Nebenwirkungen zu beobachten als unter einer Behandlung mit EFV/FTC/TDF – dies betrifft insbesondere die Nebenwirkungen, die für die Patienten besonders relevant sind: Vor allem dermatologische, psychiatrische und neurologische Nebenwirkungen führen zu einer erheblichen Beeinträchtigung des Patienten sowohl hinsichtlich seiner Funktionsfähigkeit als auch hinsichtlich seiner Lebensqualität. Derart beeinträchtigende Nebenwirkungen können die Therapieadhärenz negativ beeinflussen und somit das Risiko eines Therapieversagens mit Resistenzentwicklung und konsekutiver Krankheitsprogression erhöhen. Virologisches Versagen aufgrund von Resistenzentwicklung und nicht tolerierbarer Nebenwirkungen schränken in einer Indikation, die eine lebenslange tägliche Einnahme von Medikamenten erfordert, verfügbare Therapieoptionen ein.

In der Gesamtschau des auf Endpunktebene abgeleiteten Zusatznutzens verbleibt ein positiver Zusatznutzen für EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF, dessen Ausmaß als **gering** eingestuft wird. Da hier zwei randomisierte, aktiv-kontrollierte Studien zugrunde liegen, wird dieser Zusatznutzen mit der Wahrscheinlichkeit eines **Belegs** eingestuft.

Die von Gilead Sciences vorgeschlagene zweckmäßige Vergleichstherapie umfasste zusätzlich zu der vom G-BA festgelegten und hier untersuchten zweckmäßigen Vergleichstherapie EFV/FTC/TDF zwei weitere Regime, i.e. ATV/r+FTC/TDF und RAL+FTC/TDF (siehe Modul 3), die ebenfalls vollumfänglich untersucht wurden:

In der Gesamtschau des auf Endpunktebene abgeleiteten Zusatznutzens verbleibt ein positiver Zusatznutzen für EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF, der auf Basis einer randomisierten, aktiv-kontrollierten Studie mit der Wahrscheinlichkeit eines Hinweises eingestuft wird [11]. Im Vergleich zur zweckmäßigen Vergleichstherapie RAL+FTC/TDF wird für EVG/COBI/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 [24].

Damit ergibt sich insgesamt ein **Beleg** für einen **geringen** Zusatznutzen für EVG/COBI/FTC/TDF.

## 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 zur Substanzkombination aus Efavirenz, Emtricitabin und Tenofoviridisoproxil (im Folgenden: EFV/FTC/TDF).

#### **Patientenpopulation**

Die gemäß Indikation zugelassene Patientenpopulation umfasst zum einen nicht-vorbehandelte, erwachsene Patienten ( $\geq 18$  Jahre), und zum anderen vorbehandelte, erwachsene Patienten ( $\geq 18$  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 hier relevante und untersuchte Patientenpopulation ist definiert als erwachsene Patienten ( $\geq 18$  Jahre) mit einer HIV-1 Infektion, die antiretroviral nicht-vorbehandelt sind.

Die Patientenpopulation der nicht-vorbehandelten Patienten ( $\geq 18$  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 ( $\geq 18$  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**

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 Elvitegravir, 150 mg des Boosters Cobicistat, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil (nukleosidale Reverse-Transkriptase-Inhibitoren (NRTI)).

### **Vergleichstherapie**

Als zweckmäßige Vergleichstherapie wurde vom Gemeinsamen Bundesausschuss (G-BA) festgelegt (G-BA-Beratungsgespräch vom 5. Februar 2013):

- Efavirenz in Kombination mit Tenofoviridisoproxil/Emtricitabin oder Abacavir/Lamivudin

Die zweckmäßige Vergleichstherapie im vorliegenden Modul 4A ist Efavirenz (EFV) in Kombination mit Tenofoviridisoproxil (TDF) und Emtricitabin (FTC) in Form des Single-Tablet-Regimens EFV/FTC/TDF. EFV/FTC/TDF wird einmal täglich auf leeren Magen vor dem Zubettgehen oral eingenommen; eine Filmtablette enthält 600 mg des nicht-nukleosidalen Reverse-Transkriptase-Inhibitors (NNRTI) Efavirenz, 200 mg Emtricitabin und 245 mg Tenofoviridisoproxil.

Die Substanzen 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 gewählten Formulierung somit nicht der hier 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 European Medicines Agency (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<sup>®</sup> 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 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: „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 zum Beispiel des *Single-Tablet-Regimen*“. [25].

Darüber hinaus findet sich in der Nutzenbewertung des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) zu Pixantron [26] 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.“

### 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/ $\mu$ L oder Zellen/ $\mu$ L 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)

### 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.

Da es möglich ist, die zu bewertende Intervention in der vorliegenden Fragestellung im Rahmen von RCT zu untersuchen, 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

	<b>Einschlusskriterien</b>	<b>E</b>	<b>Ausschlusskriterien</b>	<b>A</b>
Population	Antiretroviral nicht-vorbehandelte erwachsene Patienten mit einer HIV-1 Infektion	1	<ul style="list-style-type: none"> <li>• Patienten mit anderen Infektionen</li> <li>• vorbehandelte Patienten</li> <li>• Kinder und Jugendliche</li> </ul>	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)	3	Andere Vergleichstherapie oder abweichende Dosierungen	3
Endpunkte	<p>Mindestens einer der folgenden patientenrelevanten Endpunkte wird berichtet:</p> <ul style="list-style-type: none"> <li>• Virologisches Ansprechen (Absenkung der Viruslast unter die Nachweisgrenze)</li> <li>• Zeit bis zum Verlust des virologischen Ansprechens</li> <li>• Veränderung der CD4-Zellzahl</li> <li>• Gesamtrate unerwünschter Ereignisse</li> <li>• Schwerwiegende unerwünschte Ereignisse</li> <li>• Therapieabbrüche aufgrund unerwünschter Ereignisse</li> <li>• Unerwünschte Ereignisse Grad 3-4</li> <li>• Wichtige unerwünschte Ereignisse <ul style="list-style-type: none"> <li>○ Hautausschläge</li> <li>○ Neurologische &amp; psychiatrische Ereignisse</li> </ul> </li> <li>• Unerwünschte Ereignisse von Interesse <ul style="list-style-type: none"> <li>○ Renale Ereignisse</li> <li>○ Knochenfrakturen</li> </ul> </li> <li>• Häufigste unerwünschte Ereignisse <ul style="list-style-type: none"> <li>○ Erkrankungen des Gastrointestinaltrakts (SOC)</li> <li>○ Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)</li> <li>○ Erkrankungen des Nervensystems (SOC)</li> <li>○ Psychiatrische Erkrankungen (SOC)</li> <li>○ Erkrankungen der Haut und des Unterhautzellgewebes (SOC)</li> </ul> </li> </ul>	4	Keiner der unter den Einschlusskriterien definierten Endpunkte wird berichtet	4

	<b>Einschlusskriterien</b>	<b>E</b>	<b>Ausschlusskriterien</b>	<b>A</b>
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 [27]. Hier wird die Patientenpopulation, definiert als erwachsene Patienten (≥18 Jahre) mit einer HIV-1 Infektion, die antiretroviral nicht-vorbehandelt sind, untersucht.

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.

### ***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***

Das Kriterium der Vergleichstherapie ergibt sich gemäß Festlegung des G-BA. Dieser Vorgabe wird hier gefolgt. Da diese zweckmäßige Vergleichstherapie allein jedoch nicht den allgemein anerkannten Stand der medizinischen Erkenntnis darstellt, werden zusätzlich Daten zu allen weiteren relevanten zweckmäßigen Vergleichstherapien präsentiert [11, 24]. Der allgemein anerkannte Stand der medizinischen Erkenntnis besteht aus der Substanzklassen-Kombination von 2 NRTI mit einem 3. Partner aus der Klasse der NNRTI oder der Protease-Inhibitoren oder der Integrase-Inhibitoren - die verschiedenen Substanzkombinationen werden weder priorisiert noch in eine Rangfolge gesetzt; sie werden als gleichberechtigt bewertet. Die Auswahl einer einzigen zweckmäßigen Vergleichstherapie entspricht daher weder der

klinischen Evidenz noch der Leitlinien-Empfehlung noch der klinischen Praxis und kann demzufolge auch nicht als allgemein anerkannter Stand der medizinischen Erkenntnis angesehen werden.

Die zweckmäßige Vergleichstherapie ergibt sich dabei entsprechend der klinischen Evidenz gemäß der in den Leitlinien empfohlenen und in der klinischen Praxis angewandten Therapieregime zur Behandlung einer HIV-1 Infektion. Die verschiedenen empfohlenen Therapieoptionen stellen im Hinblick auf den einzelnen Patienten keine austauschbaren Alternativen zueinander dar. Die Therapieentscheidung für eine bestimmte Kombination der empfohlenen Substanzen basiert auf virus- und substanzspezifischen sowie anamnestischen Gesichtspunkten. Die Auswahl einer einzigen zweckmäßigen Vergleichstherapie würde entsprechend nur einen Teil der für EVG/COBI/FTC/TDF in Frage kommenden Patienten abbilden und nicht die komplette Population. Aus diesem Grund sollte eine neue Therapiemöglichkeit mit allen in dieser Indikation etablierten und gleichwertig empfohlenen Therapieregimen verglichen werden.

Es werden somit, um den Vergleich der Intervention mit dem allgemein anerkannten Stand der medizinischen Erkenntnis adäquat abzubilden, im vorliegenden Dossier neben den Daten zu der vom G-BA festgelegten zweckmäßigen Vergleichstherapie EFV/FTC/TDF, auch Daten zu allen weiteren relevanten zweckmäßigen Vergleichstherapien für alle indizierten Patientenpopulationen (siehe auch [11, 24] sowie Modul 4B) berichtet.

### ***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. Da die unter Abschnitt 4.2.1 genannten patientenrelevanten Endpunkte in der vorliegenden Fragestellung mittels RCT untersucht werden können, wird nur nach RCT gesucht.

### ***Studiendauer***

Das Kriterium der Mindeststudiendauer ergibt sich aus der Leitlinie der EMA [28], 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.

#### 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](http://clinicaltrials.gov), [clinicalstudyresults.org](http://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](http://clinicaltrials.gov), [clinicalstudyresults.org](http://clinicalstudyresults.org) sowie in dem International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO) durchgeführt. Da die Seite [clinicalstudyresults.org](http://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 der zweckmäßigen Vergleichstherapie 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“ Verzerrungspotenzial wird ausgegangen, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Eine relevante Verzerrung liegt vor, falls sich die Aussagen und Schlussfolgerungen bezüglich der Ergebnisse bei Behebung des verzerrenden Aspektes verändern würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führt nicht zum Ausschluss aus der Bewertung. Die Klassifizierung dient vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

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

Eine zusammenfassende Bewertung der Verzerrungsaspekte erfolgt ausschließlich für RCT.

#### **4.2.5 Informationssynthese und -analyse**

##### **4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien**

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)<sup>1</sup>. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-<sup>2</sup> bzw. STROBE-Statements<sup>3</sup> 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 wurden die in die Bewertung einfließenden Studien anhand von Design-Charakteristika (Studiendesign, Studiendauer, Zahl der randomisierten Patienten, Ort

---

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

<sup>2</sup> 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.

<sup>3</sup> 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.

und Zeitraum der Studiendurchführung, Interventionen, primäres Zielkriterium und relevante sekundäre Zielkriterien) beschrieben. Die Beschreibung der Design-Charakteristika deckte die im CONSORT-Statement angegebenen Items 2b bis 14 ab. Darüber hinaus wurde der Patientenfluss in den Einzelstudien mithilfe der Informationen aus dem CONSORT-Flow-Chart dargestellt.

#### 4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

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

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)
- Rasse und
- Ethnische Abstammung.

Als krankheitsbeschreibende Faktoren werden

- Ausgangsviruslast (HIV-1-RNA Kopien) zu Baseline (stetig und 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 & 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)

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 [29]: 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 [30].
- 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“ [31]. Die EMA empfiehlt entsprechend die Erfassung des Erreichens und Aufrechterhaltens des virologischen Ansprechens im Rahmen klinischer Studien [28]:

„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 [32]. 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 [33]. 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 [34]. Darüber hinaus steigt das Risiko eines virologischen Rebounds mit der Komplexität des Regimes [8].

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 [35].

Ein langfristiges, dauerhaftes virologisches Ansprechen führt zu einer fortschreitenden immunologischen Rekonstitution, wie die über Jahre ansteigenden CD4-Zellzahlen belegen [36].

Des Weiteren belegen verschiedene Studien 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 [8, 37, 38]. Zudem ist das Risiko eines virologischen Rebounds bei nicht-vorbehandelten Patienten niedriger als bei behandelten Patienten, die ihre Therapie wechseln [8].

Dauer des virologischen Ansprechens und Persistenz sind somit kritische Parameter für die Reduktion von Morbidität und Mortalität, dies bedingt durch das assoziierte immunologische Ansprechen [39], 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 [40]:

„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 [28]:

„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 Centers for Disease Control and Prevention (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. [41] 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/ $\mu$ 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. [42] 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 1: 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.

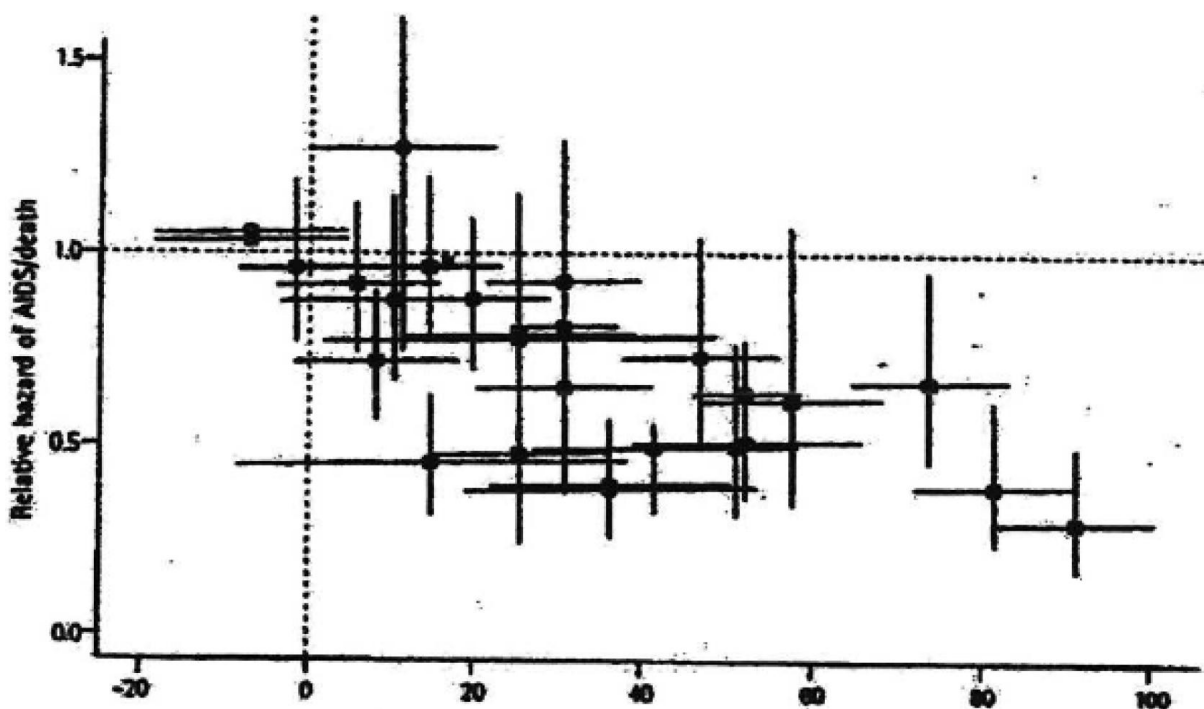


Abbildung 1: Relation zwischen Veränderung der CD4-Zellzahl und der Hazard für klinische Ereignisse (Aids/Tod) [42]



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 gemäß Leitlinien spätestens bei einer CD4-Zellzahl von 350 CD4-Zellen/ $\mu\text{L}$  begonnen werden [31]; ab einer CD4-Zellzahl von  $<200$  CD4-Zellen/ $\mu\text{L}$  muss mit zunehmender Dauer der Immunsuppression vermehrt mit dem Auftreten von Aids-definierenden Erkrankungen gerechnet werden [43]. 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 & psychiatrische Ereignisse: Diese werden in Zusammenhang mit einer Behandlung mit EFV häufig beobachtet und als übliche Nebenwirkung dieser Behandlung in der Fachinformation beschrieben [44, 45]. Sie wurden deswegen in der vorliegenden Nutzenbewertung als relevant eingestuft.
  - 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.
- Unerwünschte Ereignisse von Interesse:
  - Renale Ereignisse: Aufgrund des Zusammenhangs zwischen renalen Ereignissen und der Behandlung mit TDF, welches in den in die Nutzenbewertung eingeschlossenen Studien sowohl im Elvitegravir- als auch im Efavirenz-Arm 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 sowohl im Elvitegravir- als auch im Efavirenz-Arm gegeben wird, 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 und
- Erkrankungen der Haut und des Unterhautzellgewebes.

Die darunter betrachteten häufigsten unerwünschten Ereignisse werden als PT erfasst. Sollten zu Woche 48 Inkonsistenzen zwischen den hier eingeschlossenen Studien hinsichtlich des Auftretens von PTs bei mindestens 10% der Patienten bestehen, werden die Ereignisse aus der jeweils anderen Studie aus Vergleichbarkeitsgründen auch dann dargestellt, wenn sie nicht bei mindestens 10% der Patienten auftraten.

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, sind Ereignisse, die häufig in Zusammenhang mit EFV auftreten [44, 45] 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 - zumal dieser sich oft nicht verbergen lässt - 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-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 [7, 9, 10].

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.

Diese 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 [7, 31].

#### 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 Effekten<sup>4</sup> 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 verwendet<sup>5</sup> werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen

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

<sup>5</sup> 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.

Heterogenität der Studienergebnisse anhand des Maßes  $I^2$  und des statistischen Tests auf Vorliegen von Heterogenität<sup>6</sup> erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

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

Sind die Einzelstudien hinreichend ähnlich in ihrer Fragestellung, in ihren Populations- und Design-Charakteristika und die zu kombinierenden Studienresultate für die einzelnen Endpunkte hinreichend homogen (Bewertung auf der Heterogenität auf Basis des 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

---

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

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 [46]. Falls keine validierte Irrelevanzgrenze vorhanden ist, wird die standardisierte Mittelwertdifferenz (Hedges'  $g$ ) und die Lage des Konfidenzintervalls bezogen auf die Irrelevanzgrenze 0,2 betrachtet [47].

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 [48]. Zur Feststellung statistischer Heterogenität wird ein Signifikanzlevel von 0,2 verwendet. Für die  $I^2$ -Statistik gelten die im Cochrane-Handbuch [49] 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 z. B. 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 im Speziellen Sensitivitätsanalysen hinsichtlich der beiden Imputationsalgorithmen der FDA, TLOVR und Snapshot, durchgeführt [50].

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

- Geschlecht
- Alter
- Krankheitsschwere bzw. -stadium

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

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

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

Es werden 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, da sie im Studienprotokoll der Studie S102 präspezifiziert wurden:

- Alter (<40 Jahre versus  $\geq 40$  Jahre),

- Geschlecht (Männer versus Frauen),
- Rasse (Kaukasisch versus nicht-kaucasisch) – Ethnie wurde nicht berücksichtigt (s.u.),
- Therapieadhärenz (<95% versus ≥95%),

sowie als Maß der Krankheitsschwere

- Ausgangsviruslast ( $\leq 100.000$  versus  $> 100.000$  HIV-1-RNA-Kopien/mL) und
- CD4-Zellzahl zu Baseline ( $\leq 350$  versus  $> 350$  Zellen/ $\mu$ L).

Rasse und Ethnie sind zwei Begriffe, die im europäischen Sprachraum fast immer synonym verwendet werden. Die eigentliche 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 [51]. 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 [52]. In einer (älteren) Studie [53] 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-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. 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 Nutzaussage für die Subgruppe mit dem nicht signifikanten Effekt.

Im Falle signifikanter Interaktionen (Signifikanzniveau 0,20) 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 Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich<sup>7</sup>. 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“<sup>8</sup>, „Multiple-Treatment-Meta-Analysen“<sup>9</sup> oder auch „Netzwerk-Meta-Analysen“<sup>10</sup>, 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 ist<sup>11</sup>. Insgesamt ist es notwendig,

<sup>7</sup> 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.

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

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

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

<sup>11</sup> B. Schöttker, D. Lüthmann, D. Boulikhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen<sup>12</sup>.

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

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

Da zwei direkt vergleichende Studien vorliegen wird auf die Darstellung der Methodik des indirekten Vergleichs verzichtet.

### **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)

<sup>12</sup> 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.

- 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 [25]. 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 diesem Modul nur die Studien im entsprechenden Anwendungsgebiet aufgenommen (Infektion mit dem Humanen Immundefizienzvirus 1 (HIV-1) bei Erwachsenen im Alter von 18 Jahren und darüber, die nicht antiretroviral vorbehandelt sind).

Die Studien, die mit dem zu bewertenden Arzneimittel bei 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, durchgeführt wurden, sind in Modul 4B zu finden.

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-236-0102	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 4-4: 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-236-0103	Die Studie wurde nicht mit EFV/FTC/TDF als Vergleichstherapie durchgeführt.
GS-US-236-0128	Es liegen keine Studienergebnisse vor. Außerdem wurde die Studie nicht mit EFV/FTC/TDF als Vergleichstherapie durchgeführt.
GS-US-292-0102	Es liegen lediglich 24-Wochen-Daten vor. Außerdem wurde die Studie nicht mit EFV/FTC/TDF als Vergleichstherapie durchgeführt.
GS-US-292-0111	Es liegen keine Studienergebnisse vor. Außerdem wurde die Studie nicht mit EFV/FTC/TDF als Vergleichstherapie durchgeführt.
GS-US-292-0104	Es liegen keine Studienergebnisse vor. Außerdem wurde die Studie nicht mit EFV/FTC/TDF als Vergleichstherapie durchgeführt.

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.*

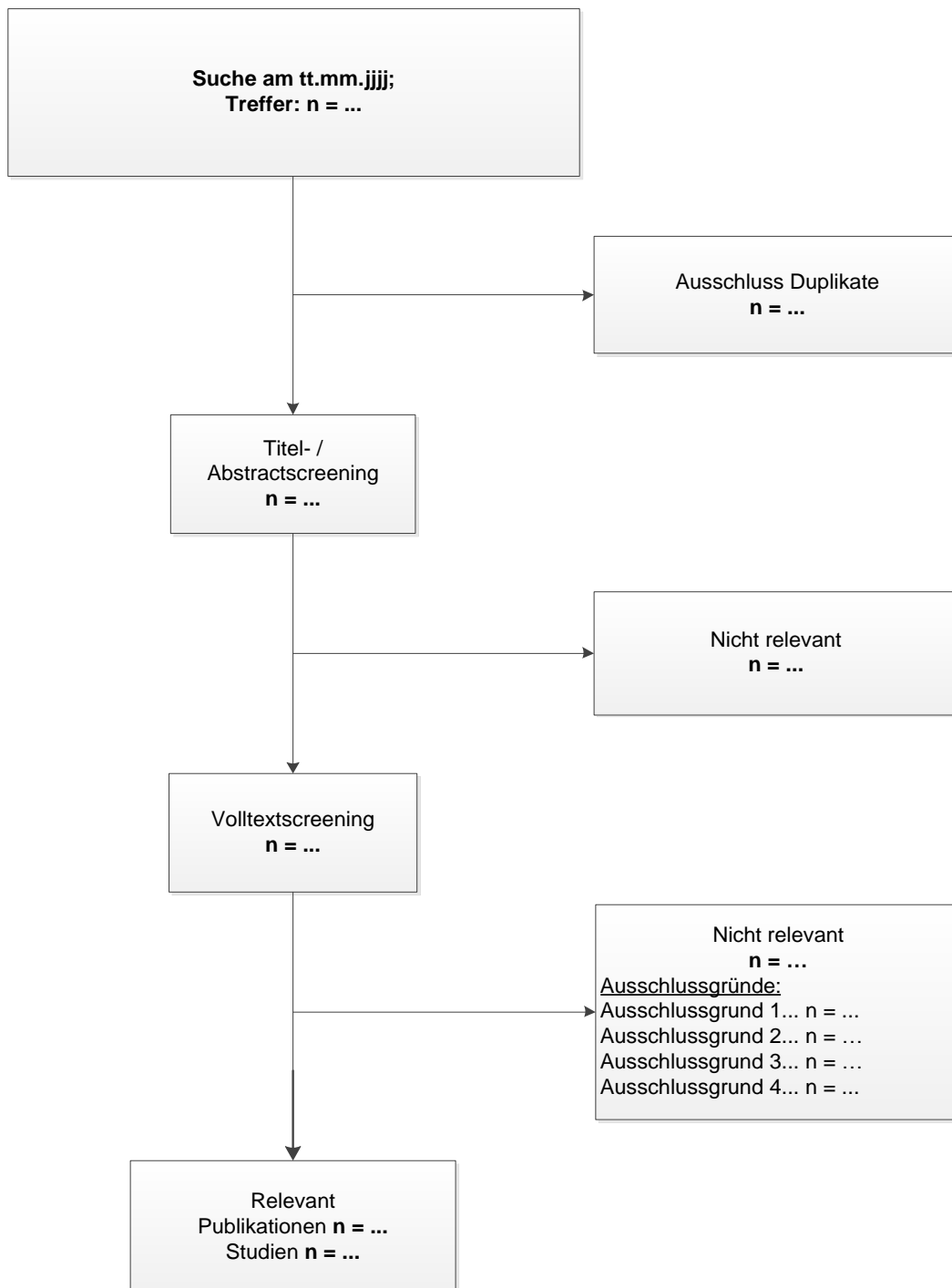


Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

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

#### 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 <sup>a</sup> )	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. <sup>b</sup> )
GS-US-236-0102	clinicaltrials.gov [NCT01095796], [54] ICTRP Search Portal [NCT01095796] [55]	ja	n. d.
GS-US-236-0104	clinicaltrials.gov [NCT00869557], [56] ICTRP Search Portal [NCT00869557] [57]	ja	n. d.

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

#### 4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

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

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



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

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

Studie	Studienkategorie			verfügbare Datenquellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
<b>ggf. Zwischenüberschrift zur Strukturierung des Studienpools</b>						
<b>placebokontrolliert</b>						
Nicht zutreffend						
<b>aktivkontrolliert, zweckmäßige Vergleichstherapie(n)</b>						
GS-US-236-0102	ja	ja	nein	ja [58, 59]	ja [54, 55]	ja [60]
GS-US-236-0104	ja	ja	nein	ja [61]	ja [56, 57]	ja [62]
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 [26]. 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 [25]. 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 EFV/FTC/TDF in der Indikation der HIV-1 Infektion bei antiretroviral nicht vorbehandelten erwachsenen Patienten, zu bewerten.

Im Weiteren werden die Studien GS-US-236-0102 und GS-US-236-0104 aus Übersichtlichkeitsgründen mit S102 bzw. 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.*

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

Tabelle 4-7: 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
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 Forschungsprogramm 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	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

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

<b>Studie</b>	<b>Studiendesign</b> <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	<b>Population</b> <relevante Charakteristika, z. B. Schweregrad>	<b>Interventionen</b> (Zahl der randomisierten Patienten)	<b>Studiendauer</b> <ggf. Run-in, Behandlung, Nachbeobachtung>	<b>Ort und Zeitraum der Durchführung</b>	<b>Primärer Endpunkt; patientenrelevante sekundäre Endpunkte</b>
				Forschungsprogramm dazu stoppt oder Follow-Up: 30 Tage		CD4-Zellzahl Mortalität Unerwünschte Ereignisse

Tabelle 4-8: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

<b>Studie</b>	<b>EVG/COBI/FTC/TDF</b>	<b>EFV/FTC/TDF</b>	<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 Tenofoviridisoproxil einmal täglich (EVG/COBI/FTC/TDF)	600 mg Efavirenz, 200 mg Emtricitabin, 245 mg Tenofoviridisoproxil 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 Tenofoviridisoproxil einmal täglich (EVG/COBI/FTC/TDF)	600 mg Efavirenz, 200 mg Emtricitabin, 245 mg Tenofoviridisoproxil 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

Tabelle 4-9: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Demografie, Safety Analyse-Set)

<b>Studie Gruppe</b>	<b>N</b>	<b>Alter (Jahre) Mittelwert (SD)</b>	<b>Geschlecht w/m (%)</b>	<b>Baseline BMI (kg/m<sup>2</sup>) Mittelwert (SD)</b>
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)
BMI = Body Mass Index, SD = Standardabweichung				

Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Rasse, Safety Analyse-Set)

Studie Gruppe	Rasse n (%)					
	Kaukasisch	Asiatisch	Schwarz (oder afrikanischer Abstammung) <sup>a</sup>	Indianer/ Urein- wohner Alaskas	Hawaiianer/ Pazifikinsulaner	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)

<sup>a</sup> Angabe in Studie S102  
k.A. = keine Angaben, n: Anzahl Patienten in der Kategorie

Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Ethnische Abstammung, Safety Analyse-Set)

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)

n: Anzahl der Patienten in der Kategorie

Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (HIV-1-RNA, Safety Analyse-Set)

Studie Gruppe	Baseline HIV-1-RNA (log <sub>10</sub> 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)

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

Tabelle 4-13: 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
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)

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

Tabelle 4-14: 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	
S102				
EVG/COBI/FTC/TDF	290 (83,3)	30 (8,6)	28 (8,0)	120,9 (31,6)
EFV/FTC/TDF	295 (83,8)	33 (9,4)	24 (6,8)	120,4 (33,7)
S104				
EVG/COBI/FTC/TDF	40 (83,3)	5 (10,4)	3 (6,3)	131,2 (33,6)
EFV/FTC/TDF	22 (95,7)	0 (0,0)	1 (4,3)	131,3 (31,2)

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 4-15: 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 <sup>a</sup> n (%)				Adhärenz -Rate <sup>a</sup> Mittel- wert (SD)
		<80%	≥80-<90%	≥90-<95%	≥95%	
<b>Woche 48</b>						
S102 <sup>b,c</sup>						
EVG/COBI/FTC/TDF	346 (99,4)	3 (0,9)	20 (5,8)	63 (18,2)	260 (75,1)	96,4 (5,0)
EFV/FTC/TDF	349 (99,1)	8 (2,3)	28 (8,0)	53 (15,2)	260 (74,5)	96,0 (6,4)
<b>Woche 60</b>						
S104 <sup>c</sup>						
EVG/COBI/FTC/TDF	46 (95,8)	1 (2,2)	1 (2,2)	11 (23,9)	33 (71,7)	95,7 (6,6)
EFV/FTC/TDF	22 (95,7)	0 (0,0)	1 (4,5)	4 (18,2)	17 (77,3)	97,0 (4,2)
<b>Woche 96</b>						
S102 <sup>b,c</sup>						
EVG/COBI/FTC/TDF	346 (99,4)	3 (0,9)	22 (6,4)	65 (18,8)	256 (74,0)	96,3 (5,1)
EFV/FTC/TDF	349 (99,1)	6 (1,7)	33 (9,5)	47 (13,5)	263 (75,4)	96,1 (6,2)
<sup>a</sup> Adhärenz-Rate berechnet sich basierend auf der Anzahl an Patienten, die mindestens eine Flasche zurückgaben <sup>b</sup> Adhärenz-Rate wurde berechnet basierend nur auf Anzahl der aktiven Medikamente <sup>c</sup> Adhärenz-Rate war wie folgt definiert: 100 x (Gesamtanzahl eingenommener Tabletten / Gesamtanzahl verschriebener Tabletten) n: Anzahl der Patienten in der Kategorie, SD = Standardabweichung						

Tabelle 4-16: Charakterisierung der Studienpopulationen (Patientendisposition) – RCT mit dem zu bewertenden Arzneimittel

Studie Disposition	EVG/COBI/FTC/TDF n (%)	EFV/FTC/TDF n (%)
S102		
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)



<b>Studie</b>	<b>EVG/COBI/FTC/TDF</b>	<b>EFV/FTC/TDF</b>
<b>Disposition</b>	<b>n (%)</b>	<b>n (%)</b>
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)
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)
<b>S104</b>		
Randomisiert	48	23
Behandlung erhalten	48 (100,0)	23 (100,0)

Studie	EVG/COBI/FTC/TDF	EFV/FTC/TDF
Disposition	n (%)	n (%)
Randomisierte Phase <sup>a</sup> bis zum Daten Cut-Off in Behandlung abgeschlossen	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)
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 <sup>a</sup> bis zum Daten Cut-Off in Studie abgeschlossen	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)
<sup>a</sup> 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

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 ( $\leq 100.000$  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-RNA Level ( $\leq 100.000$  Kopien/mL oder  $> 100.000$  Kopien/mL) stratifiziert und es wurde eine Blocklänge von 6 verwendet.

### **Studienpopulation**

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  $\geq 5.000$  Kopien/mL aufweisen, eine Lebenserwartung von  $\geq 1$  Jahr haben und außerdem eine adäquate Nierenfunktion, definiert durch eGFR  $\geq 70$  mL/min in Studie S102 bzw. eGFR  $\geq 80$  mL/min in Studie S104 gemäß Cockcroft-Gault, aufweisen. 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/m<sup>2</sup> und in der EFV/FTC/TDF-Gruppe bei durchschnittlich 26,1 kg/m<sup>2</sup>. In Studie S104 lag der durchschnittliche Baseline-BMI in der EVG/COBI/FTC/TDF-Gruppe bei 25,9 kg/m<sup>2</sup> und in der EFV/FTC/TDF-Gruppe bei 27,8 kg/m<sup>2</sup> (siehe Tabelle 4-9). 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 4-10). 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 4-11). 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 log<sub>10</sub> HIV-1-Kopien/mL im EVG/COBI/FTC/TDF-Arm und 4,78 log<sub>10</sub> Kopien/mL im EFV/FTC/TDF-Arm. In Studie S104 lag die Ausgangsviruslast im EVG/COBI/FTC/TDF-Arm bei 4,59 log<sub>10</sub> HIV-1-Kopien/mL und 4,58 log<sub>10</sub> HIV-1-Kopien/mL im EFV/FTC/TDF-Arm (siehe Tabelle 4-12). 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: 67,0%/33,0%; S104: 78,3%/21,7%). Die mittlere CD4-Zellzahl lag zu Baseline von Studie S102 bei 391 Zellen/μL im EVG/COBI/FTC/TDF-Arm und 382 Zellen/μL im EFV/FTC/TDF-Arm. In Studie S104 lag die mittlere CD4-Zellzahl zu Baseline im EVG/COBI/FTC/TDF-Arm bei 378 Zellen/μL und im EFV/FTC/TDF-Arm bei 454 Zellen/μL. 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/ $\mu$ L und 500 Zellen/ $\mu$ L. In Studie S102 waren dies im EVG/COBI/FTC/TDF-Arm in der Gruppe 201- $\leq$ 350 Zellen/ $\mu$ L 32,2% und in der Gruppe 351- $\leq$ 500 Zellen/ $\mu$ L 32,5% der Patienten. Im EFV/FTC/TDF-Arm lag dieser Anteil bei 27,3% in der 201- $\leq$ 350 Zellen/ $\mu$ L-Gruppe und 38,6% in der 351- $\leq$ 500 Zellen/ $\mu$ L-Gruppe (siehe Tabelle 4-13). 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. Zu Baseline war ein Großteil der Patienten klinisch asymptomatisch (EVG/COBI/FTC/TDF vs. EFV/FTC/TDF: S102: 83,3% vs. 83,8%; S104: 83,3% vs. 95,7%). In Studie S102 zeigten 8,6% der Patienten im EVG/COBI/FTC/TDF-Arm und 9,4% der Patienten im EFV/FTC/TDF-Arm symptomatische HIV-Infektionen und in Studie S104 waren dies 10,4% der Patienten im EVG/COBI/FTC/TDF-Arm und kein Patient im EFV/FTC/TDF-Arm. In Studie S102 hatten 8,0% der Patienten im EVG/COBI/FTC/TDF-Arm und 6,8% der Patienten im EFV/FTC/TDF-Arm zu Baseline Aids-definierende Erkrankungen. In Studie S104 waren dies 6,3% im EVG/COBI/FTC/TDF-Arm und 4,3% im EFV/FTC/TDF-Arm (siehe Tabelle 4-14). Die eGFR nach Cockcroft-Gault lag in Studie S102 zu Baseline im Mittel bei 120,9 mL/min im EVG/COBI/FTC/TDF-Arm und bei 120,4 mL/min im EFV/FTC/TDF-Arm. In Studie S104 lag die durchschnittliche eGFR bei 131,2 mL/min im EVG/COBI/FTC/TDF-Arm und bei 131,3 mL/min im EFV/FTC/TDF-Arm. Insgesamt sind die Behandlungsarme ausgeglichen. Kleinere Unbalanziertheiten, v. a. hinsichtlich der CD4-Zellzahl und dem HIV-Krankheitsstatus, indizieren hier, dass im EVG/COBI/FTC/TDF-Arm eher die etwas schwerer erkrankten HIV-1-infizierten Patienten waren.

Zur Berechnung der Adhärenz der Patienten wurde mit der Anzahl an Patienten gerechnet, die mindestens eine Flasche der Medikation wieder zurückgaben. Dies waren bis Woche 48 in Studie S102 99,4% im EVG/COBI/FTC/TDF-Arm und 99,1% im EFV/FTC/TDF-Arm (siehe Tabelle 4-15). 75,1% der Patienten im EVG/COBI/FTC/TDF-Arm und 74,5% der Patienten im EFV/FTC/TDF-Arm wiesen eine Adhärenzrate von  $\geq$ 95% auf. Die Adhärenzrate lag bei 18,2% der Patienten im EVG/COBI/FTC/TDF-Arm und 15,2% der Patienten im EFV/FTC/TDF-Arm zwischen  $\geq$ 90% und  $<$ 95%. Zu Woche 96 sind die Adhärenzraten sehr ähnlich verteilt. In Studie S104 gaben bis Woche 60 95,8% der Patienten im EVG/COBI/FTC/TDF-Arm mindestens eine Flasche wieder zurück und im EFV/FTC/TDF-Arm waren dies 95,7%. 71,7% der Patienten im EVG/COBI/FTC/TDF-Arm und 77,3% der Patienten im EFV/FTC/TDF-Arm wiesen eine Adhärenzrate  $\geq$ 95% auf. Eine Adhärenzrate unter 90% zeigten allgemein nur wenige Patienten.

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 4-16). 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; 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.

#### 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 4-17: 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			
S102	ja	ja	ja	ja	nein	nein	niedrig
S104	ja	ja	ja	ja	nein	nein	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

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 ( $\leq 100.000$  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 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.

Für Studie S104 wurden einige Endpunktdaten aus Vergleichbarkeitsgründen mit Studie S102 nachberechnet. Dies trifft auf virologisches Ansprechen (TLOVR), Zeit bis zum Verlust des virologischen Ansprechens und unerwünschte Ereignisse zu Woche 48 zu. Außerdem wurden die Woche 48-, 60- und 96-Daten für die Veränderung der CD4-Zellzahl nachberechnet [63, 64].

Außerdem wurden für beide Studien, S102 und S104, Hazard Ratios für die Zeit bis zum Verlust des virologischen Ansprechens nachberechnet [65-67].

### 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 4-18: 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 <sup>b</sup>
S102	ja	ja	ja	ja
S104	ja	nein <sup>a</sup>	ja	ja

<sup>a</sup> wurde aus Vergleichbarkeitsgründen mit Studie S102 nachberechnet.

<sup>b</sup> 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 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Gesamtrate unerwünschter Ereignisse <sup>a</sup>	Schwerwiegende unerwünschte Ereignisse <sup>a</sup>	Unerwünschte Ereignisse Grad 3-4 <sup>a</sup>	Therapieabbrüche aufgrund unerwünschter Ereignisse <sup>a</sup>
S102	ja	ja	ja	ja
S104	ja	ja	ja	ja

<sup>a</sup> 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 4-20: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Wichtige unerwünschte Ereignisse <sup>a</sup>	Unerwünschte Ereignisse von Interesse <sup>a</sup>	Häufigste unerwünschte Ereignisse (SOC) <sup>a</sup>
S102	ja	ja	ja
S104	ja	ja	ja

<sup>a</sup> 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 separat 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 4-21: 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. Virologisches Ansprechen wurde mittels Snapshot-Algorithmus oder TLOVR-Algorithmus erfasst.
S104	Anteil an Patienten mit Viruslast <50 Kopien/mL, zu Woche 24, 48 bzw. 60. 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 erfasst. Virologisches Ansprechen mittels TLOVR-Algorithmus, war in dieser Studie nicht präspezifiziert, wurde aber aus Vergleichsgründen nachanalysiert.

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

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

Tabelle 4-22: Bewertung des Verzerrungspotenzials für 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
S102	niedrig	ja	ja	nein	nein	niedrig
S104	niedrig	ja	ja	nein	nein <sup>a</sup>	niedrig

<sup>a</sup> Der Endpunkt virologisches Ansprechen (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.

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Der Endpunkt wurde in beiden Studien definiert als Absinken der HIV-1-RNA-Viruslast unter die Nachweisgrenze von 50 Kopien/mL. Dabei wurde in Studie S102 der FDA Snapshot-Algorithmus als primärer Analyse-Algorithmus gewählt, in Studie S104 waren der Missing=Failure-, Missing=Excluded- sowie der Missing/ART-Switch=Failure-Algorithmus primärer Analyse-Algorithmus. Dazu wurde das Analysefenster zu Woche 48 in beiden Studien definiert als Tag 309 bis inklusive Tag 378. 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.

### **Snapshot-Algorithmus**

Patienten wurden als **virologische Ansprecher** 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  $\geq 50$  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  $\geq 50$  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 48 in Studie S104 und Woche 96 in Studie S102 war diese Operationalisierung identisch, mit dem Unterschied dass zu Woche 96 das Woche 96-Analysefenster betrachtet wurde. Das Woche 96-Analysefenster war definiert als Tag 631 bis einschließlich Tag 714.

Zusätzlich zum Snapshot-Algorithmus wurde in Studie S102 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  $\geq 50$  Kopien/mL an zwei aufeinanderfolgenden Visiten oder zuletzt verfügbarer HIV-1-RNA-Wert während der Studie  $\geq 50$  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  $\geq 50$  Kopien/mL an zwei aufeinanderfolgenden Visiten oder zuletzt verfügbarer HIV-1-RNA-Wert  $\geq 50$  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 in Studie S102 zu Woche 96 identisch, mit dem Unterschied dass das Woche 96-Analysefenster (Tag 631 bis einschließlich Tag 714) betrachtet wurde. Zu Woche 96 wurde der Therapieabbruch aufgrund fehlender Wirksamkeit ebenfalls unter virologischem Versagen betrachtet.

A priori war für Studie S104 keine TLOVR-Analyse geplant. Aus Vergleichbarkeitsgründen wurden diese Analysen zusätzlich post-hoc durchgeführt. Da dies aber, unabhängig von den Daten, lediglich aus Vergleichbarkeitsgründen mit Studie S102 geschah, wird das Verzerrungspotenzial auf Endpunktebene dadurch nicht heraufgestuft.

Stattdessen wurde in Studie S104 als primärer Endpunkt das virologische Ansprechen zu Woche 24, erfasst mittels dreier verschiedener Algorithmen, definiert. Dieser Endpunkt wurde hinsichtlich Nichtunterlegenheit untersucht. Die drei verschiedenen Algorithmen waren wie folgt definiert:

- **Missing=Failure (M=F):** Alle fehlenden Werte werden als Versagen gewertet (HIV-1-RNA  $\geq 50$  Kopien/mL). Im Nenner wird der Anteil an Patienten mit HIV-1-RNA  $< 50$  Kopien/mL berücksichtigt.
- **Missing/ART-Switch=Failure (M/S=F):** Alle fehlenden Werte und HIV-1-RNA Daten nach einem Wechsel der ART werden als Versagen gewertet (HIV-1-RNA  $\geq 50$  Kopien/mL). Im Nenner wird der Anteil an Patienten mit HIV-1-RNA  $< 50$  Kopien/mL berücksichtigt.
- **Missing=Excluded (M=E):** Alle fehlenden Werte werden aus der Analyse ausgeschlossen, d. h. alle fehlenden Werte werden sowohl aus dem Zähler als auch aus dem Nenner der Berechnung des Anteils der virologischen Ansprecher entfernt.

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

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. Die Endpunkterheber waren verblindet. Das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine weiteren Hinweise auf ergebnisgesteuerte Berichterstattung. In Studie S104 wurde der Endpunkt virologisches Ansprechen (TLOVR) nachberechnet [64]. Da dies aber, unabhängig von den Daten, lediglich aus Vergleichbarkeitsgründen mit Studie S102 geschah, wird das Verzerrungspotenzial auf Endpunktebene für beide Studien als niedrig eingeschätzt.

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 [29].

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  $\geq 50$  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  $\geq 50$  Kopien/mL an zwei aufeinander folgenden Visiten oder zuletzt verfügbarer HIV-1 RNA-Wert  $\geq 50$  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 S102 zu Woche 96 und in Studie S104 waren 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 4-23: Ergebnisse für virologisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set)

Zeitpunkt Algorithmus Studie	EVG/COBI/FTC/ TDF		EFV/FTC/TDF		Behandlungseffekt	
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>	RD (%) [95%-KI] <sup>c</sup>
<b>Woche 24</b>						
<i>Missing =Failure<sup>d</sup></i>						
S104	48	43 (89,6)	23	20 (87,0)	1,29 [0,28;5,94], p=0,72	2,8 [-14,5;20,1]
<i>Missing/ART-Switch=Failure<sup>e</sup></i>						
S104	48	43 (89,6)	23	19 (82,6)	1,81 [0,44;7,50], p=0,39	7,2 [-11,7;26,0]
<i>Missing =Excluded<sup>f</sup></i>						
S104	45	43 (95,6)	21	20 (95,2)	1,07 [0,09;12,56], p=0,90	0,7 [-13,4;14,8]
<i>Snapshot</i>						
S104	48	43 (89,6)	23	20 (87,0)	1,29 [0,28;5,94], p=0,73	2,7 [-15,4;20,9]
<b>Woche 48</b>						
<i>Snapshot</i>						
S102	348	305 (87,6)	352	296 (84,1)	1,34 [0,87;2,06], p=0,17	3,6 [-1,6;8,8] <sup>§</sup>
S104	48	44 (91,7)	23	19 (82,6)	2,32 [0,52;10,24], p=0,26	9,2 [-9,9;28,3]
<i>TLOVR</i>						
S102	348	299 (85,9)	352	293 (83,2)	1,23 [0,81;1,85], p=0,31	2,7 [-2,6;8,1]
S104 <sup>h</sup>	48	44 (91,7)	23	19 (82,6)	2,32 [0,52;10,24], p=0,26	9,1 [-10,1;28,3]
<b>Woche 96</b>						
<i>Snapshot</i>						
S102	348	293 (84,2)	352	287 (81,5)	1,21 [0,81;1,79], p=0,35	2,7 [-2,9;8,3] <sup>§</sup>
<i>TLOVR</i>						
S102	348	276 (79,3)	352	272 (77,3)	1,13 [0,79;1,62], p=0,50	2,1 [-4,0;8,2]
<sup>a</sup> zusätzlich berechnet						
<sup>b</sup> p-Wert aus Cochrane Mantel Haenszel Test stratifiziert nach Baseline HIV-1-RNA Kategorie						
<sup>c</sup> basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen berechnet.						
<sup>d</sup> Nenner war Anzahl Patienten im ITT-Set. P-Wert und prozentuale Differenz basieren auf binärem Response: Ansprechen (HIV-1-RNA<50 Kopien/mL) und Versagen (HIV-1-RNA ≥50 Kopien/mL oder fehlend)						
<sup>e</sup> Nenner war Anzahl Patienten im ITT-Set. P-Wert und prozentuale Differenz basieren auf binärem Response: Ansprechen (HIV-1-RNA<50 Kopien/mL) und Versagen (HIV-1-RNA ≥50 Kopien/mL, fehlend oder ART-Switch). Patienten, die die Therapie abbrachen und keine follow-up Informationen auf neuer ART aufwiesen, wurden als „ART-Switch“ behandelt. Der kommerzielle EFV/FTC/TDF-Switch für Patienten, die zu EFV/FTC/TDF randomisiert und damit behandelt wurden, wurde nicht als ART-Switch eingestuft.						
<sup>f</sup> Nenner war Anzahl Patienten im ITT-Set mit nichtfehlenden HIV-1-RNA Werten zu jeder Visite. P-Wert und prozentuale Differenz basieren auf binärem Response: Ansprechen (HIV-1-RNA<50 Kopien/mL) und Versagen (HIV-1-RNA						

Zeitpunkt Algorithmus Studie	EVG/COBI/FTC/ TDF		EFV/FTC/TDF		Behandlungseffekt	
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>	RD (%) [95%-KI] <sup>c</sup>
<sup>e</sup> ≥50 Kopien/mL) <sup>g</sup> aufgrund zweier IDMC (Independent Data Monitoring Committee) Meetings, bei denen jeweils eine Wirksamkeitsanalyse zum Niveau $\alpha=0,001$ durchgeführt wurde, wurde, um das globale $\alpha$ -Niveau zu sichern, für den primären Endpunkt ein 95,2%-Konfidenzintervall (entsprechend: $\alpha=0,048$ ) berechnet. <sup>h</sup> Virologisches Ansprechen (TLOVR) wurde für Studie S104 nachberechnet (Quelle: [64])						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

In Studie S104 wurden für den primären Endpunkt drei Imputationsmethoden zur Berechnung des Anteils an virologischen Ansprechern zu Woche 24 verwendet. Mittels des Missing=Failure-Algorithmus ergibt sich das OR [95%-KI] zu 1,29 [0,28;5,94] mit  $p=0,72$ . Es besteht hier somit kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen. Die Risikodifferenz [95%-KI] beträgt 2,8 [-14,5;20,1]. Bei Verwendung des Missing/ART-Switch=Failure-Algorithmus kann die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF gezeigt werden (RD [95%-KI]: 7,2 [-11,7;26,0]). Mit einem OR [95%-KI] von 1,81 [0,44;7,50] mit  $p=0,39$  besteht jedoch keine statistisch signifikante Überlegenheit. Mittels des Missing=Excluded-Algorithmus ergibt sich das OR [95%-KI] zu 1,07 [0,09;12,56] mit  $p=0,90$ , woraus sich kein statistisch signifikanter Unterschied ableiten lässt. Die Risikodifferenz [95%-KI] beträgt 0,7 [-13,4;14,8]. Der Anteil an virologischen Ansprechern wurde auch mittels des FDA Snapshot-Algorithmus berechnet. Hier ergibt sich das OR [95%-KI] zu 1,29 [0,28;5,94] mit  $p=0,73$  und die Risikodifferenz [95%-KI] beträgt 2,7 [-15,4;20,9], was keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen anzeigt.

Zu Woche 48 zeigten sich mit EVG/COBI/FTC/TDF in beiden Studien numerisch mehr virologische Ansprecher, erfasst mittels Snapshot-Algorithmus, als im EFV/FTC/TDF-Arm. In Studie S102 sprachen 87,6% der Patienten im EVG/COBI/FTC/TDF- und 84,1% der Patienten im EFV/FTC/TDF-Arm virologisch an (OR [95%-KI]: 1,34 [0,87;2,06],  $p=0,17$ ), definiert durch einen Abfall der Viruslast unter die Nachweisgrenze (<50 Kopien/mL). In Studie S104 waren dies im EVG/COBI/FTC/TDF-Arm 91,7% der Patienten und 82,6% der Patienten im EFV/FTC/TDF-Arm (OR [95%-KI]: 2,32 [0,52;10,24],  $p=0,26$ ). In beiden Studien konnte die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze gezeigt werden (RD [95%-KI]: S102: 3,6 [-1,6;8,8], S104: 9,2 [-9,9;28,3]). Unter Verwendung des TLOVR-Algorithmus besteht in Studie S102 kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,23 [0,81;1,85],  $p=0,31$ ). Die Risikodifferenz [95%-KI] beträgt 2,7 [-2,6;8,1], was zeigt, dass hier die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze gezeigt werden konnte. In Studie S104 besteht bei Verwendung des TLOVR-Algorithmus kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 2,32 [0,52;10,24],  $p=0,26$ ). Die Risikodifferenz [95%-KI] zeigt aber auch hier



die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze (RD [95%-KI]: 9,1 [-10,1;28,3]).

In Studie S102 zeigten sich zu Woche 96 ähnliche Ergebnisse wie zu Woche 48. Die Anteile an virologischen Ansprechern reduzierten sich in beiden Behandlungsarmen leicht. Mit 84,2% virologischen Ansprechern im EVG/COBI/FTC/TDF-Arm und 81,5% virologischen Ansprechern im EFV/FTC/TDF-Arm, unter Verwendung des Snapshot-Algorithmus, sprachen numerisch mehr EVG/COBI/FTC/TDF-Patienten virologisch an als EFV/FTC/TDF-Patienten. Allerdings kann kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen gezeigt werden (OR [95%-KI]: 1,21 [0,81;1,79],  $p=0,35$ ). Die Risikodifferenz [95%-KI] von 2,7 [-2,9;8,3] zeigt jedoch, dass EVG/COBI/FTC/TDF auch zu Woche 96 gegenüber EFV/FTC/TDF nichtunterlegen zur 12%-Grenze ist. Mittels TLOVR-Algorithmus zeigt sich auch zu Woche 96 eine leichte Reduzierung der virologischen Ansprecher in beiden Behandlungsarmen (EVG/COBI/FTC/TDF: 79,3%, EFV/FTC/TDF: 77,3%). Das OR [95%-KI] beträgt hier 1,13 [0,79;1,62] mit  $p=0,50$ , was keinen statistisch signifikanten Unterschied zeigt. Anhand der Risikodifferenz [95%-KI] kann jedoch auch hier die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze gezeigt werden (RD [95%-KI]: 2,1 [-4,0;8,2]).

Tabelle 4-24: Ergebnisse für virologisches Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set)

Zeitpunkt Algorithmus Studie	EVG/COBI/FTC/ TDF		EFV/FTC/TDF		Behandlungseffekt	
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>	RD (%) [95%-KI] <sup>c</sup>
<b>Woche 24</b>						
<i>Missing=Failure<sup>d</sup></i>						
S104	45	43 (95,6)	20	19 (95,0)	1,13 [0,10;13,25], p=0,88	0,8 [-13,7;15,3]
<i>Missing/ART-Switch=Failure<sup>e</sup></i>						
S104	45	43 (95,6)	20	19 (95,0)	1,13 [0,10;13,25], p=0,88	0,8 [-13,7;15,3]
<i>Missing=Excluded<sup>f</sup></i>						
S104	45	43 (95,6)	20	19 (95,0)	1,13 [0,10;13,25], p=0,88	0,8 [-13,7;15,3]
<b>Woche 48</b>						
<i>Snapshot</i>						
S102	312	296 (94,9)	300	288 (96,0)	0,77 [0,36;1,66], p=0,54	-1,0 [-4,4;2,4] <sup>g</sup>
S104	45	44 (97,8)	20	19 (95,0)	2,32 [0,14;38,99], p=0,54	2,9 [-11,6;17,4]
<b>Woche 96</b>						
<i>Snapshot</i>						
S102	293	284 (96,9)	294	283 (96,3)	1,23 [0,50;3,01], p=0,65	0,7 [-2,4;3,7]
<sup>a</sup> zusätzlich berechnet <sup>b</sup> p-Wert aus Cochran Mantel Haenszel Test stratifiziert nach Baseline HIV-1-RNA Schicht <sup>c</sup> basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen berechnet. <sup>d</sup> Nenner war Anzahl Patienten im ITT-Set. P-Wert und prozentuale Differenz basieren auf binärem Response: Ansprechen (HIV-1-RNA<50 Kopien/mL) und Versagen (HIV-1-RNA ≥50 Kopien/mL oder fehlend) <sup>e</sup> Nenner war Anzahl Patienten im ITT-Set. P-Wert und prozentuale Differenz basieren auf binärem Response: Ansprechen (HIV-1-RNA<50 Kopien/mL) und Versagen (HIV-1-RNA ≥50 Kopien/mL, fehlend oder ART-Switch). Patienten, die die Therapie abbrachen und keine follow-up Informationen auf neuer ART aufwiesen, wurden als „ART-Switch“ behandelt. Der kommerzielle EFV/FTC/TDF-Switch für Patienten, die zu EFV/FTC/TDF randomisiert und damit behandelt wurden, wurde nicht als ART-Switch eingestuft. <sup>f</sup> Nenner war Anzahl Patienten im ITT-Set mit nichtfehlenden HIV-1-RNA Werten zu jeder Visite. P-Wert und prozentuale Differenz basieren auf binärem Response: Ansprechen (HIV-1-RNA<50 Kopien/mL) und Versagen (HIV-1-RNA ≥50 Kopien/mL) <sup>g</sup> aufgrund zweier IDMC (Independent Data Monitoring Committee) Meetings, bei denen jeweils eine Wirksamkeitsanalyse zum Niveau α=0,001 durchgeführt wurde, wurde, um das globale α-Niveau zu sichern, für den primären Endpunkt ein 95,2%-Konfidenzintervall (entsprechend: α=0,048) berechnet.						
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz						

Unter Verwendung der PP-Population zeigen sich in Studie S104 zu Woche 24 für alle drei Algorithmen die gleichen Ergebnisse. So wird mittels M=F-, M/S=F- und M=E-Algorithmus jeweils ein OR [95%-KI] von 1,13 [0,10;13,25] mit p=0,88 berechnet und eine RD [95%-KI] von 0,8 [-13,7;15,3], was keinen statistisch signifikanten Unterschied anzeigt.

Bei Verwendung des PP-Sets ergibt sich in Studie S102 mittels Snapshot ein OR [95%-KI] von 0,77 [0,36;1,66] mit  $p=0,54$ , was keinen statistisch signifikanten Unterschied anzeigt. Die Risikodifferenz [95%-KI] von -1,0 [-4,4;2,4] zeigt jedoch auch hier die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze. In Studie S104 ergibt sich ein OR [95%-KI] von 2,32 [0,14;38,99] mit  $p=0,54$  und eine RD [95%-KI] von 2,9 [-11,6;17,4], d. h. auch hier kann die Nichtunterlegenheit zur 12%-Grenze gezeigt werden.

Tabelle 4-25: Ergebnisse für virologisches Versagen (Wirksamkeit) aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 48)

Algorithmus Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
<b>Snapshot</b>					
S102					
HIV-1 RNA $\geq 50$ Kopien/mL	348	13 (3,7)	352	11 (3,1)	1,20 [0,53;2,72], $p=0,6575$
Therapieabbruch aufgrund fehlender Wirksamkeit	348	4 (1,1)	352	2 (0,6)	2,03 [0,37;11,18], $p=0,4138$
S104					
HIV-1 RNA $\geq 50$ Kopien/mL	48	1 (2,1)	23	1 (4,3)	0,47 [0,03;7,83], $p=0,5975$
Therapieabbruch aufgrund fehlender Wirksamkeit	48	0 (0,0)	23	0 (0,0)	n.b.
<b>TLOVR</b>					
S102					
Rebound	348	20 (5,7)	352	13 (3,7)	1,59 [0,78;3,25], $p=0,2034$
Never suppressed	348	4 (1,1)	352	3 (0,9)	1,35 [0,30;6,09], $p=0,6939$
S104 <sup>b</sup>					
Rebound	48	0 (0,0)	23	1 (4,3)	0,15 [0,01;3,95], $p=0,2587$
Never suppressed	48	1 (2,1)	23	0 (0,0)	1,48 [0,06;37,84], $p=0,8111$
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Quelle: [64]					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, OR = Odds Ratio					

Tabelle 4-25 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: 17 [4,9%], EFV/FTC/TDF: 13 [3,7%]) besteht in Studie S102 zu Woche 48 kein statistisch signifikanter Unterschied (OR [95%-KI]: 1,34 [0,64;2,80],  $p=0,4377$ ). In Studie S104 (EVG/COBI/FTC/TDF: 1 [2,1%], EFV/FTC/TDF: 1 [4,3%]) kann zu diesem Endpunkt ebenfalls kein statistisch signifikanter Unterschied zwischen den Behandlungen gefunden werden (OR [95%-KI]: 0,47 [0,03;7,83],  $p=0,5975$ ).

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

Auch 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, zu Woche 48 in Studie S102 (EVG/COBI/FTC/TDF: 24 [6,9%], EFV/FTC/TDF: 16 [4,5%]) kein statistisch signifikanter Unterschied (OR [95%-KI]: 1,56 [0,81;2,98],  $p=0,1833$ ). In Studie S104 besteht zu Woche 48 (EVG/COBI/FTC/TDF: 1 [2,1%], EFV/FTC/TDF: 1 [2,1%]) ebenfalls kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,47 [0,03;7,83],  $p=0,5975$ ).

Tabelle 4-26: Ergebnisse für virologisches Versagen (Wirksamkeit) aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 96)

Zeitpunkt Algorithmus Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
<b>Snapshot</b>					
S102					
HIV-1 RNA $\geq 50$ Kopien/mL	348	4 (1,1)	352	7 (2,0)	0,57 [0,17;1,98], $p=0,3779$
Therapieabbruch aufgrund fehlender Wirksamkeit	348	6 (1,7)	352	5 (1,4)	1,22 [0,37;4,03], $p=0,7471$
<b>TLOVR</b>					
S102					
Rebound	348	20 (5,7)	352	13 (3,7)	1,59 [0,78;3,25], $p=0,2034$
Never suppressed	348	3 (0,9)	352	2 (0,6)	1,52 [0,25;9,16], $p=0,6467$
Therapieabbruch aufgrund fehlender Wirksamkeit	348	0 (0,0)	352	1 (0,3)	0,34 [0,01;8,28], $p=0,5049$
<sup>a</sup> zusätzlich berechnet					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Auch zu Woche 96 kann hinsichtlich keines Kriteriums unter virologischem Versagen ein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (siehe Tabelle 4-26).

**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: 10 [2,9%], EFV/FTC/TDF: 12 [3,4%]) besteht in Studie S102 zu Woche 96 kein statistisch signifikanter Unterschied (OR [95%-KI]: 0,84 [0,36;1,97],  $p=0,6851$ ).

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

Auch mittels TLOVR-Algorithmus besteht hinsichtlich virologischen Versagens aufgrund Wirksamkeit, definiert durch die Anzahl an Patienten mit virologischem Rebound oder Patienten, die nie virologisches Ansprechen erreicht haben, zu Woche 96 in Studie S102 (ohne Berücksichtigung des Kriteriums „Therapieabbruch aufgrund fehlender Wirksamkeit“: EVG/COBI/FTC/TDF: 23 [6,6%], EFV/FTC/TDF: 15 [4,3%]; unter Berücksichtigung des Kriteriums „Therapieabbruch aufgrund fehlender Wirksamkeit“: EVG/COBI/FTC/TDF: 23 [6,6%], EFV/FTC/TDF: 16 [4,5%]) kein statistisch signifikanter Unterschied (ohne Berücksichtigung des Kriteriums „Therapieabbruch aufgrund fehlender Wirksamkeit“: OR [95%-KI]: 1,59 [0,82;3,10],  $p=0,1737$ ; unter Berücksichtigung des Kriteriums „Therapieabbruch aufgrund fehlender Wirksamkeit“: OR [95%-KI]: 1,49 [0,77;2,86],  $p=0,2365$ ).

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.*

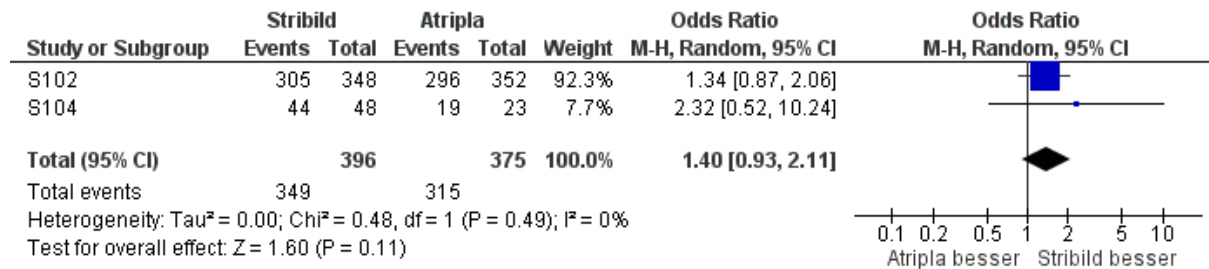


Abbildung 3: Meta-Analyse für virologisches Ansprechen (Snapshot) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, ITT-Set)

Die Meta-Analyse, basierend auf einem REM, zeigt, bei Verwendung des ITT-Sets, keinen statistisch signifikanten Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF hinsichtlich virologischem Ansprechen mittels Snapshot-Algorithmus (OR [95%-KI]: 1,40 [0,93;2,11], p=0,11). Mit p=0,49 und I<sup>2</sup>=0% zeigt der Heterogenitätstest keine signifikante Heterogenität zwischen den Studien, weswegen die Ergebnisse der Meta-Analyse, basierend auf einem FEM, nahezu identisch sind (OR [95%-KI]: 1,40 [0,92;2,11], p=0,11).

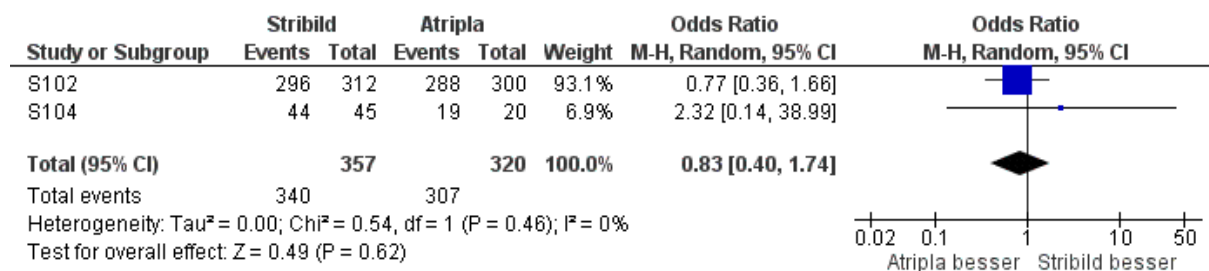


Abbildung 4: Meta-Analyse für virologisches Ansprechen (Snapshot) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, PP Set)

Auch basierend auf dem PP-Set, kann hier – bei homogener Datenlage (Heterogenitätstest: p=0,46; I<sup>2</sup>=0%) - kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen gezeigt werden (OR [95%-KI]: 0,83 [0,40;1,74], p=0,62). Eine Sensitivitätsanalyse, unter Verwendung des FEM, zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,83 [0,40;1,73], p=0,62).

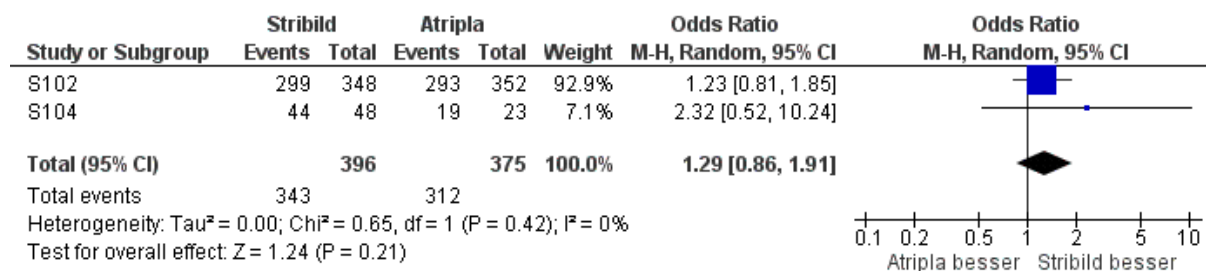


Abbildung 5: Meta-Analyse für virologisches Ansprechen (TLOVR) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, ITT-Set)

Die Meta-Analyse für virologisches Ansprechen (TLOVR) zeigt - bei homogener Datenlage ( $p=0,42$ ;  $I^2=0\%$ ) - keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen. Das OR [95%-KI] beträgt 1,29 [0,86;1,91] mit einem p-Wert von 0,21. Eine Sensitivitätsanalyse, unter Verwendung des FEM, zeigt robuste Ergebnisse (OR [95%-KI]: 1,28 [0,86;1,91]).

#### 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 4-27: Operationalisierung von Zeit bis Verlust des virologischen Ansprechens

Studie	Operationalisierung
S102	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, bestätigten HIV-1-RNA-Werts $\geq 50$ Kopien/mL oder HIV-1-RNA-Werts $\geq 50$ 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.
S104	In Studie S104 war der Endpunkt virologisches Ansprechen (TLOVR) nicht präspezifiziert. Aus Vergleichbarkeitsgründen wurde dieser Endpunkt mit der Definition wie in Studie S102 zu Woche 48 post-hoc definiert.

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

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

Tabelle 4-28: Bewertung des Verzerrungspotenzials für 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
S102	niedrig	ja	ja	nein	nein	niedrig
S104	niedrig	ja	ja	nein	nein <sup>a</sup>	niedrig

<sup>a</sup> Der Endpunkt Zeit bis zum Verlust des virologischen Ansprechens 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.

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Zeit bis zum Verlust des virologischen Ansprechens ist in Studie S102 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, Wechsel zu einem neuen Therapieregime, bestätigter HIV-1-RNA-Wert  $\geq 50$  Kopien/mL oder HIV-1-RNA-Wert  $\geq 50$  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 jeweils adäquat umgesetzt. Es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung. In Studie S104 war der Endpunkt Zeit bis zum Verlust des virologischen Ansprechens nicht präspezifiziert. Da die Nachberechnung, unabhängig von den Daten, lediglich aus Vergleichbarkeitsgründen mit Studie S102 geschah, wird das Verzerrungspotenzial für diesen Endpunkt dennoch für beide Studien als niedrig eingestuft.

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 4-29: Ergebnisse für Zeit bis Verlust des virologischen Ansprechens aus RCT mit dem zu bewertenden Arzneimittel

Studie	Patienten unter Risiko	Patienten mit Ereignis	Kaplan-Meier Schätzer (%) [95%-KI]	Behandlungseffekt
				HR <sup>a</sup> [95%-KI] <sup>b</sup> , p-Wert <sup>c</sup>
<b>Woche 48</b>				
S102				
EVG/COBI/FTC/TDF	303	49	14 [10,4;17,8]	0,92 [0,63;1,32], p=0,65
EFV/FTC/TDF	299	57	16 [12,4;20,0]	
S104				
EVG/COBI/FTC/TDF	44	4	8 [0,5;16,1]	0,47 [0,12;1,90], p=0,29 <sup>a</sup>
EFV/FTC/TDF	20	4	17 [1,9;32,9]	
<b>Woche 96</b>				
S102				
EVG/COBI/FTC/TDF	279	72	21 [16,5;25,1]	0,87 [0,63;1,19], p=0,28
EFV/FTC/TDF	279	80	23 [18,6;27,6]	
<sup>a</sup> Cox Proportional Hazards Modell, stratifiziert nach Ausgangsviruslast (≤100.000 vs.>100.000 Kopien/mL) <sup>b</sup> Wald-Konfidenzgrenzen <sup>c</sup> Log-Rank Test, stratifiziert nach Ausgangsviruslast (≤100.000 vs.>100.000 Kopien/mL)				
KI = Konfidenzintervall				

Zu Woche 48 standen in Studie S102 303 Patienten im EVG/COBI/FTC/TDF-Arm unter Risiko, wovon bei 49 Patienten ein Ereignis eintrat. Der Kaplan-Meier Schätzer [95%-KI] beträgt 14% [10,4%;17,8%]. Im EFV/FTC/TDF-Arm standen zu Woche 48 noch 299 Patienten unter Risiko, wovon bei 57 Patienten ein Ereignis eintrat. Der zugehörige Kaplan-Meier Schätzer [95%-KI] ergibt sich zu 16% [12,4%;20,0%]. Das HR [95%-KI] beträgt 0,92 [0,63;1,32] mit p=0,65, was keinen statistisch signifikanten Unterschied anzeigt.

In Studie S104 standen 44 Patienten im EVG/COBI/FTC/TDF-Arm unter Risiko, wovon bei 4 Patienten ein Ereignis eintrat. Der Kaplan-Meier Schätzer [95%-KI] beträgt 8% [0,5%;16,1%]. Im EFV/FTC/TDF-Arm standen zu Woche 48 noch 20 Patienten unter Risiko, wovon bei 4 Patienten ein Ereignis eintrat. Der zugehörige Kaplan-Meier Schätzer [95%-KI] ergibt sich zu 17% [1,9%;32,9%]. Das HR [95%-KI] beträgt 0,47 [0,12;1,90] mit p=0,29, was keinen statistisch signifikanten Unterschied anzeigt.

Zu Woche 96 standen in Studie S102 im EVG/COBI/FTC/TDF-Arm 279 Patienten unter Risiko, wovon bei 72 Patienten ein Ereignis eintrat. Der Kaplan-Meier Schätzer beträgt dabei 21% [16,5%;25,1%]. Im EFV/FTC/TDF-Arm standen zu Woche 96 279 Patienten unter

Risiko, wovon bei 80 Patienten ein Ereignis eintrat. Der daraus resultierende Kaplan-Meier Schätzer ist 23% [18,6%;27,6%]. Das HR [95%-KI] beträgt 0,87 [0,63;1,19] mit  $p=0,28$ , was keinen statistisch signifikanten Unterschied anzeigt.

In Abbildung 6 ist die Kaplan-Meier Kurve für die Zeit bis zum Verlust des virologischen Ansprechens in Studie S102 abgebildet. Gleich ab Anfang der Behandlung ist die EVG/COBI/FTC/TDF (Quad)-Kurve deutlich flacher als die EFV/FTC/TDF (ATR)-Kurve. Sie nähert sich mit der Zeit leicht an die EFV/FTC/TDF-Kurve an, bleibt jedoch stets flacher.

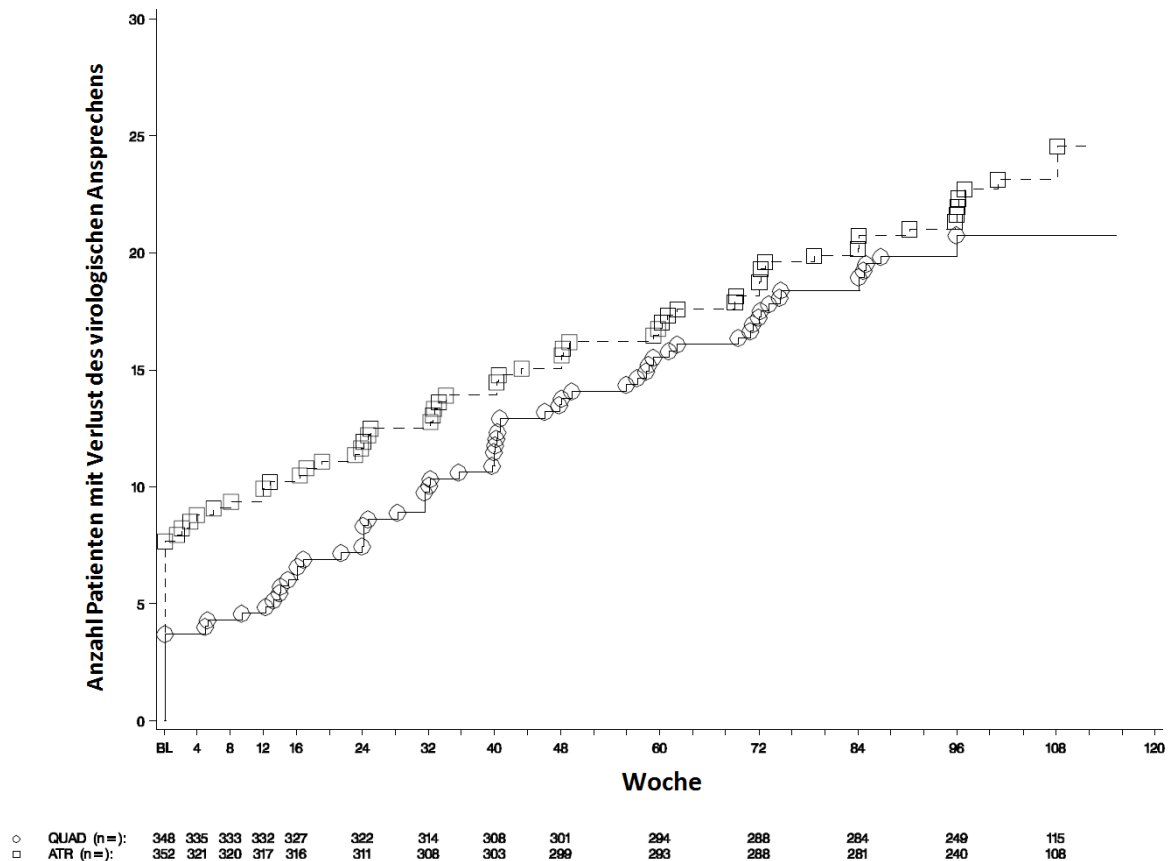


Abbildung 6: Kaplan-Meier Kurven für Zeit bis zum Verlust des virologischen Ansprechens, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Studie S102)

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

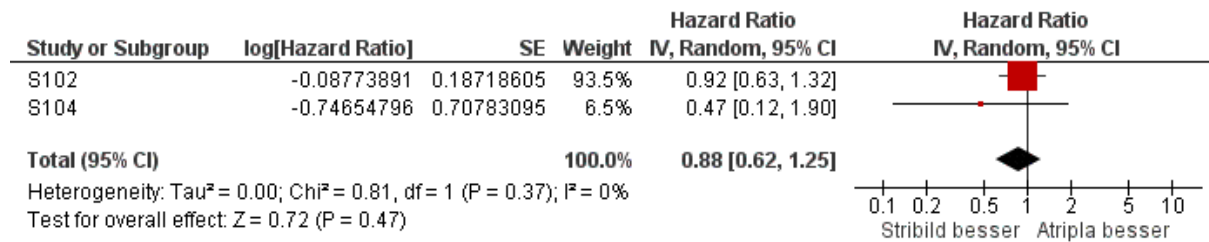


Abbildung 7: Meta-Analyse für Zeit bis Verlust des virologischen Ansprechens aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, ITT-Set)

Die Meta-Analyse zeigt – bei homogener Datenlage ( $p=0,37$ ;  $I^2=0\%$ ) – keinen statistisch signifikanten Unterschied zwischen den Behandlungen. Das HR [95%-KI] beträgt 0,88 [0,62;1,25] mit  $p=0,47$ . Eine Sensitivitätsanalyse, unter Verwendung eines FEM, zeigt robuste Ergebnisse (HR [95%-KI]: 0,88 [0,62;1,25],  $p=0,47$ ).

#### 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 4-30: 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 bzw. 60.

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 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
S102	niedrig	ja	ja	nein	nein	niedrig
S104	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-Zellen zwischen Baseline und Woche 48, Woche 60 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 4-16), 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.

Insbesondere zu Beginn einer ART ist dieser Endpunkt hoch patientenrelevant, weil durch eine raschere Dynamik der Entwicklung der CD4-Zellzahl das Immunsystem schneller rekonstituiert und so der für die Entwicklung Aids-definierender Erkrankungen relevante, niedrige CD4-Zellbereich rascher verlassen wird. Hier wird somit dem früheren Zeitpunkt zu Woche 48 eine besondere Relevanz zugemessen. Die Endpunkterheber waren verblindet. Es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotenzial beeinflussende Punkte. Das Verzerrungspotenzial wird somit für beide Studien 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.

### Woche 48 (Studie S102 und S104)

Tabelle 4-32: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 48)

Studie Behandlung	Baseline		Woche 48 <sup>a</sup>		Veränderung		Adj. <sup>b</sup> MWD [95%-KI] p-Wert <sup>a</sup>
	N	MW (SD)	N	MW (SD)	N	MW (SD)	
<b>Woche 48</b>							
S102							
EVG/COBI/FTC/ TDF	348	391 (188,6)	325	631 (249,1)	325	239 (167,2)	33 [8;58], p=0,009
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)	74 [-15;164], p=0,10
EFV/FTC/TDF	23	454 (137,6)	21	628 (201,7)	21	166 (158,5)	
<sup>a</sup> Quelle: [64, 68] <sup>b</sup> S102: ANOVA adjustiert für Baseline HIV-1-RNA Level; S104: aus ANOVA adjustiert für Baseline HIV-1-RNA Kategorie ( $\leq 100.000$ und $> 100.000$ Kopien/mL)  Adj. MWD = Adjustierte Mittelwertdifferenz, KI = Konfidenzintervall, MW = Mittelwertdifferenz, N: Anzahl Patienten in der Auswertung, SD = Standardabweichung							

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/ $\mu$ L. Diese konnte bis Woche 48 um durchschnittlich 239 Zellen/ $\mu$ L (SD: 167,2) auf 631 Zellen/ $\mu$ L (SD: 249,1) gesteigert werden. Im EFV/FTC/TDF-Arm wiesen die Patienten zu Baseline durchschnittlich eine CD4-Zellzahl von 382 Zellen/ $\mu$ L (SD: 170,2) auf; diese stieg bis Woche 48 um durchschnittlich 206 Zellen/ $\mu$ L (SD: 153,4) auf 586 Zellen/ $\mu$ 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/ $\mu$ L (SD: 179,8) im EVG/COBI/FTC/TDF-Arm bzw. 454 Zellen/ $\mu$ 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/ $\mu$ L (SD: 172,8) auf 619 Zellen/ $\mu$ L (SD: 269,5), wohingegen die CD4-Zellzahl der EFV/FTC/TDF-Patienten lediglich um 166 Zellen/ $\mu$ L (SD: 158,5) auf 628 Zellen/ $\mu$ 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).

Tabelle 4-33: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set, Woche 48)

Studie Behandlung	Baseline		Woche 48 <sup>a</sup>		Veränderung		MWD [95%-KI] p-Wert <sup>b</sup>
	N	MW (SD)	N	MW (SD)	N	MW (SD)	
<b>Woche 48</b>							
S102							
EVG/COBI/FTC/ TDF	293	392 (188,1)	289	642 (246,9)	289	246 (167,2)	36,00 [9,96;62,04] p=0,0069
EFV/FTC/TDF	294	383 (168,7)	290	592 (213,1)	290	210 (152,1)	
S104							
EVG/COBI/FTC/ TDF	45	379 (184,6)	45	619 (269,5)	45	240 (172,8)	78,00 [-9,01;165,01], p=0,0918
EFV/FTC/TDF	20	460 (144,8)	20	622 (205,2)	20	162 (161,7)	
<sup>a</sup> Quelle: [64, 68] <sup>b</sup> zusätzlich berechnet Adj. MWD = Adjustierte Mittelwertdifferenz, KI = Konfidenzintervall, MW = Mittelwertdifferenz, N: Anzahl Patienten in der Auswertung, SD = Standardabweichung							

In Studie S102 hatten die Patienten im EVG/COBI/FTC/TDF-Arm und unter Verwendung der PP-Population zu Baseline eine durchschnittliche CD4-Zellzahl von 392 Zellen/ $\mu$ L. Diese stieg bis Woche 48 um durchschnittlich 246 Zellen/ $\mu$ L (SD: 167,2) auf 642 Zellen/ $\mu$ L (SD: 246,9) an. Im EFV/FTC/TDF-Arm wiesen die Patienten zu Baseline durchschnittlich eine CD4-Zellzahl von 383 Zellen/ $\mu$ L (SD: 168,7) auf; diese stieg bis Woche 48 um durchschnittlich 210 Zellen/ $\mu$ L (SD: 152,1) auf 592 Zellen/ $\mu$ L (SD: 213,1) an. Daraus ergibt sich eine adj. MWD [95%-KI] von 36,00 [9,96;62,04] mit  $p=0,0069$ , 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 379 Zellen/ $\mu$ L (SD: 184,6) im EVG/COBI/FTC/TDF-Arm bzw. 460 Zellen/ $\mu$ L (SD: 144,8) im EFV/FTC/TDF-Arm auf. Im EVG/COBI/FTC/TDF-Arm stieg die CD4-Zellzahl um durchschnittlich 240 Zellen/ $\mu$ L (SD: 172,8) auf 619 Zellen/ $\mu$ L (SD: 269,5) an, wohingegen die CD4-Zellzahl der EFV/FTC/TDF-Patienten lediglich um 162 Zellen/ $\mu$ L (SD: 161,7) auf 622 Zellen/ $\mu$ L (SD: 205,2) anstieg. Der Unterschied zwischen den beiden Behandlungen ist allerdings nicht signifikant (adj. MWD [95%-KI]: 78,00 [-9,01;165,01],  $p=0,0918$ ).

**Woche 60 (Studie S102 und S104)**

Tabelle 4-34: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 60)

Studie Behandlung	Baseline		Woche 60 <sup>a</sup>		Veränderung		Adj. <sup>b</sup> MWD [95%-KI] p-Wert <sup>a</sup>
	N	MW (SD)	N	MW (SD)	N	MW (SD)	
<b>Woche 60</b>							
S102 <sup>b</sup>							
EVG/COBI/ FTC/TDF	348	391 (188,6)	322	646 (250,0)	322	255 (173,9)	22 [-4;47], p=0,093
EFV/FTC/TDF	352	382 (170,2)	312	616 (210,3)	312	232 (152,3)	
S104							
EVG/COBI/ FTC/TDF	48	378 (179,8)	45	618 (253,2)	45	239 (169,7)	36 [-59;130], p=0,46
EFV/FTC/TDF	23	454 (137,6)	21	665 (267,9)	21	203 (195,4)	
<sup>a</sup> Quelle: [64, 69]							
<sup>b</sup> aus ANOVA adjustiert für Baseline HIV-1-RNA Kategorie ( $\leq 100.000$ und $> 100.000$ Kopien/mL)							
Adj. MWD = Adjustierte Mittelwertdifferenz, KI = Konfidenzintervall, MW = Mittelwertdifferenz, N: Anzahl Patienten in der Auswertung, SD = Standardabweichung							

In Studie S102, erhöhte sich die CD4-Zellzahl der EVG/COBI/FTC/TDF-Patienten nach 60 Wochen um durchschnittlich 255 Zellen/ $\mu$ L (SD: 173,9) auf 646 Zellen/ $\mu$ L (SD: 250,0). Im EFV/FTC/TDF stieg sie um durchschnittlich 232 Zellen/ $\mu$ L (SD: 152,3) auf 616 Zellen/ $\mu$ L (SD: 210,3). Daraus ergibt sich eine adj. MWD [95%-KI] von 22 [-4;47] mit p=0,093, woraus sich kein statistisch signifikanter Unterschied ableiten lässt.

Nach 60 Wochen Behandlung erhöhte sich die CD4-Zellzahl in Studie S104 im EVG/COBI/FTC/TDF-Arm um durchschnittlich 239 Zellen/ $\mu$ L (SD: 169,7) auf 618 Zellen/ $\mu$ L (SD: 253,2) und im EFV/FTC/TDF-Arm stieg sie um 203 Zellen/ $\mu$ L (SD: 195,4) auf 665 Zellen/ $\mu$ L (SD: 267,9) an. Die adj. MWD [95%-KI] ergibt sich damit zu 36 [-59;130] mit p=0,46, was keinen statistisch signifikanten Unterschied zeigt.

**Woche 96 (nur Studie S102)**

Tabelle 4-35: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (ITT Analyse-Set, Woche 96)

Studie Behandlung	Baseline		Woche 96 <sup>a</sup>		Veränderung		Adj. <sup>b</sup> MWD [95%-KI] p-Wert <sup>a</sup>
	N	MW (SD)	N	MW (SD)	N	MW (SD)	
<b>Woche 96</b>							
S102							
EVG/COBI/ FTC/TDF	348	391 (188,6)	307	688 (263,6)	307	295 (213,3)	22 [-10;54], p=0,19
EFV/FTC/ TDF	352	382 (170,2)	302	655 (237,1)	302	273 (189,7)	
<sup>a</sup> Quelle: [69]							
<sup>b</sup> aus ANOVA adjustiert für Baseline HIV-1-RNA Kategorie ( $\leq 100.000$ und $> 100.000$ Kopien/mL)							
Adj. MWD = Adjustierte Mittelwertdifferenz, KI = Konfidenzintervall, MW = Mittelwertdifferenz, N: Anzahl Patienten in der Auswertung, SD = Standardabweichung							

Nach 96 Wochen Behandlung erhöhte sich die CD4-Zellzahl der EVG/COBI/FTC/TDF-Patienten durchschnittlich um 295 Zellen/ $\mu$ L (SD: 213,3) auf 688 Zellen/ $\mu$ L (SD: 263,6) und im EFV/FTC/TDF-Arm erhöhte sich die CD4-Zellzahl der Patienten um durchschnittlich 273 Zellen/ $\mu$ L (SD: 189,7) auf 655 Zellen/ $\mu$ L (SD: 237,1).

Tabelle 4-36: Ergebnisse für Veränderung der CD4-Zellzahl aus RCT mit dem zu bewertenden Arzneimittel (PP Analyse-Set, Woche 96)

Studie Behandlung	Baseline		Woche 96 <sup>a</sup>		Veränderung		Adj. <sup>b</sup> MWD [95%-KI] p-Wert <sup>a</sup>
	N	MW (SD)	N	MW (SD)	N	MW (SD)	
<b>Woche 96</b>							
S102							
EVG/COBI/ FTC/TDF	293	392 (188,1)	288	694 (265,8)	288	298 (216,0)	23,00 [-9,96;55,96], p=0,1718
EFV/FTC/TDF	294	383 (168,7)	289	656 (235,4)	289	275 (186,8)	
<sup>a</sup> Quelle: [69]							
<sup>b</sup> aus ANOVA adjustiert für Baseline HIV-1-RNA Kategorie ( $\leq 100.000$ und $> 100.000$ Kopien/mL)							
Adj. MWD = Adjustierte Mittelwertdifferenz, KI = Konfidenzintervall, MW = Mittelwertdifferenz, N: Anzahl Patienten in der Auswertung, SD = Standardabweichung							

Im EVG/COBI/FTC/TDF-Arm erhöhte sich die CD4-Zellzahl bis Woche 96 um 298 Zellen/ $\mu$ L (SD: 216,0) auf 694 Zellen/ $\mu$ L (SD: 265,8). Im EFV/FTC/TDF-Arm erhöhte sich die CD4-Zellzahl um 275 Zellen/ $\mu$ L (SD: 186,8) auf durchschnittlich 656 Zellen/ $\mu$ L (SD: 235,4). Es besteht somit in der PP-Population zu Woche 96 kein statistisch signifikanter Unterschied zwischen den Behandlungen.



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

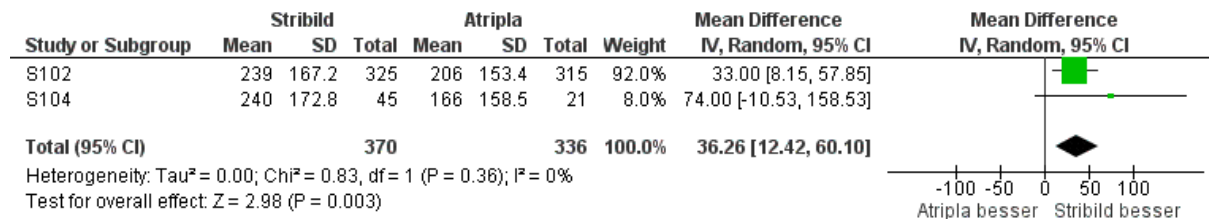


Abbildung 8: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, MWD, ITT; basierend auf den nicht-adjustierten Mittelwerten der Veränderung)

Die Meta-Analyse zeigt – bei homogener Datenlage (Heterogenitätstest:  $p=0,36$ ,  $I^2=0\%$ ) – 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.

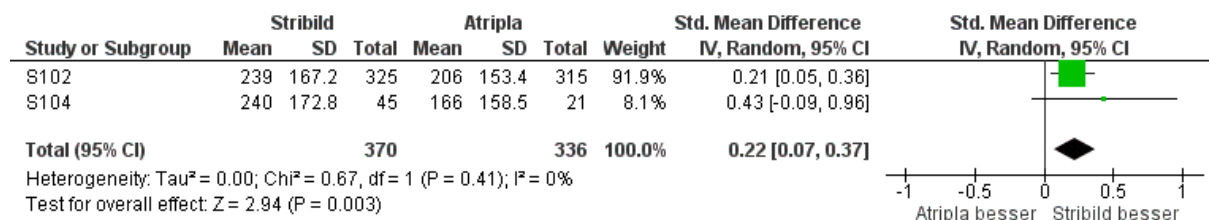


Abbildung 9: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, Hedges' g, ITT)

Die Meta-Analyse mittels Hedges' g zeigt einen statistisch signifikanten, aber, gemäß IQWiG Methoden 4.0 [47], klinisch nicht relevanten Effekt (Hedges' g [95%-KI]: 0,22 [0,07;0,37],  $p=0,003$ ), da die Untergrenze des Konfidenzintervalls die Grenze von 0,2 nicht übersteigt.

Die Meta-Analyse unter Verwendung der PP-Population ist im Folgenden dargestellt.

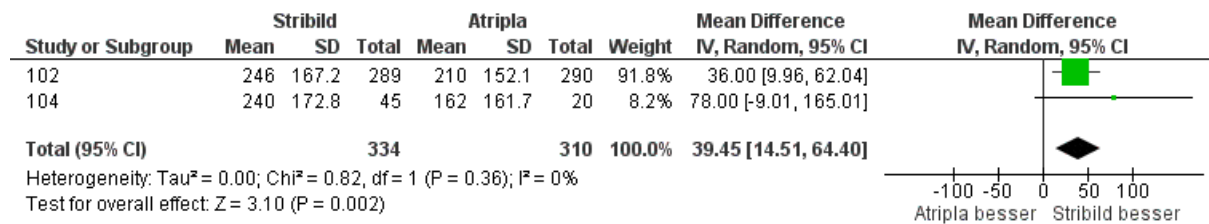


Abbildung 10: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, MWD, PP)

Unter Verwendung der PP-Population zeigt die Meta-Analyse der Veränderung der CD4-Zellzahl bis Woche 48 ähnliche Ergebnisse. Bei homogener Datenlage ( $p=0,36$ ;  $I^2=0\%$ ) ist der Unterschied zwischen den Behandlungen statistisch signifikant zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 39,45 [14,51;64,40]). Eine Sensitivitätsanalyse unter Verwendung des FEM zeigt robuste Ergebnisse.

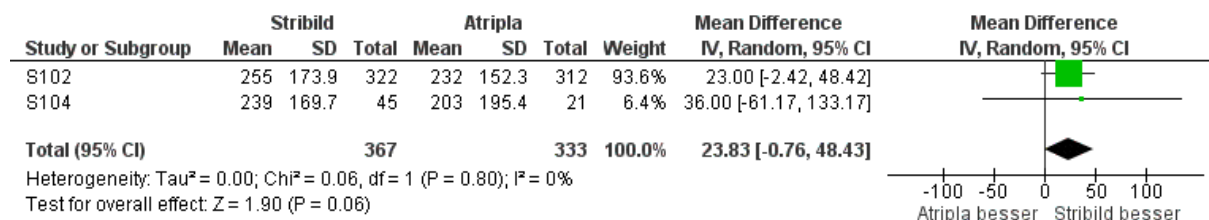


Abbildung 11: Meta-Analyse für Veränderung der CD4-Zellzahl aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 60, REM, MWD, ITT; basierend auf den nicht-adjustierten Mittelwerten der Veränderung)

Die Meta-Analyse zu Woche 60 zeigt - bei homogener Datenlage (Heterogenitätstest:  $p=0,80$ ;  $I^2=0\%$ ) - keinen statistisch signifikanten Unterschied zwischen den Behandlungen. Der OR-Gesamtschätzer [95%-KI] beträgt 23,83 [-0,76; 48,43]. Eine Sensitivitätsanalyse mittels FEM, bestätigt die Robustheit der Ergebnisse (OR [95%-KI]: 23,83 [-0,76;48,43],  $p=0,06$ ).

**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
S102	<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 neurologische &amp; psychiatrische Ereignisse sowie Hautausschläge.</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) 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 <math>\geq 10\%</math> der Patienten in einem der beiden Behandlungsarme) als PTs betrachtet.</p>
S104	<p><u>Mortalität</u> Anteil verstorbenen 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 und 60</p> <p><u>Schwerwiegende unerwünschte Ereignisse</u></p>

Studie	Operationalisierung
	<p>Anteil an Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 60</p> <p><u>Unerwünschte Ereignisse Grad 3-4</u></p> <p>Anteil an Patienten mit mindestens einem unerwünschten Ereignis Grad 3-4 (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 60</p> <p><u>Therapieabbrüche aufgrund unerwünschter Ereignisse</u></p> <p>Anteil an Patienten, die die Studie aufgrund eines unerwünschten Ereignisses abgebrochen haben (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 60</p> <p><u>Wichtige unerwünschte Ereignisse</u></p> <p>Anteil an Patienten mit mindestens einem wichtigen unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 60. Wichtige unerwünschte Ereignisse waren neurologische &amp; psychiatrische Ereignisse sowie Hautausschläge.</p> <p><u>Unerwünschte Ereignisse von Interesse</u></p> <p>Anteil an Patienten mit mindestens einem unerwünschten Ereignis von Interesse (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 60. Unerwünschte Ereignisse von Interesse waren renale Ereignisse sowie Knochenfrakturen.</p> <p><u>Häufigste unerwünschte Ereignisse</u></p> <p>Anteil an Patienten mit mindestens einem häufigsten unerwünschten Ereignis (Zeitraum zwischen erster und letzter Behandlungsdosis + 30 Tage) bis Woche 48 und 60. 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) 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 <math>\geq 10\%</math> 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
S102	niedrig	ja	ja	nein	nein	niedrig
S104	niedrig	ja	ja	unklar*	nein	niedrig

\* 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.

*Begründen Sie für jede Studie die abschließende Einschätzung.*

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 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 für Studie S104 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. Diese waren in den beiden Studien 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)

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:

- 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**.
- 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.

Des Weiteren wurden unerwünschte Ereignisse, welche besonders in Zusammenhang mit Tenofoviridisoproxil (TDF), welches in jeweils beiden Behandlungsarmen der beiden hier betrachteten Studien gegeben wird, auftreten, betrachtet. Diese umfassten:

- 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

In den beiden Studien waren die in den einzelnen Gruppen eingeschlossenen PT, bis auf neurologische & psychiatrische Ereignisse, die gleichen. Bei neurologischen & psychiatrischen Ereignissen wurden in Studie S104 in dieser Gruppe einige PT weniger betrachtet. Die PT, die in beiden Studien eingeschlossen waren, sind oben fett gedruckt.

Zu dem HLG „Frakturen“ wurde in Studie S102 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 den Studien 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

Darunter werden jeweils die am häufigsten auftretenden PT beobachtet. Dies sind die Erkrankungen, die bei mindestens 10% der Patienten in einem der Behandlungsarme auftraten. Sollten zu Woche 48 Inkonsistenzen zwischen den Studien S102 und S104 hinsichtlich des Auftretens von PT bei mindestens 10% der Patienten bestehen, werden die Ereignisse aus der jeweils anderen Studie aus Vergleichbarkeitsgründen auch dann dargestellt, wenn sie nicht bei mindestens 10% der Patienten auftraten. Allgemein gilt hier, dem Vorgehen des IQWiG folgend [29], dass die einzelnen PT als durch die jeweilige SOC

abgedeckt angesehen werden. Sie tragen somit nicht zusätzlich zur jeweiligen SOC zur Nutzenbewertung bei.

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 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.

*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 (Studie S102 und S104)

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

Endpunkt Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
<b>Mortalität</b>					
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.
<b>Gesamtrate unerwünschter Ereignisse</b>					
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
<b>Schwerwiegende unerwünschte Ereignisse</b>					
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
<b>Therapieabbrüche aufgrund unerwünschter Ereignisse</b>					
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
<b>Unerwünschte Ereignisse Grad 3-4</b>					
S102	348	45 (12,9)	352	39 (11,1)	1,19 [0,75;1,88], p=0,4514
S104	48	2 (4,2)	23	2 (8,7)	0,46 [0,06;3,46], p=0,4483
<sup>a</sup> zusätzlich berechnet KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, n.z. = nicht zutreffend, OR = Odds Ratio					



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

Endpunkt Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
<b>Neurologische &amp; psychiatrische Ereignisse</b>					
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=0,2426 <sup>a</sup>
<b>Hautausschläge</b>					
S102	348	59 (17,0)	352	98 (27,8)	0,53 [0,37;0,76], p<0,001
S104	48	9 (18,8)	23	6 (26,1)	0,65 [0,20;2,13], p=0,4802 <sup>a</sup>
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Fisher's exakter Test 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)

Endpunkt Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert
<b>Renale Ereignisse</b>					
S102	348	5 (1,4)	352	1 (0,3)	5,12 [0,59;44,02], p=0,12 <sup>b</sup>
S104	48	0 (0,0)	23	0 (0,0)	n.b.
<b>Knochenfrakturen</b>					
S102	348	6 (1,7)	352	6 (1,7)	1,01 [0,32;3,17], p=1,00 <sup>b</sup>
S104	48	1 (2,1)	23	0 (0,0)	1,48 [0,06;37,84], p=0,8111 <sup>a</sup>
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Fisher's exakter Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, n.z. = nicht zutreffend, OR = Odds Ratio					

Tabelle 4-42: Ergebnisse für Erkrankungen des Gastrointestinaltrakts (SOC, Grenze:  $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Endpunkt Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
<b>Erkrankungen des Gastrointestinaltrakts (SOC)</b>					
S102	348	188 (54,0)	352	161 (45,7)	1,39 [1,04;1,88], p=0,0286
S104	48	23 (47,9)	23	7 (30,4)	2,10 [0,73;6,03], p=0,1667
<b>Diarrhoe</b>					
S102	348	80 (23,0)	352	66 (18,8)	1,29 [0,90;1,87], p=0,1682
S104	48	10 (20,8)	23	4 (17,4)	1,25 [0,35;4,51], p=0,7333
<b>Übelkeit<sup>b</sup></b>					
S102	348	72 (20,7)	352	48 (13,6)	1,65 [1,11;2,46], p=0,0139
S104	48	4 (8,3)	23	2 (8,7)	0,95 [0,16;5,63], p=0,9590
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> In Studie S104 trat Übelkeit nicht bei mehr als 10% der Patienten in einem der Behandlungsarme auf. Aus Vergleichbarkeitsgründen werden die Ereignisse hier dennoch dargestellt.					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Tabelle 4-43: Ergebnisse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC, Grenze:  $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Endpunkt Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)</b>					
S102	348	68 (19,5)	352	99 (28,1)	0,62 [0,44;0,88], p=0,0080
S104	48	13 (27,1)	23	7 (30,4)	0,85 [0,28;2,53], p=0,7690
<b>Fatigue</b>					
S102	348	40 (11,5)	352	45 (12,8)	0,89 [0,56;1,40], p=0,6015
S104	48	8 (16,7)	23	4 (17,4)	0,95 [0,25;3,55], p=0,9392
<sup>a</sup> 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 Erkrankungen des Nervensystems (SOC, Grenze:  $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Endpunkt Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
<b>Erkrankungen des Nervensystems (SOC)</b>					
S102	348	95 (27,3)	352	146 (41,5)	0,53 [0,39;0,73], p=0,0001
S104	48	15 (31,3)	23	8 (34,8)	0,85 [0,30;2,44], p=0,7661
<b>Schwindel</b>					
S102	348	23 (6,6)	352	86 (24,4)	0,22 [0,13;0,36], p<0,00001
S104	48	0 (0,0)	23	3 (13,0)	0,06 [0,00;1,22], p=0,0674
<b>Kopfschmerzen</b>					
S102	348	49 (14,1)	352	34 (9,7)	1,53 [0,96;2,44], p=0,0719
S104	48	7 (14,6)	23	4 (17,4)	0,81 [0,21;3,11], p=0,7599
<sup>a</sup> zusätzlich berechnet KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.b. = nicht berechenbar, n.z. = nicht zutreffend, OR = Odds Ratio					

Tabelle 4-45: Ergebnisse für psychiatrische Erkrankungen (SOC, Grenze:  $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Endpunkt Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
<b>Psychiatrische Erkrankungen (SOC)</b>					
S102	348	118 (33,9)	352	163 (46,3)	0,59 [0,44;0,81], p=0,0009
S104	48	15 (31,3)	23	11 (47,8)	0,50 [0,18;1,38], p=0,1780
<b>Abnorme Träume</b>					
S102	348	53 (15,2)	352	95 (27,0)	0,47 [0,32;0,69], p=0,0001
S104	48	5 (10,4)	23	8 (34,8)	0,22 [0,06;0,77], p=0,0180
<b>Schlaflosigkeit</b>					
S102	348	30 (8,6)	352	49 (13,9)	0,58 [0,36;0,94], p=0,0280
S104	48	5 (10,4)	23	2 (8,7)	1,22 [0,22;6,82], p=0,8202
<b>Depression<sup>b</sup></b>					
S102	348	33 (9,5)	352	39 (11,1)	0,84 [0,52;1,37], p=0,4872
S104	48	4 (8,3)	23	2 (8,7)	0,95 [0,16;5,63], p=0,9590
<b>Angst<sup>c</sup></b>					
S102	348	16 (4,6)	352	14 (4,0)	1,16 [0,56;2,42], p=0,6856
S104	48	2 (4,2)	23	3 (13,0)	0,29 [0,04;1,87], p=0,1930
<sup>a</sup> zusätzlich berechnet <sup>b</sup> In Studie S104 trat Depression nicht bei mehr als 10% der Patienten in einem der Behandlungsarme auf. Aus Vergleichbarkeitsgründen werden die Ereignisse hier dennoch dargestellt. <sup>c</sup> In Studie S102 trat Angst nicht bei mehr als 10% der Patienten aus einem der Behandlungsarme auf. Aus Vergleichbarkeitsgründen werden die Ereignisse hier dennoch dargestellt.  KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, n.z. = nicht zutreffend, OR = Odds Ratio					

Tabelle 4-46: Ergebnisse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC, Grenze:  $\geq 10\%$ ) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 48)

Endpunkt Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
<b>Erkrankungen der Haut und des Unterhautzellgewebes (SOC)</b>					
S102	348	87 (25,0)	352	129 (36,6)	0,58 [0,42;0,80], p=0,0009
S104	48	15 (31,3)	23	8 (34,8)	0,85 [0,30;2,44], p=0,7661
<b>Hautausschlag</b>					
S102	348	22 (6,3)	352	43 (12,2)	0,48 [0,28;0,83], p=0,0082
S104	48	5 (10,4)	23	4 (17,4)	0,55 [0,13;2,29], p=0,4130
<sup>a</sup> zusätzlich berechnet KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

### ***Unerwünschte Ereignisse***

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). 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, 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).

In Studie S102 hatten 11,8% der Patienten im EVG/COBI/FTC/TDF-Arm und 6,8% der Patienten 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). 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).

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). 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],  $p=0,2587$ ).

Unerwünschte Ereignisse Grad 3-4 traten in Studie S102 bei 12,9% der EVG/COBI/FTC/TDF-Patienten und bei 11,1% der EFV/FTC/TDF-Patienten auf. Der Unterschied war nicht statistisch signifikant (OR [95%-KI]: 1,19 [0,75;1,88],  $p=0,4514$ ). In Studie S104 hatten 4,2% der EVG/COBI/FTC/TDF-Patienten und 8,7% der EFV/FTC/TDF-Patienten unerwünschte Ereignisse Grad 3-4. Der Unterschied zwischen den Behandlungen war nicht statistisch signifikant (OR [95%-KI]: 0,46 [0,06;3,46],  $p=0,4483$ ).

### ***Wichtige unerwünschte Ereignisse***

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  $p<0,001$ . 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],  $p=0,2426$ ).

Hautausschläge traten in Studie S102 bei 17,0% der EVG/COBI/FTC/TDF-Patienten und 27,8% der EFV/FTC/TDF-Patienten auf. Es besteht somit ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,53 [0,37;0,76],  $p<0,001$ ). Ein numerischer Vorteil von EVG/COBI/FTC/TDF hinsichtlich des Auftretens von Hautausschlägen besteht auch in Studie S104 (EVG/COBI/FTC/TDF: 18,8%, EFV/FTC/TDF: 26,1%), statistische Signifikanz erreicht dieser Unterschied zwischen den Behandlungen jedoch nicht (OR [95%-KI]: 0,65 [0,20;2,13],  $p=0,4802$ ).

Da für die Endpunkte neurologische & psychiatrische Ereignisse und Hautausschläge das Verzerrungspotenzial in Studie S104 nicht sicher ausgeschlossen werden kann, ist die Aussagekraft dieser Endpunkte für Studie S104 niedriger einzustufen.

### ***Unerwünschte Ereignisse von Interesse***

In Studie S102 traten bei 1,4% der EVG/COBI/FTC/TDF-Patienten renale Ereignisse auf und im EFV/FTC/TDF-Arm waren dies 0,3%. Der Unterschied ist nicht statistisch signifikant (OR [95%-KI]: 5,12 [0,59;44,02],  $p=0,12$ ). In Studie S104 traten keine renalen Ereignisse auf.

In Studie S102 traten Knochenfrakturen bei jeweils 1,7% der Patienten pro Behandlungsarm auf. Der Unterschied zwischen den Behandlungen ist somit nicht statistisch signifikant (OR [95%-KI]: 1,01 [0,32;3,17],  $p=1,00$ ). In Studie S104 traten bei 2,1% der EVG/COBI/FTC/TDF-Patienten Knochenfrakturen auf, während im EFV/FTC/TDF-Arm keine Knochenfraktur beobachtet werden konnte. Der Unterschied zwischen den Behandlungen ist statistisch nicht signifikant (OR [95%-KI]: 1,48 [0,06;37,84],  $p=0,8111$ ).

Da für die Endpunkte renale Ereignisse und Knochenfrakturen das Verzerrungspotenzial in Studie S104 nicht sicher ausgeschlossen werden kann, ist die Aussagekraft dieser Endpunkte für Studie S104 niedriger einzustufen.

### **Häufigste unerwünschte Ereignisse**

*Erkrankungen des Gastrointestinaltrakts (SOC)* traten in Studie S102 bei 54,0% der Patienten im EVG/COBI/FTC/TDF-Arm und 45,7% der Patienten im EFV/FTC/TDF-Arm auf. Es besteht hier somit ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,39 [1,04;1,88], p=0,0286). In Studie S104 traten bei 47,9% der EVG/COBI/FTC/TDF-Patienten und bei 30,4% der EFV/FTC/TDF-Patienten Erkrankungen des Gastrointestinaltrakts auf. Der Unterschied zwischen den Behandlungen ist jedoch nicht statistisch signifikant (OR [95%-KI]: 2,10 [0,73;6,03], p=0,1667).

23,0% der EVG/COBI/FTC/TDF-Patienten in Studie S102 hatten Diarrhoe und im EFV/FTC/TDF-Arm waren dies 18,8% der Patienten. Der Unterschied zwischen den Behandlungen ist statistisch nicht signifikant (OR [95%-KI]: 1,29 [0,90;1,87], p=0,1682). 20,8% der EVG/COBI/FTC/TDF-Patienten und 17,4% der EFV/FTC/TDF-Patienten in Studie S104 hatten Diarrhoe. Der Unterschied zwischen den Behandlungen ist nicht statistisch signifikant (OR [95%-KI]: 1,25 [0,35;4,51], p=0,7333).

Übelkeit trat bei 20,7% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 13,6% der EFV/FTC/TDF-Patienten in Studie S102 auf. Es besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,65 [1,11;2,46], p=0,0139). In Studie S104 trat bei 8,3% der EVG/COBI/FTC/TDF-Patienten und 8,7% der EFV/FTC/TDF-Patienten Übelkeit auf. Der Unterschied ist nicht statistisch signifikant (OR [95%-KI]: 0,95 [0,16;5,63], p=0,9590).

*Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)* traten in Studie S102 bei 19,5% der EVG/COBI/FTC/TDF-Patienten und 28,1% der EFV/FTC/TDF-Patienten auf. Es besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,62 [0,44;0,88], p=0,0080). In Studie S104 traten allgemeine Erkrankungen und Beschwerden am Verabreichungsort bei 27,1% der EVG/COBI/FTC/TDF-Patienten und bei 30,4% der EFV/FTC/TDF-Patienten auf. Der Unterschied zwischen den Behandlungen ist nicht statistisch signifikant (OR [95%-KI]: 0,85 [0,28;2,53], p=0,7690).

Fatigue trat in Studie S102 bei 11,5% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 12,8% der Patienten im EFV/FTC/TDF-Arm auf. Es besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,89 [0,56;1,40], p=0,6015). In Studie S104 hatten 16,7% der Patienten im EVG/COBI/FTC/TDF-Arm und 17,4% der Patienten im EFV/FTC/TDF-Arm Fatigue. Auch hier ist der Unterschied zwischen den Behandlungen statistisch nicht signifikant (OR [95%-KI]: 0,95 [0,25;3,55], p=0,9392).

*Erkrankungen des Nervensystems (SOC)* traten in Studie S102 bei 27,3% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 41,5% der Patienten im EFV/FTC/TDF-Arm auf. Es besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,53 [0,39;0,73],  $p=0,0001$ ). In Studie S104 traten Erkrankungen des Nervensystems bei 31,3% der EVG/COBI/FTC/TDF-Patienten und 34,8% der EFV/FTC/TDF-Patienten auf. Mit einem OR [95%-KI] von 0,85 [0,30;2,44] und  $p=0,7661$  ist dieser Unterschied statistisch nicht signifikant.

Bei 6,6% der EVG/COBI/FTC/TDF-Patienten und 24,4% der EFV/FTC/TDF-Patienten in Studie S102 trat Schwindel auf. Das OR [95%-KI] ergibt sich zu 0,22 [0,13;0,36] mit  $p<0,00001$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF anzeigt. Auch in Studie S104 besteht ein numerischer Vorteil von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF, da im EVG/COBI/FTC/TDF-Arm kein Patient Schwindel aufwies, während er im EFV/FTC/TDF bei 13,0% der Patienten auftrat. Das OR [95%-KI] von 0,06 [0,00;1,22] mit  $p=0,0674$  zeigt allerdings keinen statistisch signifikanten Unterschied.

Kopfschmerzen traten in Studie S102 bei 14,1% der EVG/COBI/FTC/TDF-Patienten und 9,7% der EFV/FTC/TDF-Patienten auf. Der Unterschied zwischen den Behandlungen ist statistisch nicht signifikant (OR [95%-KI]: 1,53 [0,96;2,44],  $p=0,0719$ ). In Studie S104 traten Kopfschmerzen bei 14,6% der EVG/COBI/FTC/TDF-Patienten und 17,4% der EFV/FTC/TDF-Patienten auf. Auch hier ist der Unterschied zwischen den Behandlungen nicht statistisch signifikant (OR [95%-KI]: 0,81 [0,21;3,11],  $p=0,7599$ ).

*Psychiatrische Erkrankungen (SOC)* traten in Studie S102 bei 33,9% der EVG/COBI/FTC/TDF-Patienten und bei 46,3% der EFV/FTC/TDF-Patienten auf. Es besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,59 [0,44;0,81],  $p=0,0009$ ). In Studie S104 besteht ein numerischer Vorteil von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF (EVG/COBI/FTC/TDF: 31,3%, EFV/FTC/TDF: 47,8%) hinsichtlich der Vermeidung psychiatrischer Erkrankungen. Allerdings erreicht dieser Unterschied hier keine statistische Signifikanz (OR [95%-KI]: 0,50 [0,18;1,38],  $p=0,1780$ ).

Abnorme Träume traten in Studie S102 bei 15,2% der EVG/COBI/FTC/TDF-Patienten und 27,0% der EFV/FTC/TDF-Patienten auf. Es besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,47 [0,32;0,69],  $p=0,0001$ ). Auch in Studie S104 besteht hinsichtlich der Vermeidung abnormer Träume ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,22 [0,06;0,77],  $p=0,0180$ ).

Schlaflosigkeit trat in Studie S102 bei 8,6% der EVG/COBI/FTC/TDF-Patienten und 13,9% der EFV/FTC/TDF-Patienten auf. Hier besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,58 [0,36;0,94],  $p=0,0280$ ). In Studie S104 trat Schlaflosigkeit bei 10,4% der Patienten im



EVG/COBI/FTC/TDF-Arm und bei 8,7% der Patienten im EFV/FTC/TDF-Arm auf. Der Unterschied zwischen den Behandlungen ist statistisch nicht signifikant (OR [95%-KI]: 1,22 [0,22;6,82],  $p=0,8202$ ).

Depressionen traten in Studie S102 bei 9,5% der EVG/COBI/FTC/TDF-Patienten und 11,1% der EFV/FTC/TDF-Patienten auf. Es besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,84 [0,52;1,37],  $p=0,4872$ ). In Studie S104 traten Depressionen bei 8,3% der EVG/COBI/FTC/TDF-Patienten und 8,7% der EFV/FTC/TDF-Patienten auf. Der Unterschied ist ebenfalls nicht statistisch signifikant (OR [95%-KI]: 0,95 [0,16;5,63],  $p=0,9590$ ). Angst trat in Studie S102 bei 4,6% der EVG/COBI/FTC/TDF-Patienten und 4,0% der EFV/FTC/TDF-Patienten auf. Ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht nicht (OR [95%-KI]: 1,16 [0,56;2,42],  $p=0,6856$ ). In Studie S104 trat bei 4,2% der EVG/COBI/FTC/TDF-Patienten und 13,0% der EFV/FTC/TDF-Patienten Angst auf. Ein statistisch signifikanter Unterschied besteht zwischen den Behandlungen jedoch nicht (OR [95%-KI]: 0,29 [0,04;1,87],  $p=0,1930$ ).

*Erkrankungen der Haut und des Unterhautzellgewebes (SOC)* traten in Studie S102 bei 25,0% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 36,6% der Patienten im EFV/FTC/TDF-Arm auf. Dieser Unterschied zwischen den Behandlungen ist statistisch signifikant zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,58 [0,42;0,80],  $p=0,0009$ ). In Studie S104 traten Erkrankungen der Haut und des Unterhautzellgewebes bei 31,3% der EVG/COBI/FTC/TDF-Patienten und 34,8% der EFV/FTC/TDF-Patienten auf. Statistisch signifikant ist dieser Unterschied jedoch nicht (OR [95%-KI]: 0,85 [0,30;2,44],  $p=0,7661$ ). Hautausschlag trat in Studie S102 bei 6,3% der EVG/COBI/FTC/TDF-Patienten und 12,2% der EFV/FTC/TDF-Patienten auf. Der Unterschied zwischen den Behandlungen ist statistisch signifikant zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,48 [0,28;0,83],  $p=0,0082$ ). Auch in Studie S104 besteht hinsichtlich der Vermeidung von Hautausschlag ein numerischer Vorteil zugunsten von EVG/COBI/FTC/TDF. Statistische Signifikanz erreicht dieser Unterschied jedoch nicht (OR [95%-KI]: 0,55 [0,13;2,29],  $p=0,4130$ ).

**Woche 60 (nur Studie 60)**

Tabelle 4-47: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 60)

Studie Endpunkt	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
S104					
Mortalität	48	0 (0,0)	23	0 (0,0)	n.b.
Gesamtrate unerwünschter Ereignisse	48	44 (91,7)	23	21 (91,3)	1,05 [0,18;6,18], p=0,9590
Schwerwiegende unerwünschte Ereignisse	48	2 (4,2)	23	1 (4,3)	0,96 [0,08;11,12], p=0,9717
Therapieabbrüche aufgrund unerwünschter Ereignisse	48	0 (0,0)	23	1 (4,3)	0,15 [0,01;3,95], p=0,2587
Unerwünschte Ereignisse Grad 3-4	48	2 (4,2)	23	2 (8,7)	0,46 [0,06;3,46], p=0,4483
<sup>a</sup> 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 4-48: Ergebnisse für wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 60)

Studie Endpunkt	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
S104					
Neurologische & psychiatrische Ereignisse	48	20 (41,7)	23	13 (56,5)	0,55 [0,20;1,50], p=0,31
Hautausschläge	48	9 (18,8)	23	6 (26,1)	0,65 [0,20;2,13], p=0,54
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Fisher's exakter Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Tabelle 4-49: Ergebnisse für unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 60)

Studie Endpunkt	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
S104					
Renale Ereignisse	48	0 (0,0)	23	0 (0,0)	n.b.
Knochenfrakturen	48	1 (2,1)	23	0 (0,0)	1,48 [0,06;37,84], p=0,8111

<sup>a</sup> 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 4-50: Ergebnisse für die häufigsten unerwünschten Ereignisse (SOC, Grenze: ≥10%) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 60)

Studie Endpunkt (SOC) häufigster PT	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
S104					
Erkrankungen des Gastrointestinaltrakts (SOC)	48	26 (54,2)	23	7 (30,4)	2,70 [0,94;7,75], p=0,0647
<i>Diarrhoe</i>	48	13 (27,1)	23	4 (17,4)	1,76 [0,50;6,17], p=0,3742
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)	48	12 (25,0)	23	7 (30,4)	0,76 [0,25;2,29], p=0,6288
<i>Fatigue</i>	48	8 (16,7)	23	4 (17,4)	0,95 [0,25;3,55], p=0,9392
Erkrankungen des Nervensystems (SOC)	48	15 (31,3)	23	8 (34,8)	0,85 [0,30;2,44], p=0,7661
<i>Kopfschmerzen</i>	48	7 (14,6)	23	4 (17,4)	0,81 [0,21;3,11], p=0,7599
<i>Schwindel</i>	48	0 (0,0)	23	3 (13,0)	0,06 [0,00;1,22], p=0,0674
Psychiatrische Erkrankungen (SOC)	48	16 (33,3)	23	11 (47,8)	0,55 [0,20;1,50], p=0,2417
<i>Schlaflosigkeit</i>	48	6 (12,5)	23	2 (8,7)	1,50 [0,28;8,08], p=0,6370
<i>Abnorme Träume</i>	48	5 (10,4)	23	8 (34,8)	0,22 [0,06;0,77], p=0,0180
<i>Angst</i>	48	2 (4,2)	23	3 (13,0)	0,29 [0,04;1,87], p=0,1930
Erkrankungen der Haut und des Unterhautzellgewebes (SOC)	48	15 (31,3)	23	8 (34,8)	0,85 [0,30;2,44], p=0,7661
<i>Hautausschlag</i>	48	5 (10,4)	23	4 (17,4)	0,55 [0,13;2,29], p=0,4130

<sup>a</sup> zusätzlich berechnet  
KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, PT = Preferred term

### ***Unerwünschte Ereignisse***

In Studie S104 traten bis Woche 60 keine Todesfälle unter den behandelten Patienten auf. Die Gesamtrate unerwünschter Ereignisse war mit 91,7% der Patienten im EVG/COBI/FTC/TDF-Arm und 91,3% der Patienten im EFV/FTC/TDF-Arm in beiden Behandlungsarmen nahezu identisch. Entsprechend beträgt das OR [95%-KI] 1,05 [0,18;6,18] mit  $p=0,9590$  und zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen. Auch hinsichtlich des Auftretens schwerwiegender unerwünschter Ereignisse kann kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen gefunden werden (OR [95%-KI]: 0,96 [0,08;11,12],  $p=0,9717$ ). Im EVG/COBI/FTC/TDF-Arm brach kein Patient die Behandlung aufgrund unerwünschter Ereignisse vorzeitig ab und im EFV/FTC/TDF-Arm brach 1 (4,3%) Patient die Behandlung aufgrund unerwünschter Ereignisse ab. Damit beträgt das OR [95%-KI] 0,15 [0,01;3,95] mit  $p=0,2587$ , was keinen statistisch signifikanten Unterschied zeigt. Jeweils 2 Patienten pro Behandlungsarm (EVG/COBI/FTC/TDF: 4,2%, EFV/FTC/TDF: 8,7%) hatten unerwünschte Ereignisse Grad 3-4 (OR [95%-KI]: 0,46 [0,06;3,46],  $p=0,4483$ ). Somit besteht auch hier kein statistisch signifikanter Unterschied zwischen den Behandlungen.

### ***Wichtige unerwünschte Ereignisse***

Neurologische & psychiatrische Ereignisse traten in Studie S104 im EVG/COBI/FTC/TDF-Arm numerisch seltener auf. Hier hatten 20 (41,7%) Patienten ein neurologisches bzw. psychiatrisches Ereignis, wohingegen im EFV/FTC/TDF-Arm deutlich mehr als die Hälfte der Patienten (13 Patienten, 56,5%) ein unerwünschtes Ereignis neurologischer bzw. psychiatrischer Art hatten. Das OR [95%-KI] beträgt 0,55 [0,20;1,50] mit  $p=0,31$  und zeigt keinen statistisch signifikanten Unterschied zwischen den Behandlungen. Auch Hautausschläge traten im EVG/COBI/FTC/TDF-Arm (9 Patienten, 18,8%) seltener auf als im EFV/FTC/TDF-Arm (6 Patienten, 26,1%; OR [95%-KI]: 0,65 [0,20;2,13],  $p=0,54$ ). Allerdings erreicht auch hier der Unterschied zwischen den Behandlungen keine statistische Signifikanz. Insgesamt sei an dieser Stelle an das generell bestehende Problem der geringen Patientenzahl und der damit einhergehenden niedrigen Power der Studie erinnert. Da für diesen Endpunkt das Verzerrungspotenzial in Studie S104 nicht sicher ausgeschlossen werden kann, ist die Aussagekraft dieser Endpunkte für Studie S104 niedriger einzustufen.

### ***Unerwünschte Ereignisse von Interesse***

In keinem der beiden Behandlungsarme wurden renale Ereignisse beobachtet und lediglich ein Patient (2,1%) im EVG/COBI/FTC/TDF-Arm wies eine Knochenfraktur auf. Im EFV/FTC/TDF-Arm traten keine Knochenfrakturen auf (OR [95%-KI]: 1,48 [0,06;37,84],  $p=0,8111$ ). Da für diesen Endpunkt das Verzerrungspotenzial in Studie S104 nicht sicher ausgeschlossen werden kann, ist die Aussagekraft dieser Endpunkte für Studie S104 niedriger einzustufen.

### **Häufigste unerwünschte Ereignisse**

*Erkrankungen des Gastrointestinaltrakts (SOC)* traten bei 26 (54,2%) Patienten im EVG/COBI/FTC/TDF-Arm und bei 7 (30,4%) Patienten im EFV/FTC/TDF-Arm auf. Mit einem OR [95%-KI] von 2,70 [0,94;7,75] mit  $p=0,0647$  kann hier allerdings kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen festgestellt werden. Die am häufigsten auftretende Erkrankung des Gastrointestinaltrakts war Diarrhoe, was bei 13 (27,1%) Patienten im EVG/COBI/FTC/TDF-Arm und bei 4 (17,4%) Patienten im EFV/FTC/TDF-Arm beobachtet wurde. Allerdings besteht auch hier kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,76 [0,50;6,17],  $p=0,3742$ ).

*Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)* traten bei 12 (25,0%) Patienten im EVG/COBI/FTC/TDF-Arm und bei 7 (30,4%) Patienten im EFV/FTC/TDF-Arm auf (OR [95%-KI]: 0,76 [0,25;2,29],  $p=0,6288$ ). Es besteht somit kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen. Die am häufigsten beobachtete allgemeine Erkrankung und Beschwerde am Verabreichungsort war Fatigue. Diese wurden bei 8 (16,7%) Patienten im EVG/COBI/FTC/TDF-Arm und bei 4 (17,4%) Patienten im EFV/FTC/TDF-Arm beobachtet. Damit beträgt das OR [95%-KI] 0,95 [0,25;3,55] mit  $p=0,9392$ , was keinen statistisch signifikanten Unterschied anzeigt.

*Erkrankungen des Nervensystems (SOC)* wurden bei 31,3% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 34,8% der Patienten im EFV/FTC/TDF-Arm beobachtet. Es besteht somit kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,85 [0,30;2,44],  $p=0,7661$ ). Unter Erkrankungen des Nervensystems wurden am häufigsten Fälle von Kopfschmerzen und Schwindel beobachtet. Im EVG/COBI/FTC/TDF-Arm zeigten 14,6% der Patienten Kopfschmerzen und im EFV/FTC/TDF-Arm waren dies 17,4%, was keinen statistisch signifikanten Unterschied anzeigt (OR [95%-KI]: 0,81 [0,21;3,11],  $p=0,7599$ ). Auch hinsichtlich Schwindel kann trotz des numerischen Unterschiedes kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (EVG/COBI/FTC/TDF: 0,0%, EFV/FTC/TDF: 13,0%) beobachtet werden (OR [95%-KI]: 0,06 [0,00;1,22],  $p=0,0674$ ).

*Psychiatrische Erkrankungen (SOC)* traten im EVG/COBI/FTC/TDF-Arm bei 33,3% der Patienten und im EFV/FTC/TDF-Arm bei 47,8% der Patienten auf. Ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen kann allerdings nicht festgestellt werden (OR [95%-KI]: 0,55 [0,20;1,50],  $p=0,2417$ ). Die darunter am häufigsten betrachteten Ereignisse waren Schlaflosigkeit (EVG/COBI/FTC/TDF: 12,5%; EFV/FTC/TDF: 8,7%), abnorme Träume (EVG/COBI/FTC/TDF: 10,4%; EFV/FTC/TDF: 34,8%) und Angst (EVG/COBI/FTC/TDF: 4,2%; EFV/FTC/TDF: 13,0%). Hinsichtlich Schlaflosigkeit (OR [95%-KI]: 1,50 [0,28;8,08],  $p=0,6370$ ) und Angst (OR [95%-KI]: 0,29 [0,04;1,87],  $p=0,1930$ ) können keine statistisch signifikanten Unterschiede gefunden werden. Allerdings zeigte sich für abnorme Träume ein statistisch signifikanter Unterschied zwischen

den beiden Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,22 [0,06;0,77], p=0,0180).

*Erkrankungen der Haut und des Hautunterzellgewebes (SOC)* wurden bei 31,3% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 34,8% der Patienten im EFV/FTC/TDF-Arm beobachtet. Es besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,85 [0,30;2,44], p=0,7661). Die darunter am häufigsten beobachtete Hauterkrankung war Hautausschlag. Dieser trat numerisch seltener in der EVG/COBI/FTC/TDF-Gruppe (10,4%) als in der EFV/FTC/TDF-Gruppe (17,4%) auf. Allerdings kann kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen festgestellt werden (OR [95%-KI]: 0,55 [0,13;2,29], p=0,4130).

### Woche 96 (nur Studie S102)

Tabelle 4-51: Ergebnisse für unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)

Studie Endpunkt	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
S102					
Mortalität	348	1 (0,3)	352	2 (0,6)	0,50 [0,05;5,59], p=0,5769
Gesamtrate unerwünschter Ereignisse	348	337 (96,8)	352	342 (97,2)	0,90 [0,38;2,14], p=0,8041
Schwerwiegende unerwünschte Ereignisse	348	56 (16,1)	352	33 (9,4)	1,85 [1,17;2,93], p=0,0083
Therapieabbrüche aufgrund unerwünschter Ereignisse	348	17 (4,9)	352	24 (6,8)	0,70 [0,37;1,33], p=0,2782
Unerwünschte Ereignisse Grad 3-4	348	61 (17,5)	352	51 (14,5)	1,25 [0,84;1,88], p=0,2733
<sup>a</sup> zusätzlich berechnet KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Tabelle 4-52: Ergebnisse für wichtige unerwünschte Ereignisse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)

Studie Endpunkt	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
S102					
Neurologische & psychiatrische Ereignisse	348	162 (46,6)	352	232 (65,9)	0,45 [0,33;0,61], p<0,001
Hautausschläge	348	74 (21,3)	352	108 (30,7)	0,61 [0,43;0,86], p=0,006
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Fisher's exakter Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Tabelle 4-53: Ergebnisse für unerwünschte Ereignisse von Interesse aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)

Studie Endpunkt	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
S102					
Renale Ereignisse	348	7 (2,0)	352	1 (0,3)	7,21 [0,88;58,87], p=0,037 Peto Odds Ratio <sup>c</sup> : 4,60 [1,14;18,54]
Knochenfrakturen	348	10 (2,9)	352	8 (2,3)	1,27 [0,50;3,26], p=0,64
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Fisher's exakter Test <sup>c</sup> aufgrund inkonsistenter Ergebnisse zwischen zusätzlich berechnetem Odds Ratio mit 95%-KI (Normalapproximation) und p-Wert aus exaktem Fisher Test, wird hier zusätzlich das Peto Odds Ratio berechnet. KI = Konfidenzintervall, N: Anzahl der Patienten in der Analyse, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

Tabelle 4-54: Ergebnisse für die häufigsten unerwünschten Ereignisse (SOC, Grenze: ≥10%) aus RCT mit dem zu bewertenden Arzneimittel (Safety Analyse-Set, Woche 96)

Studie Endpunkt (SOC) Preferred term	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
S102					
Erkrankungen des Gastrointestinaltrakts (SOC)	348	211 (60,6)	352	188 (53,4)	1,34 [1,00;1,81], p=0,0538
<i>Diarrhoe</i>	348	87 (25,0)	352	83 (23,6)	1,08 [0,76;1,53], p=0,6613
<i>Übelkeit</i>	348	76 (21,8)	352	53 (15,1)	1,58 [1,07;2,32], p=0,0213
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)	348	85 (24,4)	352	109 (31,0)	0,72 [0,52;1,01], p=0,0537
<i>Fatigue</i>	348	45 (12,9)	352	52 (14,8)	0,86 [0,56;1,32], p=0,4810
Erkrankungen des Nervensystems (SOC)	348	112 (32,2)	352	159 (45,2)	0,58 [0,42;0,78], p=0,0004
<i>Schwindel</i>	348	25 (7,2)	352	91 (25,9)	0,22 [0,14;0,36], p<0,00001
<i>Kopfschmerzen</i>	348	56 (16,1)	352	40 (11,4)	1,50 [0,97;2,31], p=0,0702
Psychiatrische Erkrankungen (SOC)	348	138 (39,7)	352	179 (50,9)	0,64 [0,47;0,86], p=0,0030
<i>Abnorme Träume</i>	348	53 (15,2)	352	97 (27,6)	0,47 [0,32;0,69], p=0,0001
<i>Schlaflosigkeit</i>	348	38 (10,9)	352	57 (16,2)	0,63 [0,41;0,99], p=0,0428
<i>Depression</i>	348	42 (12,1)	352	49 (13,9)	0,85 [0,55;1,32], p=0,4668

Studie Endpunkt (SOC) Preferred term	EVG/COBI/FTC/TDF		EFV/FTC/TDF		Behandlungseffekt
	N	n (%)	N	n (%)	OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a</sup>
Erkrankungen der Haut und des Unterhautzellgewebes (SOC)	348	111 (31,9)	352	147 (41,8)	0,65 [0,48;0,89], p=0,0070
<i>Hautausschlag</i>	348	26 (7,5)	352	49 (13,9)	0,50 [0,30;0,82], p=0,0066

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

### ***Unerwünschte Ereignisse***

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). 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). 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). 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) und 17,5% der Patienten im EVG/COBI/FTC/TDF-Arm und 14,5% der Patienten im EFV/FTC/TDF-Arm erlebten unerwünschte Ereignisse Grad 3-4 (OR [95%-KI]: 1,25 [0,84;1,88], p=0,2733).

### ***Wichtige unerwünschte Ereignisse***

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 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). Auch hinsichtlich des Auftretens von Hautausschlägen besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,61 [0,43;0,86], p=0,006). Mit 21,3% der Patienten im EVG/COBI/FTC/TDF-Arm und 30,7% im EFV/FTC/TDF-Arm traten Hautausschläge im EVG/COBI/FTC/TDF-Arm deutlich seltener auf.

### ***Unerwünschte Ereignisse von Interesse***

Renale Ereignisse traten bei 2,0% der Patienten im EVG/COBI/FTC/TDF-Arm und bei 0,3% der Patienten im EFV/FTC/TDF-Arm auf. Hier besteht eine Inkonsistenz zwischen dem Konfidenzintervall, welches mittels Normalapproximation berechnet wird und keinen



signifikanten Unterschied anzeigt, und dem p-Wert der mittels Fisher's exaktem Test berechnet wird und signifikant unter 0,05 liegt. Aus diesem Grund wird ein Peto Odds Ratio [95%-KI] berechnet, welches 4,60 [1,14;18,54] beträgt. Der p-Wert aus dem exaktem Test nach Fisher beträgt  $p=0,037$ . Somit besteht hier ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF. 2,9% der Patienten im EVG/COBI/FTC/TDF-Arm und 2,3% der Patienten im EFV/FTC/TDF-Arm erlebten Knochenfrakturen. Dabei kann allerdings kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen festgestellt werden (OR [95%-KI]: 1,27 [0,50;3,26],  $p=0,64$ ).

### **Häufigste unerwünschte Ereignisse**

*Erkrankungen des Gastrointestinaltrakts (SOC)* traten bei 60,6% der Patienten im EVG/COBI/FTC/TDF-Arm und 53,4% der Patienten im EFV/FTC/TDF-Arm auf. Ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen kann dabei nicht festgestellt werden (OR [95%-KI]: 1,34 [1,00;1,81],  $p=0,0538$ ). Die darunter 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],  $p=0,6613$ ) 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],  $p=0,0213$ ).

*Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)* traten bei weniger Patienten im EVG/COBI/FTC/TDF-Arm (24,4%) auf im EFV/FTC/TDF-Arm (31,0%). Allerdings ist dieser Unterschied nicht statistisch signifikant (OR [95%-KI]: 0,72 [0,52;1,01],  $p=0,0537$ ). Am häufigsten beobachtet wurde in dem Zusammenhang Fatigue (EVG/COBI/FTC/TDF: 12,9%; EFV/FTC/TDF: 14,8%; OR [95%-KI]: 0,86 [0,56;1,32],  $p=0,4810$ ).

Bei 32,2% der Patienten im EVG/COBI/FTC/TDF-Arm und 45,2% der Patienten im EFV/FTC/TDF-Arm wurden *Erkrankungen des Nervensystems (SOC)* beobachtet. Diese traten statistisch signifikant weniger häufig im EVG/COBI/FTC/TDF-Arm auf als im EFV/FTC/TDF-Arm (OR [95%-KI]: 0,58 [0,42;0,78],  $p=0,0004$ ). Die häufigsten darunter betrachteten Ereignisse 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],  $p=0,0702$ ). 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],  $p<0,00001$ ).

Bei 39,7% der EVG/COBI/FTC/TDF-Patienten und 50,9% der EFV/FTC/TDF-Patienten wurden *psychiatrische Erkrankungen (SOC)* beobachtet. Dieser Unterschied zwischen den Behandlungen ist statistisch signifikant zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,64 [0,47;0,86],  $p=0,0030$ ). Die häufigsten darunter beobachteten Ereignisse 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). Hinsichtlich abnormer Träume (OR [95%-KI]: 0,47 [0,32;0,69], p=0,0001) und Schlaflosigkeit (OR [95%-KI]: 0,63 [0,41;0,99], p=0,0428) besteht dabei jeweils ein statistisch signifikanter Vorteil der EVG/COBI/FTC/TDF-Behandlung gegenüber der EFV/FTC/TDF-Behandlung.

*Erkrankungen der Haut und des Unterhautzellgewebes (SOC)* traten statistisch signifikant seltener im EVG/COBI/FTC/TDF-Arm auf als im EFV/FTC/TDF-Arm (EVG/COBI/FTC/TDF: 31,9%; EFV/FTC/TDF: 41,8%; OR [95%-KI]: 0,65 [0,48;0,89], p=0,0070). Dabei war Hautausschlag das darunter 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).

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

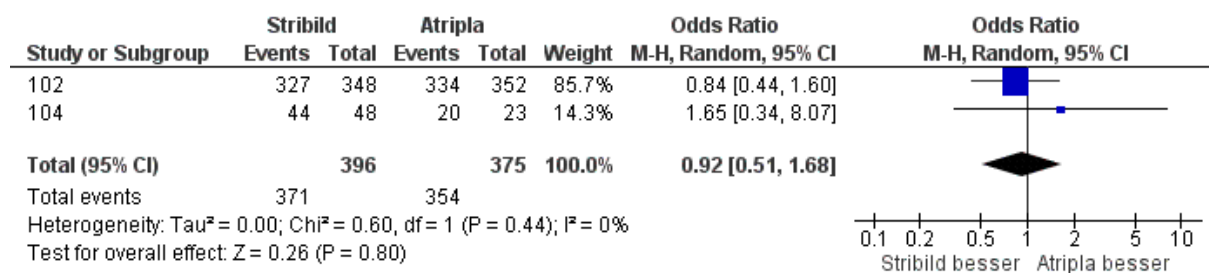


Abbildung 12: Meta-Analyse für die Gesamtrate unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zur **Gesamtrate unerwünschter Ereignisse** zeigt - bei homogener Datenlage (p=0,44; I<sup>2</sup>=0%) - keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungen. Das OR [95%-KI] beträgt 0,92 [0,51;1,68] mit p=0,80. Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,92 [0,51;1,68], p=0,79).

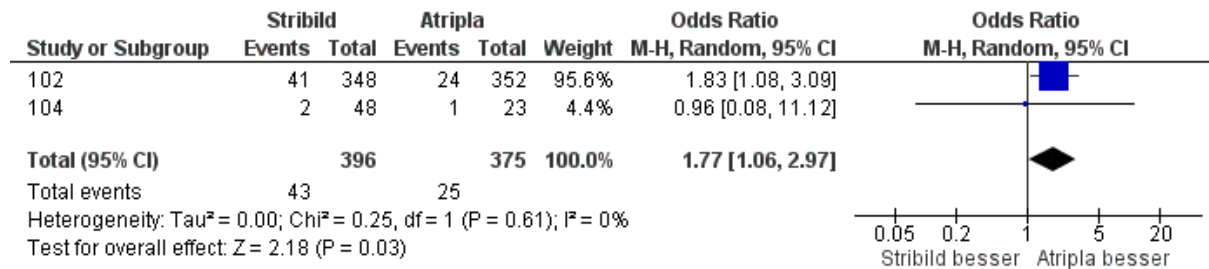


Abbildung 13: Meta-Analyse für die Gesamtrate schwerwiegender unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse der **schwerwiegenden unerwünschten Ereignisse** zeigt einen statistisch signifikanten Unterschied zwischen den beiden Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,77 [1,06;2,97], p=0,03). Dabei ist die Datenlage homogen (p=0,61; I<sup>2</sup>=0%) und eine Sensitivitätsanalyse unter Verwendung des FEM zeigt robuste Ergebnisse (OR [95%-KI]: 1,77 [1,06;2,97], p=0,03).

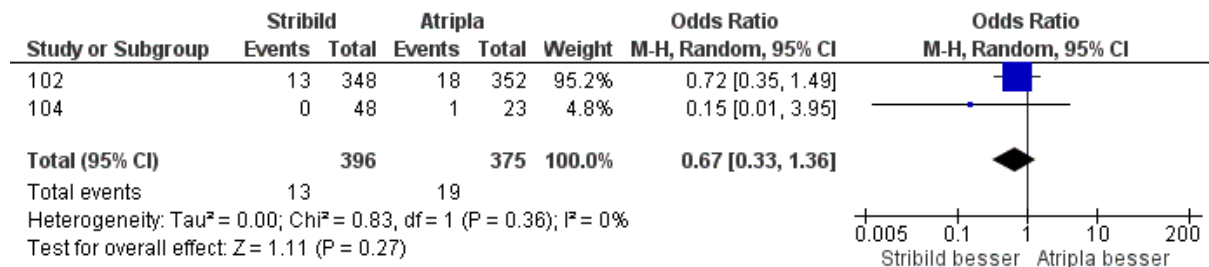


Abbildung 14: Meta-Analyse für Therapieabbrüche aufgrund unerwünschter Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

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). Es besteht keine bedeutsame Heterogenität (p=0,36; I<sup>2</sup>=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).

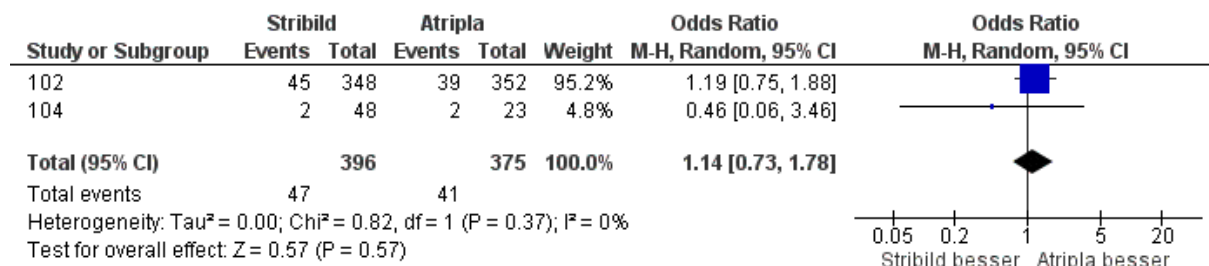


Abbildung 15: Meta-Analyse für unerwünschte Ereignisse Grad 3-4 aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

In der Meta-Analyse zu **unerwünschten Ereignissen Grad 3-4** zeigt sich kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,14 [0,73;1,78],  $p=0,57$ ). Die Datenlage ist homogen ( $p=0,37$ ;  $I^2=0\%$ ) und eine Sensitivitätsanalyse mittels eines FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 1,14 [0,73;1,78],  $p=0,57$ ).

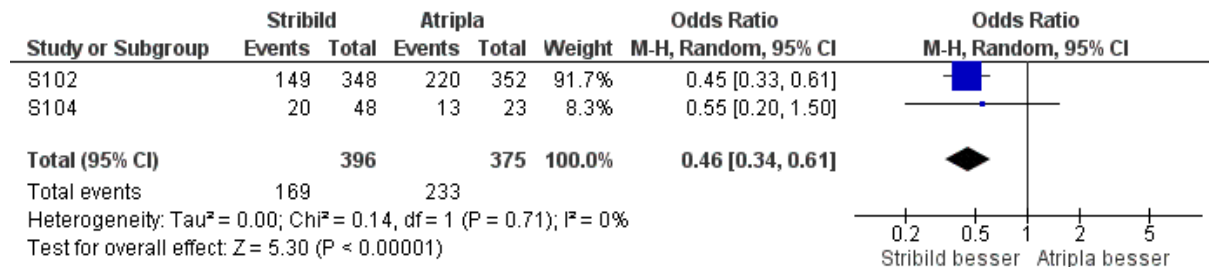


Abbildung 16: Meta-Analyse für neurologische & psychiatrische Ereignisse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

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],  $p<0,00001$ ). 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 ( $p=0,71$ ;  $I^2=0\%$ ) und eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,46 [0,34;0,61],  $p<0,00001$ ).

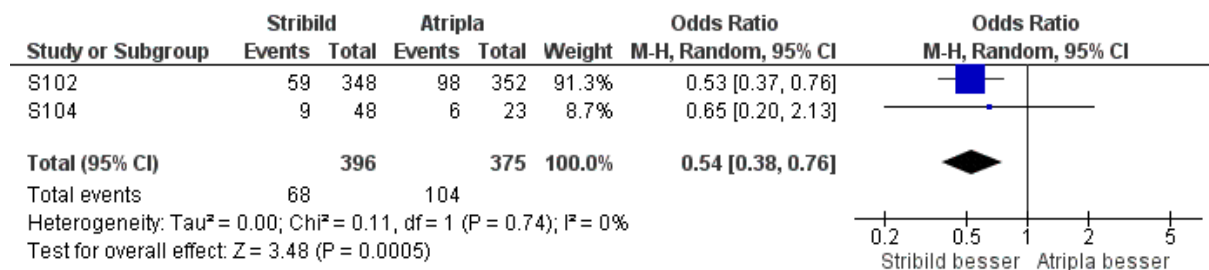


Abbildung 17: Meta-Analyse für Hautausschläge aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt, dass **Hautausschläge** unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auftreten als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,54 [0,38;0,76],  $p=0,0005$ ). Die Datenlage ist dabei homogen ( $p=0,74$ ;  $I^2=0\%$ ) und eine Sensitivitätsanalyse unter Verwendung des FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,54 [0,38;0,76],  $p=0,0005$ ).

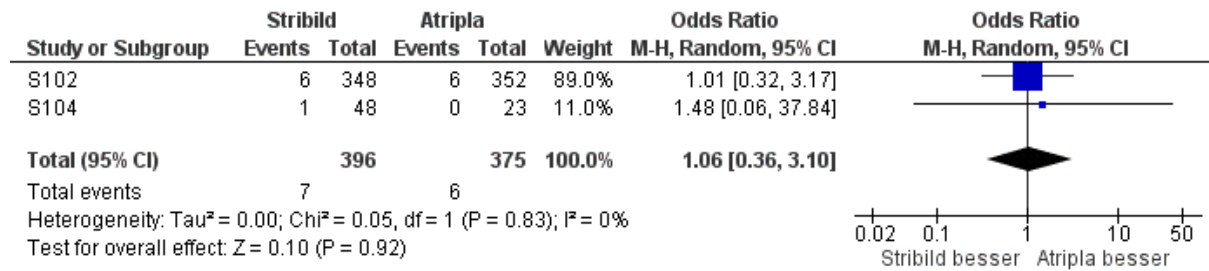


Abbildung 18: Meta-Analyse für Knochenfrakturen aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse der **Knochenfrakturen** zeigt - bei homogener Datenlage ( $p=0,83$ ;  $I^2=0\%$ ) - keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,06 [0,36;3,10],  $p=0,92$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 1,06 [0,36;3,09],  $p=0,92$ ).

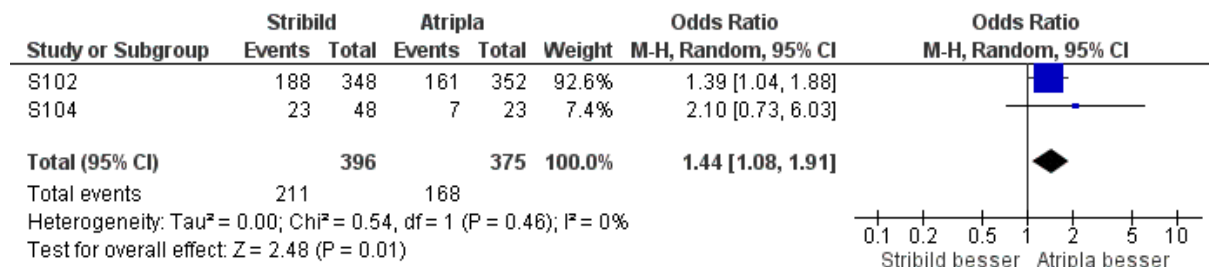


Abbildung 19: Meta-Analyse für Erkrankungen des Gastrointestinaltrakts (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Für **Erkrankungen des Gastrointestinaltrakts (SOC)** zeigt die Meta-Analyse einen statistisch signifikanten Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,44 [1,08;1,91],  $p=0,01$ ). Die Datenlage ist homogen ( $p=0,46$ ;  $I^2=0\%$ ) und eine Sensitivitätsanalyse zeigt robuste Ergebnisse (OR [95%-KI]: 1,44 [1,08;1,91],  $p=0,01$ ).

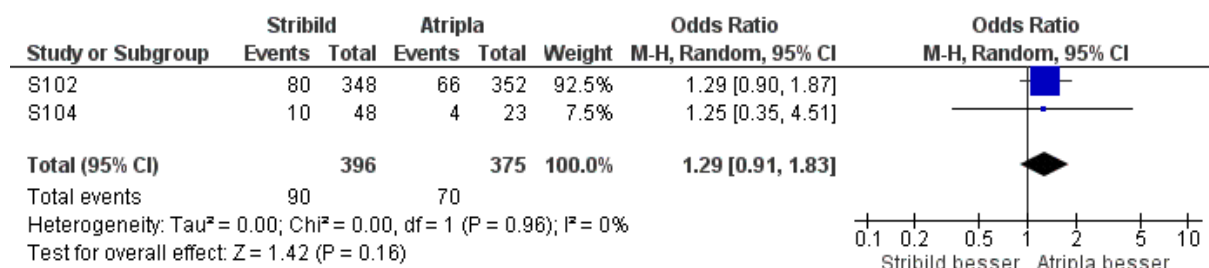


Abbildung 20: Meta-Analyse für Diarrhoe aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt - bei homogener Datenlage ( $p=0,96$ ;  $I^2=0\%$ ) -, dass hinsichtlich des Auftretens von **Diarrhoe** kein statistisch signifikanter Unterschied zwischen den beiden

Behandlungen besteht (OR [95%-KI]: 1,29 [0,91;1,83],  $p=0,16$ ). Eine Sensitivitätsanalyse unter Verwendung des FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 1,29 [0,91;1,83],  $p=0,16$ ).

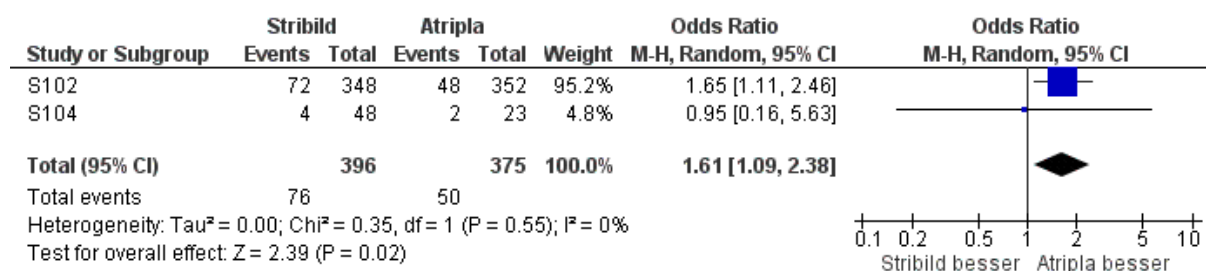


Abbildung 21: Meta-Analyse für Übelkeit aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**Übelkeit** trat laut Meta-Analyse zu Woche 48 statistisch signifikant häufiger im EVG/COBI/FTC/TDF-Arm auf als im EFV/FTC/TDF-Arm (OR [95%-KI]: 1,61 [1,09;2,38],  $p=0,02$ ). Die Datenlage ist dabei homogen ( $p=0,55$ ;  $I^2=0\%$ ) und eine Sensitivitätsanalyse unter Verwendung des FEM zeigt robuste Ergebnisse (OR [95%-KI]: 1,61 [1,09;2,38],  $p=0,02$ ).

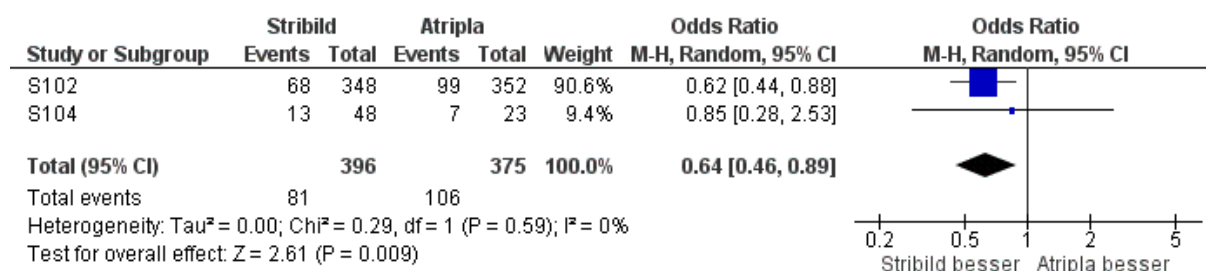


Abbildung 22: Meta-Analyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt, dass **allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)** statistisch signifikant seltener unter einer Behandlung mit EVG/COBI/FTC/TDF auftraten als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,64 [0,46;0,89],  $p=0,009$ ). Die Datenlage ist dabei homogen und eine Sensitivitätsanalyse unter Verwendung des FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,64 [0,46;0,89],  $p=0,009$ ).

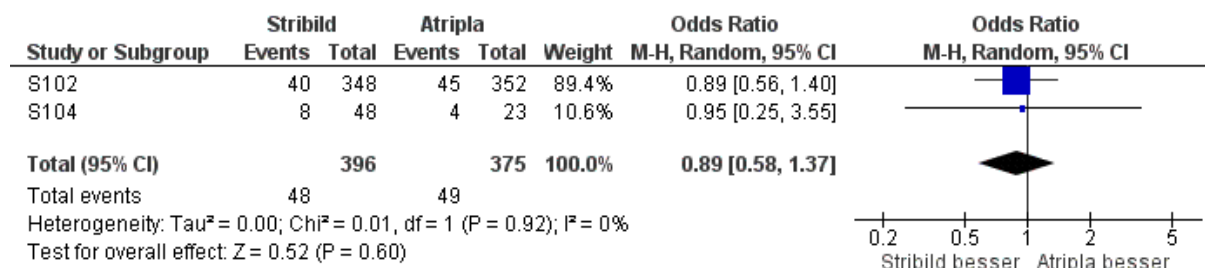


Abbildung 23: Meta-Analyse für Fatigue aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt hinsichtlich von **Fatigue** keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,89 [0,58;1,37], p=0,60). Die Datenlage ist dabei homogen (p=0,92; I<sup>2</sup>=0%) und die Sensitivitätsanalyse unter Verwendung des FEM zeigt robuste Ergebnisse (OR [95%-KI]: 0,89 [0,58;1,37], p=0,60).

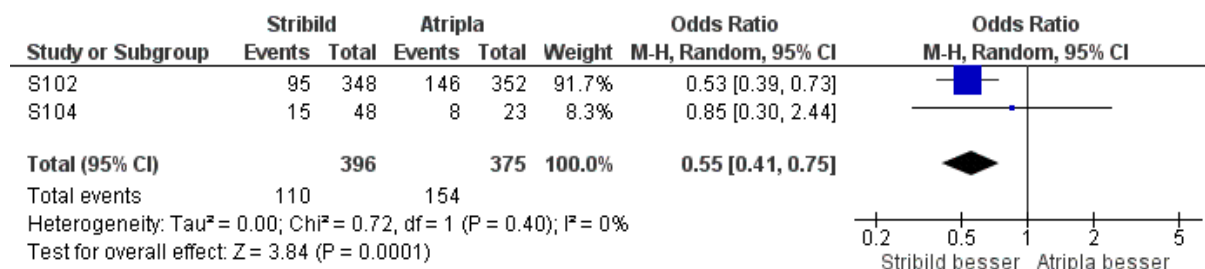


Abbildung 24: Meta-Analyse für Erkrankungen des Nervensystems (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**Erkrankungen des Nervensystems (SOC)** traten, laut Meta-Analyse, statistisch signifikant seltener unter einer Behandlung mit EVG/COBI/FTC/TDF auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,55 [0,41;0,75], p=0,0001). Dabei ist die Datenlage homogen (p=0,40; I<sup>2</sup>=0%) und eine Sensitivitätsanalyse unter Verwendung des FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,55 [0,41;0,75], p=0,0001).

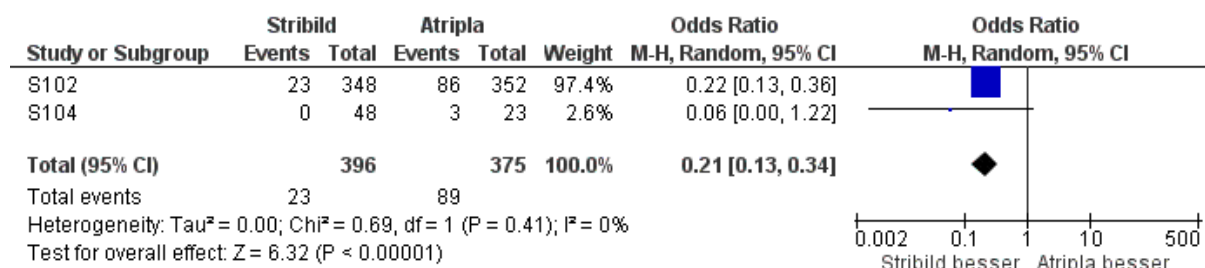


Abbildung 25: Meta-Analyse für Schwindel aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zu **Schwindel** zeigt einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,21 [0,13;0,34],

$p < 0,00001$ ). Schwindel tritt somit unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF. Die Datenlage ist dabei homogen ( $p = 0,41$ ;  $I^2 = 0\%$ ) und die Sensitivitätsanalyse unter Verwendung des FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,21 [0,13;0,34],  $p < 0,00001$ ).

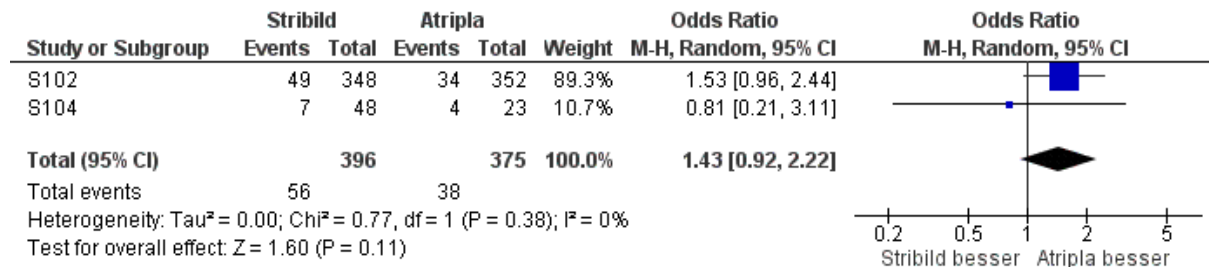


Abbildung 26: Meta-Analyse für Kopfschmerzen aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt - bei homogener Datenlage ( $p = 0,38$ ;  $I^2 = 0\%$ ) -, dass hinsichtlich von **Kopfschmerzen** zwischen den beiden Behandlungen kein statistisch signifikanter Unterschied besteht (OR [95%-KI]: 1,43 [0,92;2,22],  $p = 0,11$ ). Eine Sensitivitätsanalyse unter Verwendung des FEM zeigt robuste Ergebnisse (OR [95%-KI]: 1,43 [0,92;2,23],  $p = 0,11$ ).

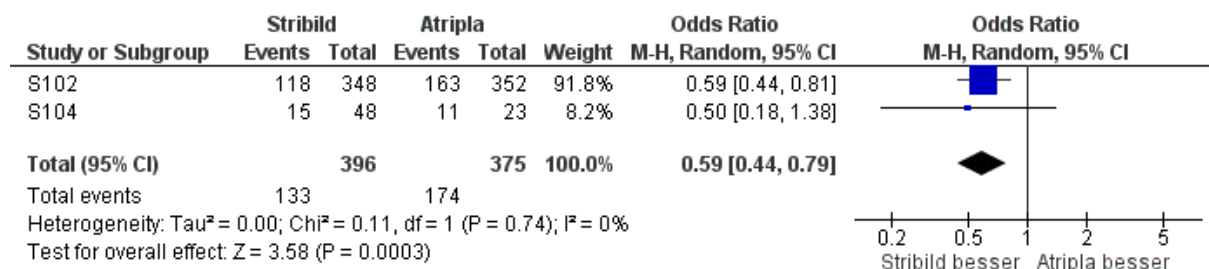


Abbildung 27: Meta-Analyse für psychiatrische Erkrankungen (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt, dass unter einer Behandlung mit EVG/COBI/FTC/TDF **psychiatrische Erkrankungen (SOC)** statistisch signifikant seltener auftraten als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,59 [0,44;0,79],  $p = 0,0003$ ). Dabei ist die Datenlage homogen ( $p = 0,74$ ;  $I^2 = 0\%$ ) und eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,59 [0,44;0,79],  $p = 0,0003$ ).



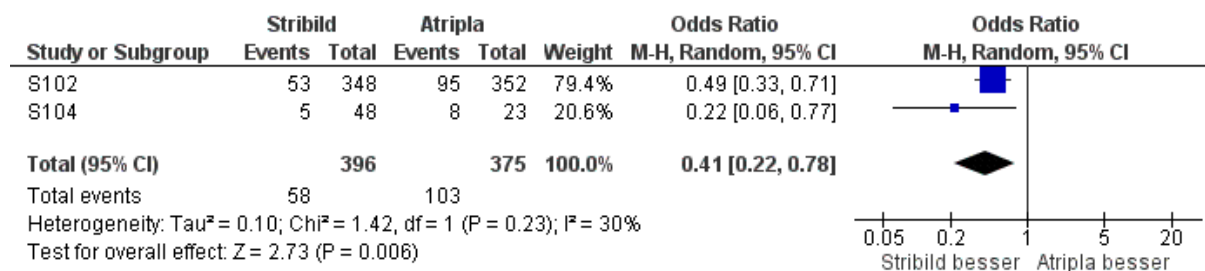


Abbildung 28: Meta-Analyse für abnorme Träume aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**Abnorme Träume** traten, laut Meta-Analyse, statistisch signifikant seltener unter einer Behandlung mit EVG/COBI/FTC/TDF auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,41 [0,22;0,78],  $p=0,006$ ). Es besteht keine bedeutsame Heterogenität ( $p=0,23$ ;  $I^2=30\%$ ) zwischen den Studien und die Sensitivitätsanalyse mittels FEM zeigt robuste Ergebnisse (OR [95%-KI]: 0,46 [0,32;0,65],  $p<0,0001$ ).

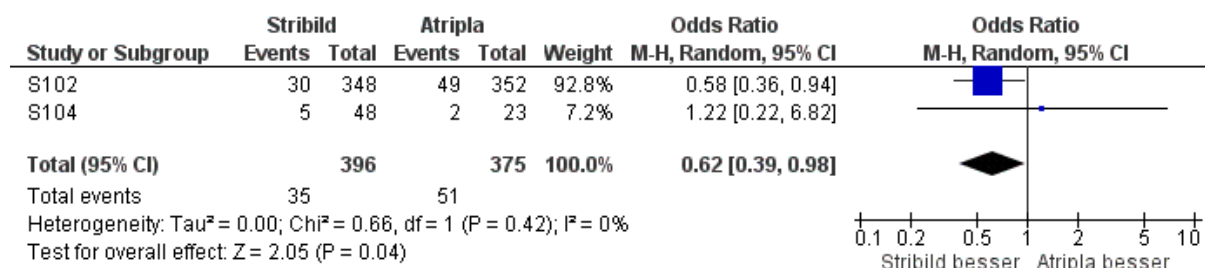


Abbildung 29: Meta-Analyse für Schlaflosigkeit aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**Schlaflosigkeit** tritt, laut Meta-Analyse bei homogener Datenlage ( $p=0,42$ ;  $I^2=0\%$ ), unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,62 [0,39;0,98],  $p=0,04$ ). Eine Sensitivitätsanalyse unter Verwendung des FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,62 [0,39;0,98],  $p=0,04$ ).

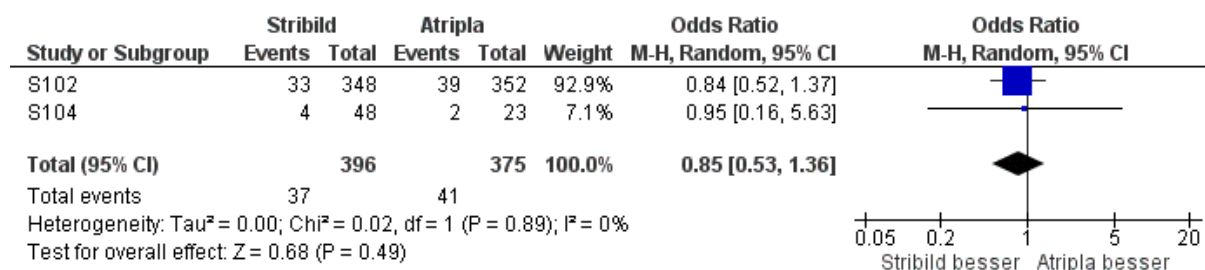


Abbildung 30: Meta-Analyse für Depression aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt, dass **Depressionen** in beiden Behandlungsarmen vergleichbar oft auftreten (OR [95%-KI]: 0,85 [0,53;1,36],  $p=0,49$ ). Die Datenlage ist dabei homogen ( $p=0,89$ ;  $I^2=0\%$ ) und eine Sensitivitätsanalyse zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,85 [0,53;1,36],  $p=0,49$ ).

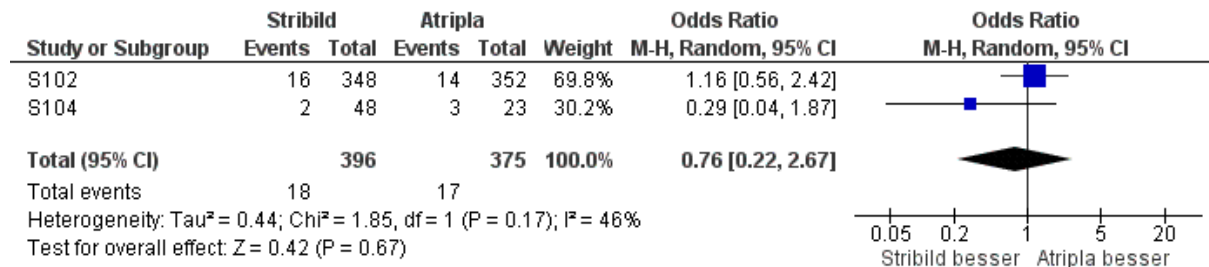


Abbildung 31: Meta-Analyse für Angst aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt dass hinsichtlich des Auftretens von **Angst** kein statistisch signifikanter Unterschied zwischen den Behandlungen besteht (OR [95%-KI]: 0,76 [0,22;2,67],  $p=0,67$ ). Zwischen den beiden Studien besteht jedoch Heterogenität ( $p=0,17$ ;  $I^2=46\%$ ). Unter Verwendung des FEM beträgt das OR [95%-KI] 0,97 [0,49;1,90] mit  $p=0,92$ .

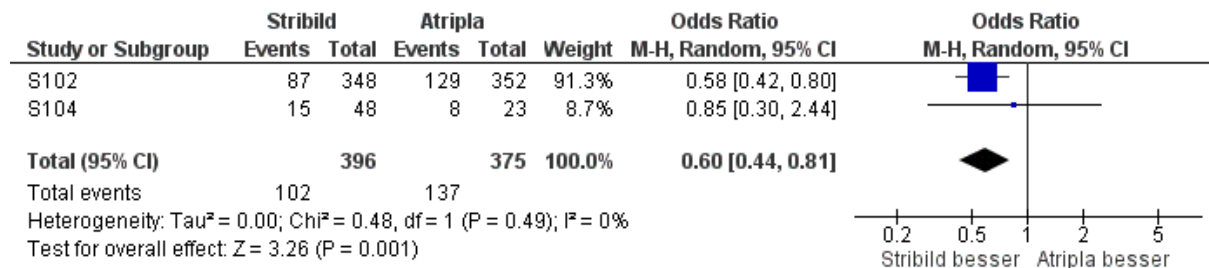


Abbildung 32: Meta-Analyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC) aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**Erkrankungen der Haut und des Unterhautzellgewebes (SOC)** treten, laut Meta-Analyse, statistisch signifikant seltener unter einer Behandlung mit EVG/COBI/FTC/TDF auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,60 [0,44;0,81],  $p=0,001$ ). Die Studie sind untereinander homogen ( $p=0,49$ ;  $I^2=0\%$ ) und eine Sensitivitätsanalyse unter Verwendung des FEM zeigt robuste Ergebnisse (OR [95%-KI]: 0,60 [0,44;0,81],  $p=0,001$ ).

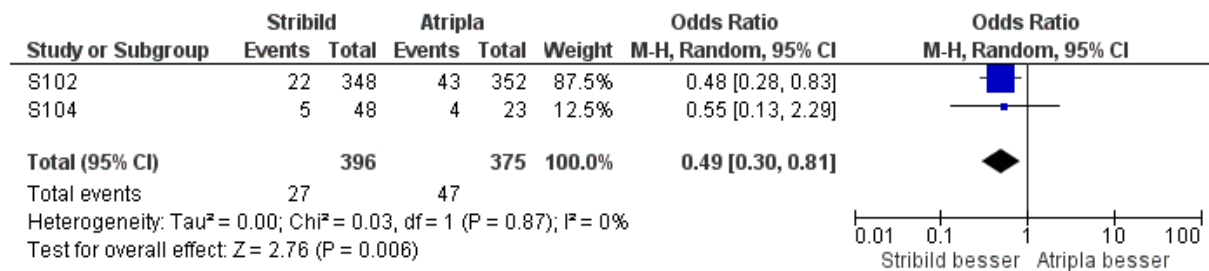


Abbildung 33: Meta-Analyse für Hautausschlag aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

Die Meta-Analyse zeigt, dass **Hautausschlag** statistisch signifikant seltener unter einer Behandlung mit EVG/COBI/FTC/TDF auftrat als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,49 [0,30;0,81], p=0,006). Dabei ist die Datenlage homogen (p=0,74; I<sup>2</sup>=0%) und eine Sensitivitätsanalyse unter Verwendung des FEM zeigt die Robustheit der Ergebnisse (OR [95%-KI]: 0,49 [0,30;0,81], p=0,006).

#### 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 in Studie S102 Subgruppenanalysen 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 ( $\leq 100.000$  Kopien/mL,  $> 100.000$  Kopien/mL)
- CD4-Zellzahl zu Baseline ( $\leq 350$  Zellen/ $\mu$ L,  $> 350$  Zellen/ $\mu$ L)
- Therapieadhärenz ( $< 95\%$ ,  $\geq 95\%$ )

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 [65, 66, 68, 69]. 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.

In Studie S104 wurden keine Subgruppenanalysen geplant und durchgeführt. Folglich sind alle hier für diese Studie dargestellten Subgruppenanalysen im Nachhinein durchgeführt worden [63, 64, 67]. Außerdem wird für die Endpunkte neurologische & psychiatrische Ereignisse, Hautausschläge, renale Ereignisse und Knochenfrakturen die Robustheit der Ergebnisse überprüft, indem die Studie S104 ausgeschlossen wird.

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 nicht nur eine Tablette täglich einnehmen mussten, wie für EVG/COBI/FTC/TDF und EFV/FTC/TDF in der Praxis vorgesehen, sondern zusätzlich eine Placebo-Tablette. Die Adhärenz-Daten in diesen Studien sind somit nicht auf die Praxis übertragbar. Aus diesem Grund werden die Daten hier vollständig dargestellt, jedoch nicht in die Bewertung und Nutzenableitung aufgenommen.

Zwei Patienten im EVG/COBI/FTC/TDF-Arm und drei Patienten im EFV/FTC/TDF-Arm der Studie S102 gaben zu Woche 48 und 96 nicht mindestens eine Flasche der Studienmedikation zurück, weswegen diese Patienten nicht in der Analyse der Therapieadhärenz berücksichtigt werden können. In Studie S104 gaben zwei Patienten im EVG/COBI/FTC/TDF-Arm und ein

Patient im EFV/FTC/TDF-Arm zu Woche 48, 60 bzw. 96 nicht mindestens eine Flasche der Studienmedikation wieder zurück. Diese Patienten können nicht in der Analyse der Therapieadhärenz berücksichtigt werden.

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, 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, nahezu bei 1 liegt. Bei einer Irrtumswahrscheinlichkeit von  $\alpha = 0,05$ , sind bei 462 durchgeführten Subgruppenanalysen ca. 23 fälschlicherweise als signifikant eingestufte Interaktionen zu erwarten.

Detaillierte Ergebnisse werden für alle präspezifizierten (nur S102) 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 4-55). Details zu allen durchgeführten Subgruppenanalysen werden aus Übersichtlichkeitsgründen in der Anlage [70] vollständig dargestellt.

In der folgenden Tabelle 4-55 werden alle präspezifizierten (nur in Studie S102) und post-hoc durchgeführten Interaktionstests mittels p-Wert zusammengefasst. Dabei werden die Analysezeitpunkte (Woche 48, Woche 60 und Woche 96) durch unterschiedliche Grauschattierungen differenziert dargestellt.

Tabelle 4-55: Übersicht über alle präspezifizierten (fett gedruckt) und post-hoc durchgeführten Interaktionstests

Endpunkt	p-Wert aus Interaktionstest					
	Alter	Geschlecht	Rasse	Ausgangsviruslast	CD4-Zellzahl zu Baseline	Therapieadhärenz
Woche 48						
Woche 60						
Woche 96						
Virologisches Ansprechen (Snapshot)	<b>0,50</b>	<b>0,61</b>	<b>0,45</b>	<b>0,45</b>	<b>0,99</b>	<b>0,78</b>
	<b>0,33</b>	<b>0,94</b>	<b>0,26</b>	<b>0,30</b>	<b>0,30</b>	<b>0,18</b>
Virologisches Ansprechen (TLOVR-Algorithmus)	0,73	0,85	0,39	0,43	0,91	0,83
	0,51	0,50	0,19	0,41	0,07	0,47
Zeit bis zum Verlust des	0,86	0,99	0,37	0,40	0,89	0,85

Endpunkt	p-Wert aus Interaktionstest					
	Alter	Geschlecht	Rasse	Ausgangsviruslast	CD4-Zellzahl zu Baseline	Therapieadhärenz
Woche 48						
Woche 60						
Woche 96						
virologischen Ansprechens	0,53	0,66	0,21	0,43	0,24	0,68
Veränderung der CD4-Zellzahl	0,35	0,70	0,43	0,84	0,57	0,35
	0,41	0,55	0,88	0,86	0,28	0,10
Mortalität	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.
	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.
	n.b.	n.b.	0,57	n.b.	0,65	0,32
Gesamt UE	0,62	0,47	0,48	0,56	0,65	0,93
	n.b.	n.b.	0,86	0,96	0,32	0,66
	0,1979	0,80	0,75	0,62	0,94	0,48
Gesamt SUE	0,50	0,64	0,33	0,96	0,28	0,24
	0,51	0,51	0,73	0,59	n.b.	n.b.
	0,49	0,64	0,14	0,85	0,53	0,57
Abbrüche aufgrund UE	0,67	0,41	0,009	0,58	0,21	0,65
	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.
	0,83	0,46	0,04	0,30	0,87	0,33
UE Grad 3-4	0,06	0,71	0,14	0,58	0,44	0,90
	n.b.	0,31	0,77	0,36	0,58	0,42
	0,41	0,90	0,18	0,91	0,54	0,76
Neurologische & psychiatrische Ereignisse	0,55	0,31	0,81	0,74	0,47	0,16
	0,73	0,54	0,83	0,71	0,08	0,66
	0,18	0,24	0,37	0,94	0,77	0,81
Hautausschläge	0,12	0,72	0,03	0,57	0,72	0,95
	0,64	0,49	0,71	0,16	0,63	0,17
	0,26	0,75	0,11	0,40	0,99	0,35
Renale Ereignisse	0,82	n.b.	n.b.	0,11	0,55	n.b.
	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.
	0,58	n.b.	0,94	0,1976	0,40	0,69
Knochenfrakturen	0,51	n.b.	0,61	0,28	0,39	0,50
	n.b.	n.b.	n.b.	n.b.	n.b.	n.b.
	0,84	n.b.	0,58	0,40	0,18	0,37
Erkrankungen des Gastrointestinaltrakts (SOC)	0,45	0,41	0,87	0,28	0,15	0,60
	0,96	0,16	0,45	0,43	0,52	0,28
	0,47	0,38	0,80	0,03	0,15	0,09

Endpunkt	p-Wert aus Interaktionstest					
	Alter	Geschlecht	Rasse	Ausgangsviruslast	CD4-Zellzahl zu Baseline	Therapieadhärenz
Woche 48						
Woche 60						
Woche 96						
<i>Diarrhoe</i>	0,57	0,53	0,84	0,92	0,19	0,38
	0,72	0,51	0,43	0,11	0,58	0,10
	0,95	0,45	0,58	0,21	0,48	0,17
<i>Übelkeit</i>	0,54	0,46	0,39	0,18	0,89	0,39
	n.z.	n.z.	n.z.	n.z.	n.z.	n.z.
	0,68	0,21	0,88	0,11	0,52	0,36
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)	0,77	0,34	0,50	0,85	0,08	0,93
	1,00	0,63	0,66	0,65	0,66	0,44
	0,55	0,03	0,27	0,40	0,04	0,79
<i>Fatigue</i>	0,84	0,2017	0,85	0,79	0,90	0,94
	0,81	0,62	0,36	0,21	0,70	0,30
	0,46	0,02	0,47	0,48	0,87	0,71
Erkrankungen des Nervensystems (SOC)	0,46	0,24	0,50	0,39	0,64	0,0532
	0,88	0,92	0,14	0,63	0,44	0,80
	0,33	0,22	0,78	0,35	0,94	0,18
<i>Schwindel</i>	0,50	1,00	0,87	0,70	0,45	0,74
	n.b.	n.b.	0,99	0,87	0,97	0,94
	0,71	0,92	0,76	0,84	0,68	0,57
<i>Kopfschmerzen</i>	0,38	0,18	0,19	0,31	0,45	0,18
	0,14	0,25	0,22	0,65	0,75	0,93
	0,61	0,23	0,56	0,47	0,75	0,24
Psychiatrische Erkrankungen (SOC)	0,87	0,61	0,78	0,85	0,79	0,35
	0,45	0,71	0,17	0,44	0,35	0,33
	0,63	0,41	0,84	0,50	0,66	0,86
<i>Abnorme Träume</i>	0,64	0,36	0,13	0,61	0,89	0,88
	0,08	0,19	0,14	0,66	0,58	0,31
	0,36	0,71	0,14	0,76	0,69	0,69
<i>Schlaflosigkeit</i>	0,57	0,52	0,84	0,84	0,69	0,57
	0,56	0,78	0,47	0,91	0,69	0,40
	0,85	0,42	0,45	0,51	0,38	0,61
<i>Depression</i>	0,15	0,25	0,94	0,74	0,88	0,40
	n.z.	n.z.	n.z.	n.z.	n.z.	n.z.
	0,15	0,09	0,82	0,35	0,96	0,54

Endpunkt	p-Wert aus Interaktionstest					
	Alter	Geschlecht	Rasse	Ausgangsviruslast	CD4-Zellzahl zu Baseline	Therapieadhärenz
Woche 48						
Woche 60						
Woche 96						
<i>Angst</i>	0,50	0,51	0,41	0,53	0,97	0,27
	0,50	0,51	0,41	0,53	0,97	0,29
	n.z.	n.z.	n.z.	n.z.	n.z.	n.z.
Erkrankungen der Haut und des Unterhautzellgewebes (SOC)	0,32	0,45	0,04	0,51	0,83	0,17
	0,34	0,12	0,44	0,14	0,71	0,51
	0,34	0,78	0,07	0,76	0,92	0,06
Hautausschlag	0,17	0,57	0,55	0,36	0,85	0,95
	0,24	n.b.	0,90	0,09	0,22	0,71
	0,46	0,58	0,44	0,88	0,29	0,36

Im Studienprotokoll der Studie S102 präspezifizierte Subgruppenanalysen sind fett gedruckt  
n.b. = nicht berechenbar, n.z. = nicht zutreffend, SOC = System Organ Class, SUE = Schwerwiegendes unerwünschtes Ereignis, UE = Unerwünschtes Ereignis

## Zusammenfassung der Subgruppenanalysen

### *Virologisches Ansprechen (Snapshot)*

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,50$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,61$ ;  $I^2=0\%$ ), Rasse ( $p=0,45$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,45$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,99$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,78$ ;  $I^2=0\%$ ).

Es bestehen zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,33$ ;  $I^2=0,2\%$ ), Geschlecht ( $p=0,94$ ;  $I^2=0\%$ ), Rasse ( $p=0,26$ ;  $I^2=22,2\%$ ), Ausgangsviruslast ( $p=0,30$ ;  $5,9\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,30$ ;  $I^2=7,8\%$ ). Es ergibt sich allerdings ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,18$ ;  $I^2=44,8\%$ ). In der Gruppe der Patienten die nach 96 Wochen eine Therapieadhärenz von  $<95\%$  aufwiesen, beträgt das OR [95%-KI] 1,73 [0,91;3,29] mit  $p=0,0969$  und in der Gruppe der Patienten die zu  $\geq 95\%$  adhärenz waren, liegt das OR [95%-KI] bei 0,97 [0,57;1,66] mit  $p=0,9115$ . 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]:  $<95\%$ : 11,1 [-2,5;24,8];  $\geq 95\%$ : -0,3 [-5,9;5,3]) zeigen außerdem, dass EVG/COBI/FTC/TDF in beiden Subgruppen gegenüber EFV/FTC/TDF nichtunterlegen zur 12%-Grenze ist.

### *Virologisches Ansprechen (TLOVR)*

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,73$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,85$ ;  $I^2=0\%$ ), Rasse ( $p=0,39$ ;  $I^2=0\%$ ), Ausgangsviruslast



( $p=0,43$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,91$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,83$ ;  $I^2=0\%$ ).

Zu Woche 96 kann kein Hinweis oder Beleg auf Effektmodifikation durch Alter ( $p=0,51$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,50$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,41$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,47$ ;  $I^2=0\%$ ) gefunden werden. Bezüglich Rasse ( $p=0,19$ ;  $I^2=42,4\%$ ) besteht allerdings ein Hinweis auf Effektmodifikation. Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht (OR [95%\_KI]: kaukasisch: 0,93 [0,58;1,49],  $p=0,7605$ ; nicht-kaukasisch: 1,52 [0,87;2,67],  $p=0,1432$ ), wirkt sich diese Interaktion nicht auf das Gesamtfazit aus. Auch hinsichtlich CD4-Zellzahl zu Baseline ( $p=0,07$ ;  $I^2=69,1\%$ ) besteht ein Hinweis auf Effektmodifikation. Doch auch hier besteht in keiner Subgruppe ein signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]:  $\leq 350$  Zellen/ $\mu\text{L}$ : 0,81 [0,49;1,37],  $p=0,4355$ ;  $>350$  Zellen/ $\mu\text{L}$ : 1,59 [0,95;2,66],  $p=0,0774$ ), weswegen sich diese Interaktion nicht auf das Gesamtfazit auswirkt.

### ***Zeit bis zum Verlust des virologischen Ansprechens***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,86$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,99$ ;  $I^2=0\%$ ), Rasse ( $p=0,37$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,40$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,89$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,85$ ;  $I^2=0\%$ ).

Auch zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,53$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,66$ ;  $I^2=0\%$ ), Rasse ( $p=0,21$ ;  $I^2=35,9\%$ ), Ausgangsviruslast ( $p=0,43$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,24$ ;  $I^2=26,1\%$ ) oder Therapieadhärenz ( $p=0,68$ ;  $I^2=0\%$ ).

### ***Veränderung der CD4-Zellzahl***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,35$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,70$ ;  $I^2=0\%$ ), Rasse ( $p=0,43$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,84$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,57$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,35$ ;  $I^2=0\%$ ).

Zu Woche 96 zeigen sich keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,41$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,55$ ;  $I^2=0\%$ ), Rasse ( $p=0,88$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,86$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,28$ ;  $I^2=15,2\%$ ). Es besteht allerdings ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,10$ ;  $I^2=63,5\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhären waren, beträgt die MWD [95%-KI] 72,00 [5,52;138,48] mit  $p=0,0386$ , was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF anzeigt (Hedges'  $g$  [95%-KI]: 0,37 [0,02;0,71]), und in der Gruppe der zu  $\geq 95\%$  adhären Patienten liegt die MWD [95%-KI] bei 8,00 [-28,45;44,45] mit  $p=0,6671$ .

**Mortalität**

Zu Woche 48 sind aufgrund zu geringer Ereignisraten keine Interaktionstests durchführbar.

Auch zu Woche 60 sind die Ereignisraten zu gering um Interaktionstests durchführen zu können.

Zu Woche 96 sind Interaktionstests, stratifiziert nach Alter, Geschlecht und Ausgangsviruslast aufgrund zu geringer Ereignisraten nicht durchführbar. Doch auch hinsichtlich Rasse ( $p=0,57$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,65$ ;  $I^2=0\%$ ) und Therapieadhärenz ( $p=0,32$ ;  $I^2=0\%$ ) kann kein Hinweis oder Beleg auf Effektmodifikation gefunden werden.

**Gesamtrate unerwünschter Ereignisse**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,62$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,47$ ;  $I^2=0\%$ ), Rasse ( $p=0,48$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,56$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,65$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,93$ ;  $I^2=0\%$ ).

Zu Woche 60 sind Interaktionstests, stratifiziert nach Alter und Geschlecht aufgrund zu geringer Ereignisraten nicht durchführbar. Es bestehen keine Hinweise oder Belege auf Effektmodifikation durch Rasse ( $p=0,86$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,96$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,32$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,66$ ;  $I^2=0\%$ ).

Zu Woche 96 zeigen sich keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ( $p=0,80$ ;  $I^2=0\%$ ), Rasse ( $p=0,75$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,62$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,94$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,48$ ;  $I^2=0\%$ ). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Alter ( $p=0,1979$ ;  $I^2=39,7\%$ ). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 1,37 [0,47;4,02] mit  $p=0,5681$  und für die mindestens 40-jährigen Patienten ergibt sich das OR [95%-KI] 0,37 [0,07;1,96] mit  $p=0,2441$ . Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

**Gesamtrate schwerwiegender unerwünschter Ereignisse**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,50$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,64$ ;  $I^2=0\%$ ), Rasse ( $p=0,33$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,96$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,28$ ;  $I^2=12,7\%$ ) oder Therapieadhärenz ( $p=0,24$ ;  $I^2=26,7\%$ ).

Es bestehen zu Woche 60 keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,51$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,51$ ;  $I^2=0\%$ ), Rasse ( $p=0,73$ ;  $I^2=0\%$ ) oder Ausgangsviruslast ( $p=0,59$ ;  $I^2=0\%$ ). Subgruppenanalysen, stratifiziert nach CD4-Zellzahl zu Baseline und Therapieadhärenz waren aufgrund zu geringer Ereignisraten nicht durchführbar.

Zu Woche 96 zeigen sich keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,49$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,64$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,85$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,53$ ;  $I^2=0\%$ ) oder durch Therapieadhärenz ( $p=0,57$ ;  $I^2=0\%$ ). Es kann allerdings ein Hinweis auf Effektmodifikation durch Rasse gefunden werden ( $p=0,14$ ;  $I^2=53,0\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 2,49 [1,33;4,66] mit  $p=0,0043$ , was einen statistisch signifikanten Unterschied zwischen den beiden Behandlungen zuungunsten von EVG/COBI/FTC/TDF anzeigt, und in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 1,25 [0,63;2,48] mit  $p=0,5268$ .

### ***Therapieabbrüche aufgrund unerwünschter Ereignisse***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,67$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,41$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,58$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,21$ ;  $I^2=35,5\%$ ) oder Therapieadhärenz ( $p=0,65$ ;  $I^2=0\%$ ). Es besteht jedoch ein Beleg für Effektmodifikation durch Rasse ( $p=0,009$ ;  $I^2=85,4\%$ ). In der Gruppe der kaukasischen Patienten besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,48 [0,58;3,76],  $p=0,41$ ). In der Gruppe der nicht-kaukasischen Patienten besteht jedoch ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,16 [0,04;0,64],  $p=0,01$ ). Nicht-kaukasische Patienten brechen somit unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener die Behandlung aufgrund unerwünschter Ereignisse ab als unter einer Behandlung mit EFV/FTC/TDF.

Zu Woche 60 waren Subgruppenanalysen aufgrund zu geringer Ereignisraten nicht durchführbar.

Es können zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,83$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,46$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,30$ ;  $I^2=5,3\%$ ), CD4-Zellzahl zu Baseline ( $p=0,87$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,33$ ;  $I^2=0\%$ ) gefunden werden. Es besteht jedoch ein Beleg auf Effektmodifikation durch Rasse ( $p=0,04$ ;  $I^2=76\%$ ). Hier ergibt sich in der Gruppe der Kaukasier ein OR [95%-KI] von 1,15 [0,53;2,51] mit  $p=0,7214$  und in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,24 [0,06;0,87] mit  $p=0,0303$ . D. h. in der Gruppe der nicht-kaukasischen Patienten besteht ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen zugunsten von EVG/COBI/FTC/TDF.

### ***Unerwünschte Ereignisse Grad 3-4***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ( $p=0,71$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,58$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,44$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,90$ ;  $I^2=0\%$ ). Allerdings gibt es einen Hinweis auf Effektmodifikation durch Alter ( $p=0,06$ ;  $I^2=71,5\%$ ). In der Gruppe der <40-jährigen Patienten beträgt das OR [9%-KI] 1,77 [0,92;3,39] mit  $p=0,0877$  und für die mindestens 40-jährigen Patienten ergibt sich das OR [95%-KI] zu 0,74 [0,39;1,39] mit  $p=0,35$ . 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 ( $p=0,14$ ;  $I^2=54,3\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 1,55 [0,82;2,90] mit  $p=0,18$  und in der Gruppe der nicht-kaukasischen Patienten ist das OR [95%-KI] 0,78 [0,41;1,49] mit  $p=0,46$ . 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 60 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ( $p=0,31$ ;  $I^2=4,5\%$ ), Rasse ( $p=0,77$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,36$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,58$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,42$ ;  $I^2=0\%$ ). Eine Subgruppenanalyse, stratifiziert nach Alter, konnte aufgrund zu geringer Ereignisraten nicht durchgeführt werden.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,41$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,90$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,91$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,54$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,76$ ;  $I^2=0\%$ ). Allerdings kann ein Hinweis auf Effektmodifikation durch Rasse gefunden werden ( $p=0,18$ ;  $I^2=43,4\%$ ). Da allerdings in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht (OR [95%-KI]: Kaukasisch: 1,60 [0,91;2,80],  $p=0,1009$ ; Nicht-kaukasisch: 0,92 [0,50;1,67],  $p=0,7740$ ) wirkt sich diese Effektmodifikation nicht auf das Gesamtfazit aus.

### ***Neurologische & psychiatrische Ereignisse***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,55$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,31$ ;  $I^2=2,7\%$ ), Rasse ( $p=0,81$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,74$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,47$ ;  $I^2=0\%$ ). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,16$ ;  $I^2=48,3\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhären waren beträgt das OR [95%-KI] 0,65 [0,36;1,16] mit  $p=0,14$  und für Patienten, die  $\geq 95\%$  adhären waren ergibt sich das OR [95%-KI] zu 0,40 [0,29;0,57] mit  $p<0,00001$ . Wenn die Medikation also zu mindestens 95% adhären eingenommen wird, treten neurologische & psychiatrische Ereignisse unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF. Dieser Hinweis besteht auch unter Ausschluss der für diesen Endpunkt als nicht sicher niedrig verzerrt bewerteten Studie S104 ( $p=0,17$ ;  $I^2=47,4\%$ ).

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,73$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,54$ ;  $I^2=0\%$ ), Rasse ( $p=0,83$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,71$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,66$ ;  $I^2=0\%$ ). Allerdings gibt es einen Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,08$ ;  $I^2=67,0\%$ ). In der Gruppe der Patienten die zu Baseline eine CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  aufwiesen beträgt das OR [95%-KI] 1,67 [0,32;8,59] mit  $p=0,5415$  und in der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl  $>350$  Zellen/ $\mu\text{L}$  aufwiesen berechnet sich das OR [95%-KI] zu 0,25 [0,06;0,98] mit  $p=0,0471$ . Es besteht somit für Patienten, die zu Baseline eine höhere CD4-

Zellzahl zu Baseline aufweisen, ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF.

Zu Woche 96 können keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ( $p=0,24$ ;  $I^2=27,2\%$ ), Rasse ( $p=0,37$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,94$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,77$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,81$ ;  $I^2=0\%$ ) gefunden werden. Allerdings besteht ein Hinweis auf eine Effektmodifikation durch Alter ( $p=0,18$ ;  $I^2=44,2\%$ ). In der Gruppe der <40-jährigen Patienten ergibt sich das OR [95%-KI] zu 0,54 [0,36;0,81] mit  $p=0,0027$ , und in der Gruppe der  $\geq 40$ -jährigen beträgt das OR [95%-KI] 0,35 [0,22;0,56] mit  $p<0,0001$ . Da dies jedoch nur eine quantitative Interaktion ist, wirkt sie sich nicht auf das Gesamtfazit aus.

### **Hautausschläge**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ( $p=0,72$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,57$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,72$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,95$ ;  $I^2=0\%$ ). Allerdings besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,12$ ;  $I^2=58,3\%$ ). In der Gruppe der <40-jährigen Patienten beträgt das OR [95%-KI] 0,73 [0,46;1,16] mit  $p=0,18$ . Für  $\geq 40$ -jährige Patienten berechnet sich das OR [95%-KI] zu 0,30 [0,10;0,84] mit  $p=0,02$ . Unter den älteren Patienten besteht somit ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Dieser Hinweis besteht auch unter Ausschluss der für diesen Endpunkt als nicht sicher niedrig verzerrt bewerteten Studie S104 ( $p=0,12$ ;  $I^2=59,4\%$ ). Ein Beleg für Effektmodifikation besteht durch Rasse ( $p=0,03$ ;  $I^2=78,4\%$ ). In der kaukasischen Patientengruppe beträgt das OR [95%-KI] 0,42 [0,28;0,65] mit  $p<0,0001$  und für die nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,97 [0,52;1,84] mit  $p=0,94$ . Es besteht hier also in der kaukasischen Patientengruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Dieser Beleg besteht auch unter Ausschluss der für diesen Endpunkt als nicht sicher niedrig verzerrt bewerteten Studie S104 ( $p=0,03$ ;  $I^2=78,4\%$ ).

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,64$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,49$ ;  $I^2=0\%$ ), Rasse ( $p=0,71$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,63$ ;  $I^2=0\%$ ). Allerdings gibt es einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,16$ ;  $I^2=49,8\%$ ). In der Gruppe der Patienten mit Ausgangsviruslast  $\leq 100.000$  Kopien/mL beträgt das OR [95%-KI] 0,39 [0,10;1,44] mit  $p=0,1566$ . Für Patienten, die zu Baseline eine Viruslast  $>100.000$  Kopien/mL aufwiesen, ergibt sich das OR [95%-KI] zu 4,53 [0,19;105,84] mit  $p=0,3475$ . 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 ( $p=0,17$ ;  $I^2=46,3\%$ ). Für Patienten, die zu  $<95\%$  adhären waren, ergibt sich das OR [95%-KI] zu 0,13 [0,01;1,89] mit  $p=0,1331$ , und in der  $\geq 95\%$  adhären Patientengruppe beträgt das OR [95%-KI] 1,04 [0,26;4,11] mit  $p=0,9554$ . 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 ( $p=0,26$ ;  $I^2=19,6\%$ ), Geschlecht ( $p=0,75$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,40$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,99$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,35$ ;  $I^2=0\%$ ) gefunden werden. Es besteht jedoch ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,11$ ;  $I^2=59,9\%$ ). In der Gruppe der Patienten kaukasischer Herkunft beträgt das OR [95%-KI] 0,51 [0,33;0,77] mit  $p=0,0016$ , was einen statistisch signifikanten Unterschied zwischen den beiden Behandlungen zugunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der nicht-kaukasischen Patienten besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,92 [0,50;1,67],  $p=0,7740$ ).

### **Renale Ereignisse**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,82$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,55$ ;  $I^2=0\%$ ). Subgruppenanalysen, stratifiziert nach Geschlecht, Rasse und Therapieadhärenz sind aufgrund zu geringer Ereignisraten nicht durchführbar. Es besteht jedoch ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,11$ ;  $I^2=61,8\%$ ). In der Gruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 11,54 [0,63;209,84] mit  $p=0,0985$ . Und für Patienten, die zu Baseline eine Viruslast  $>100.000$  Kopien/mL aufwiesen, ergibt sich das OR [95%-KI] 0,32 [0,01;8,06] mit  $p=0,4925$ . 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 60 sind Subgruppenanalysen aufgrund zu geringer Ereignisraten nicht durchführbar.

Zu Woche 96 können keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,58$ ;  $I^2=0\%$ ), Rasse ( $p=0,94$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,40$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,69$ ;  $I^2=0\%$ ) gefunden werden. Der Interaktionstest, stratifiziert nach Geschlecht, ist aufgrund zu geringer Ereignisraten nicht berechenbar. Es besteht ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,1976$ ;  $I^2=39,8\%$ ). In der Gruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 13,69 [0,77;244,51] mit  $p=0,0751$ , und in der Gruppe der Patienten mit höherer Viruslast zu Baseline ergibt sich das OR [95%-KI] zu 0,98 [0,06;15,90] mit  $p=0,9903$ . Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

### **Knochenfrakturen**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,51$ ;  $I^2=0\%$ ), Rasse ( $p=0,61$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,28$ ;  $I^2=13,2\%$ ), CD4-Zellzahl zu Baseline ( $p=0,39$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,50$ ;  $I^2=0\%$ ). Eine

Subgruppenanalyse, stratifiziert nach Geschlecht, ist aufgrund zu geringer Ereignisraten nicht durchführbar.

Zu Woche 60 sind Subgruppenanalysen aufgrund zu geringer Ereignisraten nicht durchführbar.

Es können zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,84$ ;  $I^2=0\%$ ), Rasse ( $p=0,58$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,40$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,37$ ;  $I^2=0\%$ ) gefunden werden. Der Interaktionstest, stratifiziert nach Geschlecht, war aufgrund zu geringer Ereignisraten nicht berechenbar. Eine Subgruppenanalyse, stratifiziert nach CD4-Zellzahl zu Baseline, ergibt jedoch einen Hinweis auf Effektmodifikation ( $p=0,18$ ;  $I^2=45,5\%$ ). Da aber in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht (OR [95%-KI]:  $\leq 350$  Zellen/ $\mu\text{L}$ : 2,92 [0,58;14,70],  $p=0,1939$ ;  $>350$  Zellen/ $\mu\text{L}$ : 0,70 [0,20;2,53],  $p=0,5881$ ), wirkt sich diese Interaktion nicht auf das Gesamtfazit aus.

### ***Erkrankungen des Gastrointestinaltrakts (SOC)***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,45$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,41$ ;  $I^2=0\%$ ), Rasse ( $p=0,87$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,28$ ;  $I^2=14,5\%$ ) oder Therapieadhärenz ( $p=0,60$ ;  $I^2=0\%$ ). Es gibt jedoch einen Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,15$ ;  $I^2=52,6\%$ ). In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  aufwiesen, beträgt das OR [95%-KI] 1,87 [1,20;2,93] mit  $p=0,006$  und in der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl  $>350$  Zellen/ $\mu\text{L}$  aufwiesen, ergibt sich das OR [95%-KI] zu 1,22 [0,83;1,77] mit  $p=0,31$ . Das heißt, dass sich der statistisch signifikante Unterschied zuungunsten von EVG/COBI/FTC/TDF auf die Gruppe der Patienten beschränkt, die zu Baseline eine CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  aufwiesen.

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,96$ ;  $I^2=0\%$ ), Rasse ( $p=0,45$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,43$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,52$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,28$ ;  $I^2=15,0\%$ ). Es gibt allerdings einen Hinweis auf Effektmodifikation durch Geschlecht ( $p=0,16$ ;  $I^2=49,2\%$ ). In der Gruppe der männlichen Patienten beträgt das OR [95%-KI] 2,00 [0,68;5,91] mit  $p=0,2096$  und für die weiblichen Patienten ergibt sich das OR [95%-KI] zu 45,00 [0,67;3042,57] mit  $p=0,0766$ . 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 ( $p=0,47$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,38$ ;  $I^2=0\%$ ) oder Rasse ( $p=0,80$ ;  $I^2=0\%$ ). Allerdings gibt es einen Beleg für Effektmodifikation durch Ausgangsviruslast ( $p=0,03$ ;  $I^2=78,6\%$ ). In der Subgruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 1,06 [0,74;1,53] mit  $p=0,7464$ . In der Subgruppe der Patienten, die zu Baseline eine Viruslast  $>100.000$  Kopien/mL aufwiesen, besteht ein statistisch signifikanter Unterschied und das OR [95%-KI] beträgt 2,16 [1,27;3,65] mit  $p=0,0042$ . Des

Weiteren besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,15$ ;  $I^2=51,9\%$ ). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline  $\leq 350$  Zellen/ $\mu\text{L}$  beträgt das OR [95%-KI] 1,74 [1,10;2,75] mit  $p=0,0171$ , was einen statistisch signifikanten Unterschied zuungunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline  $>350$  Zellen/ $\mu\text{L}$  ergibt sich das OR [95%-KI] zu 1,11 [0,75;1,66] mit  $p=0,5966$ , was keinen statistisch signifikanten Unterschied anzeigt. Auch hinsichtlich Therapieadhärenz kann ein Hinweis auf Effektmodifikation gefunden werden ( $p=0,09$ ;  $I^2=64,2\%$ ). In der Gruppe der Patienten mit einer Therapieadhärenz von  $<95\%$  beträgt das OR [95%-KI] 2,10 [1,14;3,85] mit  $p=0,0173$  und zeigt einen statistisch signifikanten Unterschied an. In der Gruppe der Patienten die zu  $\geq 95\%$  adhärenz waren, liegt das OR [95%-KI] bei 1,15 [0,81;1,63] mit  $p=0,4243$ .

### **Diarrhoe**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,57$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,53$ ;  $I^2=0\%$ ), Rasse ( $p=0,84$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,92$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,38$ ;  $I^2=0\%$ ). Allerdings gibt es einen Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,19$ ;  $I^2=40,9\%$ ). In der Gruppe der Patienten mit einer Baseline CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  beträgt das OR [95%-KI] 1,73 [0,99;3,03] mit  $p=0,06$ , und in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline  $>350$  Zellen/ $\mu\text{L}$  ergibt sich das OR [95%-KI] zu 1,07 [0,68;1,69] mit  $p=0,77$ . 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 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,72$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,51$ ;  $I^2=0\%$ ), Rasse ( $p=0,43$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,58$ ;  $I^2=0\%$ ). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,11$ ;  $I^2=60,2\%$ ). In der Gruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 3,38 [0,66;17,28] mit  $p=0,1427$ , und in der Subgruppe der Patienten, die zu Baseline eine Viruslast  $>100.000$  Kopien/mL aufwiesen, ergibt sich das OR [95%-KI] zu 0,33 [0,03;3,51] mit  $p=0,3607$ . 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 ( $p=0,10$ ;  $I^2=63,5\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhärenz waren, liegt das OR [95%-KI] bei 12,69 [0,58;276,08] mit  $p=0,1059$  und in der Gruppe der Patienten, die eine Therapieadhärenz  $\geq 95\%$  aufwiesen, ergibt sich das OR [95%-KI] zu 0,72 [0,17;3,01] mit  $p=0,6551$ . 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 ( $p=0,95$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,45$ ;  $I^2=0\%$ ), Rasse ( $p=0,58$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,21$ ;  $I^2=37,3\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,48$ ;  $I^2=0\%$ ). Es besteht jedoch ein



Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,17$ ;  $I^2=47,7\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhärenz waren, ergibt sich das OR [95%-KI] zu 1,65 [0,82;3,32] mit  $p=0,1612$  und in der Gruppe der zu  $\geq 95\%$  adhärenz Patienten beträgt das OR [95%-KI] 0,93 [0,63;1,39] mit  $p=0,7372$ . Da aber in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich diese Interaktion nicht auf das Gesamtfazit aus.

### **Übelkeit**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,54$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,46$ ;  $I^2=0\%$ ), Rasse ( $p=0,39$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,89$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,39$ ;  $I^2=0\%$ ). Es gibt allerdings einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,18$ ;  $I^2=45,0\%$ ). In der Gruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 1,36 [0,84;2,21] mit  $p=0,2172$  und in der Gruppe der Patienten mit Ausgangsviruslast  $>100.000$  Kopien/mL ergibt sich das OR [95%-KI] zu 2,46 [1,20;5,04] mit  $p=0,0135$ . Der statistisch signifikante Unterschied zuungunsten von EVG/COBI/FTC/TDF beschränkt sich hier also auf die Gruppe der Patienten, die zu Baseline eine Viruslast  $>100.000$  Kopien/mL aufwiesen.

Da in Studie S104 nicht mindestens 10% der Patienten in einem der Behandlungsarme Übelkeit aufwiesen, werden für diesen Endpunkt zu Woche 60 keine Subgruppenanalysen durchgeführt.

Es können zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,68$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,21$ ;  $I^2=36,6\%$ ), Rasse ( $p=0,88$ ;  $I^2=0\%$ ), CD4-Zellzahl ( $p=0,52$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,36$ ;  $I^2=0\%$ ) gefunden werden. Es gibt jedoch einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,11$ ;  $I^2=59,8\%$ ). In der Gruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen beträgt das OR [95%-KI] 1,26 [0,79;2,02] mit  $p=0,3321$ , und in der Gruppe der Patienten mit Ausgangsviruslast  $>100.000$  Kopien/mL beträgt das OR [95%-KI] 2,48 [1,24;4,98] mit  $p=0,0104$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen anzeigt.

### **Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,77$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,34$ ;  $I^2=0\%$ ), Rasse ( $p=0,50$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,85$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,93$ ;  $I^2=0\%$ ). Allerdings gibt es einen Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,08$ ;  $I^2=67,4\%$ ). In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  aufwiesen, beträgt das OR [95%-KI] 0,46 [0,27;0,75] mit  $p=0,002$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der Patienten mit höherer CD4-Zellzahl zu Baseline ergibt sich das OR [95%-KI] zu 0,83 [0,53;1,31] mit  $p=0,43$ .

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=1,00$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,63$ ;  $I^2=0\%$ ), Rasse ( $p=0,66$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,65$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,66$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,44$ ;  $I^2=0\%$ ).

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,55$ ;  $I^2=0\%$ ), Rasse ( $p=0,27$ ;  $I^2=19,1\%$ ), Ausgangsviruslast ( $p=0,40$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,79$ ;  $I^2=0\%$ ). Es besteht jedoch ein Beleg für Effektmodifikation durch Geschlecht ( $p=0,03$ ;  $I^2=78,8\%$ ). Unter den Männern traten allgemeine Erkrankungen und Beschwerden am Verabreichungsort statistisch signifikant seltener im EVG/COBI/FTC/TDF-Arm auf als im EFV/FTC/TDF-Arm (OR [95%-KI]: 0,63 [0,44;0,90],  $p=0,0111$ ). In der Gruppe der Frauen kann kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen beobachtet werden (OR [95%-KI]: 2,15 [0,75;6,13],  $p=0,1527$ ). Auch für CD4-Zellzahl zu Baseline kann ein Beleg für Effektmodifikation gefunden werden ( $p=0,04$ ;  $I^2=75,6\%$ ): Patienten mit  $\leq 350$  CD4-Zellen/ $\mu\text{L}$  zu Baseline weisen unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant weniger allgemeine Erkrankungen und Beschwerden am Verabreichungsort auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,49 [0,29;0,81],  $p=0,0052$ ). In der Gruppe der Patienten, die zu Baseline  $>350$  Zellen/ $\mu\text{L}$  aufwiesen, besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,98 [0,62;1,53],  $p=0,9214$ ).

### ***Fatigue***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,84$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,2017$ ;  $I^2=38,6\%$ ), Rasse ( $p=0,85$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,79$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,90$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,94$ ;  $I^2=0\%$ ).

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,81$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,62$ ;  $I^2=0\%$ ), Rasse ( $p=0,36$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,21$ ;  $I^2=36,4\%$ ), CD4-Zellzahl zu Baseline ( $p=0,70$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,30$ ;  $I^2=8,0\%$ ).

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,46$ ;  $I^2=0\%$ ), Rasse ( $p=0,47$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,48$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,87$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,71$ ;  $I^2=0\%$ ). Es besteht allerdings ein Beleg für Effektmodifikation durch Geschlecht ( $p=0,02$ ;  $I^2=80,1\%$ ). Da allerdings in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden kann (OR [95%-KI]: Männer: 0,71 [0,45;1,12],  $p=0,1396$ ; Frauen: 4,78 [0,96;23,83],  $p=0,0563$ ), wirkt sich diese Interaktion nicht auf das Gesamtfazit aus.

### ***Erkrankungen des Nervensystems (SOC)***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,46$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,24$ ;  $I^2=27,4\%$ ), Rasse ( $p=0,50$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,39$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,64$ ;  $I^2=0\%$ ). Es gibt jedoch einen

Hinweis für Effektmodifikation durch Therapieadhärenz ( $p=0,0532$ ;  $I^2=73,2\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhären waren, ergibt sich das OR [95%-KI] zu 0,94 [0,52;1,72] mit  $p=0,84$ , und für Patienten, die eine Therapieadhärenz  $\geq 95\%$  aufwiesen, beträgt das OR [95%-KI] 0,47 [0,33;0,67] mit  $p<0,0001$ . Es besteht somit in der Gruppe der Patienten, die die Studienmedikation  $\geq 95\%$  adhären einnahmen, ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Dort traten Erkrankungen des Nervensystems statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF.

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,88$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,92$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,63$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,44$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,80$ ;  $I^2=0\%$ ). Es besteht ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,14$ ;  $I^2=53,0\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 1,31 [0,34;5,06] mit  $p=0,6929$  und in der Gruppe der nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,17 [0,01;1,88] mit  $p=0,1472$ . Da aber in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich diese Interaktion nicht auf das Gesamtfazit aus.

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,33$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,22$ ;  $I^2=32,8\%$ ), Rasse ( $p=0,78$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,35$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,94$ ;  $I^2=0\%$ ). Lediglich die Subgruppenanalyse, stratifiziert nach Therapieadhärenz nach 96 Wochen, ergibt einen Hinweis auf Effektmodifikation ( $p=0,18$ ;  $I^2=44,8\%$ ). In der Gruppe der Patienten, die eine Therapieadhärenz  $<95\%$  aufwiesen, ergibt sich das OR [95%-KI] zu 0,84 [0,46;1,54] mit  $p=0,5772$ . Und in der Gruppe der Patienten, die zu  $\geq 95\%$  adhären waren, beträgt das OR [95%-KI] 0,52 [0,36;0,75] mit  $p=0,0004$ , was einen statistisch signifikanten Unterschied in der Vermeidung von Erkrankungen des Nervensystems zugunsten von EVG/COBI/FTC/TDF anzeigt.

### **Schwindel**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,50$ ;  $I^2=0\%$ ), Geschlecht ( $p=1,00$ ;  $I^2=0\%$ ), Rasse ( $p=0,87$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,70$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,45$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,74$ ;  $I^2=0\%$ ).

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Rasse ( $p=0,99$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,87$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,97$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,94$ ;  $I^2=0\%$ ). Subgruppenanalysen, stratifiziert nach Alter und Geschlecht, können aufgrund zu geringer Ereignisraten nicht durchgeführt werden.

Es können zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,71$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,92$ ;  $I^2=0\%$ ), Rasse ( $p=0,76$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,84$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,68$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,57$ ;  $I^2=0\%$ ) gefunden werden.

### ***Kopfschmerzen***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,38$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,31$ ;  $I^2=3,1\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,45$ ;  $I^2=0\%$ ). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Geschlecht ( $p=0,18$ ;  $I^2=43,3\%$ ). In der Gruppe der männlichen Patienten beträgt das OR [95%-KI] 1,66 [1,01;2,73] mit  $p=0,04$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der weiblichen Patienten ergibt sich das OR [95%-KI] zu 0,68 [0,20;2,30] mit  $p=0,54$ . Ein weiterer Hinweis auf Effektmodifikation besteht durch Rasse ( $p=0,19$ ;  $I^2=43,0\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 1,90 [1,06;3,39] mit  $p=0,03$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF anzeigt, und in der Gruppe der nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,75 [0,22;2,61] mit  $p=0,65$ . Außerdem besteht ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,18$ ;  $I^2=43,9\%$ ). In der Gruppe der zu  $<95\%$  adhärennten Patienten beträgt das OR [95%-KI] 2,72 [0,93;7,96] mit  $p=0,07$  und in der Gruppe der Patienten die zu  $\geq 95\%$  adhärennt waren, ergibt sich zu 1,22 [0,75;1,99] mit  $p=0,43$ . 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 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ( $p=0,25$ ;  $I^2=23,8\%$ ), Rasse ( $p=0,22$ ;  $I^2=33,5\%$ ), Ausgangsviruslast ( $p=0,65$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,75$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,93$ ;  $I^2=0\%$ ). Allerdings besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,14$ ;  $I^2=55,0\%$ ). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 1,68 [0,30;9,47] mit  $p=0,5564$  und in der Gruppe der älteren Patienten ergibt sich das OR [95%-KI] zu 0,16 [0,01;2,11] mit  $p=0,1620$ . 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 ( $p=0,61$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,23$ ;  $I^2=31,0\%$ ), Rasse ( $p=0,56$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,47$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,75$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,24$ ;  $I^2=26,5\%$ ).

### ***Psychiatrische Erkrankungen (SOC)***

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,87$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,61$ ;  $I^2=0\%$ ), Rasse ( $p=0,78$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,85$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,79$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,35$ ;  $I^2=0\%$ ).

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,45$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,71$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,44$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,35$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,33$ ;  $I^2=0\%$ ). Allerdings

besteht ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,17$ ;  $I^2=47,2\%$ ). In der Gruppe der kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,32 [0,09;1,08] mit  $p=0,0671$  und in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 1,71 [0,22;13,41] mit  $p=0,6075$ . 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 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,63$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,41$ ;  $I^2=0\%$ ), Rasse ( $p=0,84$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,50$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,66$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,86$ ;  $I^2=0\%$ ).

### **Abnorme Träume**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,64$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,36$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,61$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,89$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,88$ ;  $I^2=0\%$ ). Allerdings besteht ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,13$ ;  $I^2=56,0\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 0,26 [0,08;0,87] mit  $p=0,03$  und in der Gruppe der nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,73 [0,40;1,33] mit  $p=0,30$ . Somit treten bei kaukasischen Patienten unter eine Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener abnorme Träume auf als unter einer Behandlung mit EFV/FTC/TDF.

Zu Woche 60 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Ausgangsviruslast ( $p=0,66$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,58$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,31$ ;  $I^2=4,2\%$ ). Allerdings besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,08$ ;  $I^2=67,0\%$ ). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,09 [0,02;0,51] mit  $p=0,0064$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der  $\geq 40$ -jährigen Patienten ergibt sich das OR [95%-KI] zu 1,29 [0,11;15,00] mit  $p=0,8411$ . Ein weiterer Hinweis auf Effektmodifikation besteht durch Geschlecht ( $p=0,19$ ;  $I^2=42,4\%$ ). In der Gruppe der männlichen Patienten beträgt das OR [95%-KI] 0,16 [0,04;0,63] mit  $p=0,0085$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF zeigt. Für die Frauen beträgt das OR [95%-KI] 2,14 [0,06;77,54] mit  $p=0,6772$ . Außerdem besteht ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,14$ ;  $I^2=53,6\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 0,10 [0,02;0,56] mit  $p=0,0089$ . In der Gruppe der nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 1,00 [0,08;12,56] mit  $p=1,0000$ . Somit treten abnorme Träume bei kaukasischen Patienten unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF.

Es gibt zu Woche 96 keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,36$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,71$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,76$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,69$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,69$ ;  $I^2=0\%$ ). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,14$ ;  $I^2=54,4\%$ ). In der Gruppe der

kaukasischen Patienten wurden im EVG/COBI/FTC/TDF-Arm statistisch signifikant weniger abnorme Träume beobachtet als im EFV/FTC/TDF-Arm (OR [95%-KI]: 0,38 [0,24;0,62],  $p=0,0001$ ), während unter den nicht-kaukasischen Patienten kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen beobachtet werden kann (OR [95%-KI]: 0,69 [0,37;1,26],  $p=0,2270$ ).

### **Schlaflosigkeit**

Zu Woche 48 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,57$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,52$ ;  $I^2=0\%$ ), Rasse ( $p=0,84$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,84$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,69$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,57$ ;  $I^2=0\%$ ).

Zu Woche 60 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,56$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,78$ ;  $I^2=0\%$ ), Rasse ( $p=0,47$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,91$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,69$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,40$ ;  $I^2=0\%$ ).

Zu Woche 96 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,85$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,42$ ;  $I^2=0\%$ ), Rasse ( $p=0,45$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,51$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,38$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,61$ ;  $I^2=0\%$ ).

### **Depression**

Zu Woche 48 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ( $p=0,25$ ;  $I^2=25,8\%$ ), Rasse ( $p=0,94$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,74$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,88$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,40$ ;  $I^2=0\%$ ). Allerdings besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,15$ ;  $I^2=51,0\%$ ). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 1,17 [0,60;2,26] mit  $p=0,6485$  und für die mindestens 40-jährigen Patienten ergibt sich ein OR [95%-KI] von 0,56 [0,27;1,19] mit  $p=0,1324$ . Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Da in Studie S104 nicht mindestens 10% der Patienten in einem der Behandlungsarme Depressionen hatten, werden für diesen Endpunkt zu Woche 60 keine Subgruppenanalysen durchgeführt.

Zu Woche 96 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Rasse ( $p=0,82$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,35$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,96$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,54$ ;  $I^2=0\%$ ). Es liegt jedoch ein Hinweis auf Effektmodifikation durch Alter vor ( $p=0,15$ ;  $I^2=52,7\%$ ). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht (OR [95%-KI] <40 Jahre: 1,12 [0,63;2,00],  $p=0,7014$ ;  $\geq 40$  Jahre: 0,57 [0,28;1,15],  $p=0,1167$ ), wirkt sich diese Interaktion nicht auf das Gesamtfazit aus. Des Weiteren besteht ein Hinweis auf Effektmodifikation durch Geschlecht ( $p=0,09$ ;  $I^2=65,5\%$ ). Da jedoch auch hier in keiner Subgruppe ein statistisch signifikanter Unterschied besteht (OR [95%-KI]: Männer: 0,73

[0,46;1,17],  $p=0,1952$ ; Frauen: 2,67 [0,65;10,94],  $p=0,1733$ ), ist diese Interaktion nicht relevant für das Gesamtfazit.

### **Angst**

Zu Woche 48 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,50$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,51$ ;  $I^2=0\%$ ), Rasse ( $p=0,41$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,53$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,97$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,27$ ;  $I^2=18,6\%$ ).

Zu Woche 60 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,50$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,51$ ;  $I^2=0\%$ ), Rasse ( $p=0,41$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,53$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,97$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,29$ ;  $I^2=10,1\%$ ).

Da in Studie S102 zu Woche 96 Angst nicht bei mindestens 10% der Patienten in einem der Behandlungsarme auftrat, werden hierfür keine Subgruppenanalysen durchgeführt.

### **Erkrankungen der Haut und des Unterhautzellgewebes (SOC)**

Zu Woche 48 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,32$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,45$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,51$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,83$ ;  $I^2=0\%$ ). Allerdings besteht ein Beleg für Effektmodifikation durch Rasse ( $p=0,04$ ;  $I^2=76,6\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 0,47 [0,32;0,69] mit  $p=0,0001$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF zeigt. In der Gruppe der nicht kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,94 [0,55;1,60] mit  $p=0,83$ . Ein Hinweis auf Effektmodifikation besteht durch Therapieadhärenz ( $p=0,17$ ;  $I^2=46,0\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhären waren, beträgt das OR [95%-KI] 0,99 [0,53;1,85] mit  $p=0,97$ , und in der Gruppe der zu  $\geq 95\%$  adhären Patienten ergibt sich das OR [95%-KI] zu 0,54 [0,30;0,98] mit  $p=0,04$ . Somit besteht in der Gruppe der hochadhären Patienten ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF.

Zu Woche 60 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,34$ ;  $I^2=0\%$ ), Rasse ( $p=0,44$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,71$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,51$ ;  $I^2=0\%$ ). Es besteht allerdings ein Hinweis auf Effektmodifikation durch Geschlecht ( $p=0,12$ ;  $I^2=57,9\%$ ). In der Gruppe der männlichen Patienten beträgt das OR [95%-KI] 0,61 [0,20;1,84] mit  $p=0,3786$  und in der Gruppe der weiblichen Patienten ergibt sich das OR [95%-KI] zu 11,67 [0,32;422,14] mit  $p=0,1797$ . 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 Ausgangsviruslast ( $p=0,14$ ;  $I^2=54,6\%$ ). In der Gruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 0,53 [0,16;1,70] mit  $p=0,2845$ , und in der Gruppe der Patienten mit Ausgangsviruslast  $>100.000$  Kopien/mL ergibt sich das OR [95%-KI] zu 6,60 [0,29;149,77] mit  $p=0,2361$ . Da jedoch auch hier 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 ( $p=0,34$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,78$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,76$ ;  $I^2=0\%$ ) oder CD4-Zellzahl zu Baseline ( $p=0,92$ ;  $I^2=0\%$ ). Es besteht jedoch ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,07$ ;  $I^2=70,2\%$ ). In der Gruppe der kaukasischen Patienten besteht ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,53 [0,36;0,78],  $p=0,0013$ ), der in der Gruppe der nicht-kaukasischen Patienten nicht mehr beobachtet werden kann (OR [95%-KI]: 0,96 [0,58;1,61],  $p=0,8896$ ). Des Weiteren besteht ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,06$ ;  $I^2=71,4\%$ ). Bei Patienten, die zu Woche 96 eine Therapieadhärenz  $<95\%$  aufwiesen, besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,08 [0,58;2,00],  $p=0,8052$ ). In der Gruppe der Patienten die zu  $\geq 95\%$  adhärenz waren, wurden statistisch signifikant weniger Erkrankungen der Haut und des Unterhautzellgewebes im EVG/COBI/FTC/TDF-Arm beobachtet als im EFV/FTC/TDF-Arm (OR [95%-KI]: 0,55 [0,38;0,78],  $p=0,0010$ ).

### **Hautausschlag**

Zu Woche 48 bestehen keine Hinweise oder Belege auf Effektmodifikation durch Geschlecht ( $p=0,57$ ;  $I^2=0\%$ ), Rasse ( $p=0,55$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,36$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,85$ ;  $I^2=0\%$ ) oder Therapieadhärenz ( $p=0,95$ ;  $I^2=0\%$ ). Es gibt jedoch einen Hinweis auf Effektmodifikation durch Alter ( $p=0,17$ ;  $I^2=48,1\%$ ). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,65 [0,35;1,24] mit  $p=0,19$  und in der Gruppe der mindestens 40-jährigen Patienten ergibt sich das OR [95%-KI] 0,31 [0,13;0,72] mit  $p=0,006$ . Somit besteht in der Gruppe der  $\geq 40$ -jährigen Patienten ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF.

Zu Woche 60 gibt es keine Hinweise oder Belege auf Effektmodifikation durch Alter ( $p=0,24$ ;  $I^2=26,8\%$ ), Rasse ( $p=0,90$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,22$ ;  $I^2=33,3\%$ ) oder Therapieadhärenz ( $p=0,71$ ;  $I^2=0\%$ ). Die Subgruppenanalyse stratifiziert nach Geschlecht war aufgrund zu geringer Ereignisraten nicht durchführbar. Es gibt einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,09$ ;  $I^2=64,7\%$ ). In der Gruppe der Patienten mit Ausgangsviruslast  $\leq 100.000$  Kopien/mL beträgt das OR [95%-KI] 0,20 [0,03;1,22] mit  $p=0,0809$  und in der Gruppe von Patienten, die zu Baseline eine Viruslast  $>100.000$  Kopien/mL aufwiesen, ergibt sich das OR [95%-KI] zu 4,53 [0,19;105,84] mit  $p=0,3475$ . 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 ( $p=0,46$ ;  $I^2=0\%$ ), Geschlecht ( $p=0,58$ ;  $I^2=0\%$ ), Rasse ( $p=0,44$ ;  $I^2=0\%$ ), Ausgangsviruslast ( $p=0,88$ ;  $I^2=0\%$ ), CD4-Zellzahl zu Baseline ( $p=0,29$ ;  $I^2=11,8\%$ ) oder Therapieadhärenz ( $p=0,36$ ;  $I^2=0\%$ ).



Generell sind die 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, nicht nur eine Tablette täglich einnehmen mussten, wie für EVG/COBI/FTC/TDF und EFV/FTC/TDF in der Praxis vorgesehen, sondern zusätzlich eine Placebo-Tablette, verteilt über zwei Dosierungen am Tag. Verschiedene klinische Daten zeigen jedoch einen Unterschied zwischen einmal und zweimal täglicher Gabe [71, 72], was bedeutet, dass die Adhärenz-Daten in diesen Studien somit nicht auf die Praxis übertragbar sind. Aus diesem Grund wurden die Subgruppenanalysen zwar vollständig dargestellt, finden jedoch keine weitere Berücksichtigung in der Bewertung und Nutzenableitung.

Im Folgenden werden die Ergebnisse der signifikanten Subgruppenanalysen detailliert dargestellt.

#### 4.3.1.3.2.1 Virologisches Ansprechen (Snapshot-Algorithmus; in Studie S102 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 4-56: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Alter (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
	N	n (%)	N	n (%)	
<b>&lt;40 Jahre</b>					
S102	200	171 (85,5)	197	169 (85,8)	1,00 [0,57;1,75], p=0,9351
S104	31	28 (90,3)	16	12 (75,0)	3,11 [0,60;16,08], p=0,1757
<b>≥40 Jahre</b>					
S102	148	134 (90,5)	155	127 (81,9)	2,07 [1,04;4,12], p=0,0329
S104	17	16 (94,1)	7	7 (100,0)	0,73 [0,03;20,18], p=0,8545
<b>p-Wert Interaktionstest<sup>b,c</sup>:</b>					0,50
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					0%
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet <sup>c</sup> Chi Quadrat Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.1.1.2 Virologisches Ansprechen (Snapshot) stratifiziert nach Geschlecht**

Tabelle 4-57: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Geschlecht (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
	N	n (%)	N	n (%)	
<b>Männlich</b>					
S102	307	270 (87,9)	316	266 (84,2)	1,39 [0,88;2,19], p=0,1758
S104	44	41 (93,2)	21	17 (81,0)	3,22 [0,65;15,93], p=0,1525
<b>Weiblich</b>					
S102	41	35 (85,4)	36	30 (83,3)	1,13 [0,33;3,88], p=0,8063
S104	4	3 (75,0)	2	2 (100,0)	0,47 [0,01;16,89], p=0,6772
<b>p-Wert Interaktionstest<sup>b,c</sup>:</b>					0,61
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					0%
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet <sup>c</sup> Chi Quadrat Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.1.1.3 Virologisches Ansprechen (Snapshot) stratifiziert nach Rasse**

Tabelle 4-58: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Rasse (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
	N	n (%)	N	n (%)	
<b>Kaukasisch</b>					
S102	214	191 (89,3)	227	199 (87,7)	1,18 [0,66;2,12], p=0,6027
S104	33	30 (90,9)	18	15 (83,3)	2,00 [0,36;11,13], p=0,4286
<b>Nicht Kaukasisch</b>					
S102	134	114 (85,1)	125	97 (77,6)	1,65 [0,87;3,11], p=0,1240
S104	15	14 (93,3)	5	4 (80,0)	3,50 [0,18;69,34], p=0,4109
<b>p-Wert Interaktionstest<sup>b,c</sup>:</b>					0,45
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					0%
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet <sup>c</sup> Chi Quadrat Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.1.1.4 Virologisches Ansprechen (Snapshot) stratifiziert nach Ausgangsviruslast**

Tabelle 4-59: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Ausgangsviruslast (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
	N	n (%)	N	n (%)	
<b>≤100.000 Kopien/mL</b>					
S102	230	206 (89,6)	236	201 (85,2)	1,49 [0,86;2,60], p=0,1555
S104	37	35 (94,6)	18	15 (83,3)	3,50 [0,53;23,14], p=0,1936
<b>&gt;100.000 Kopien/mL</b>					
S102	118	99 (83,9)	116	95 (81,9)	1,15 [0,58;2,28], p=0,6844
S104	11	9 (81,8)	5	4 (80,0)	1,13 [0,08;16,31], p=0,9312
<b>p-Wert Interaktionstest<sup>b,c</sup>:</b>					0,45
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					0%
<sup>a</sup> Logistisches Regressionsmodell mit Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet <sup>c</sup> Chi Quadrat Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

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

Tabelle 4-60: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
	N	n (%)	N	n (%)	
<b>≤350 Zellen/μL</b>					
S102	155	129 (83,2)	147	123 (83,7)	0,96 [0,52;1,77], p=0,9167
S104	24	23 (95,8)	8	6 (75,0)	7,67 [0,59;99,48], p=0,1193
<b>&gt;350 Zellen/μL</b>					
S102	193	176 (91,2)	205	173 (84,4)	1,93 [1,03;3,60], p=0,0415
S104	24	21 (87,5)	15	13 (86,7)	1,08 [0,16;7,33], p=0,9396
<b>p-Wert Interaktionstest<sup>b,c</sup>:</b>					0,99
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					0%
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet <sup>c</sup> Chi Quadrat Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.1.1.6 Virologisches Ansprechen (Snapshot) stratifiziert nach Therapieadhärenz**

Tabelle 4-61: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Therapieadhärenz (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>
	N	n (%)	N	n (%)	
<b>&lt;95%</b>					
S102	86	68 (79,1)	89	67 (75,3)	1,29 [0,63;2,62], p=0,5511
S104	13	12 (92,3)	5	4 (80,0)	3,00 [0,15;59,89], p=0,4720
<b>≥95%</b>					
S102	260	237 (91,2)	260	229 (88,1)	1,39 [0,78;2,45], p=0,2517
S104	33	32 (97,0)	17	15 (88,2)	4,27 [0,36;50,83], p=0,2511
<b>p-Wert Interaktionstest<sup>b,c</sup>:</b>					0,78
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					0%
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet <sup>c</sup> Chi Quadrat Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

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

Tabelle 4-62: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Alter (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/ TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>	RD (%) [95%-KI] <sup>c</sup>
	N	n (%)	N	n (%)		
<40 Jahre	200	165 (82,5)	197	162 (82,2)	1,02 [0,61;1,72], p=0,9445	0,4 [-7,2;7,9]
≥40 Jahre	148	128 (86,5)	155	125 (80,6)	1,53 [0,82;2,84], p=0,1729	5,8 [-2,6;14,3]
<b>p-Wert Interaktionstest<sup>d</sup>:</b>					0,33	
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					0,2%	
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet (Chi-Quadrat Test) <sup>c</sup> RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet <sup>d</sup> 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 4-63: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Geschlecht (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p- Wert <sup>b</sup>	RD (%) [95%-KI] <sup>c</sup>
	N	n (%)	N	n (%)		
Männlich	307	260 (84,7)	316	259 (82,0)	1,22 [0,80;1,86], p=0,3617	2,8 [-3,1;8,7]
Weiblich	41	33 (80,5)	36	28 (77,8)	1,17 [0,39;3,52], p=0,7701	3,1 [-16,3;22,5]
<b>p-Wert Interaktionstest<sup>d</sup>:</b>						0,94
<b>I<sup>2</sup> (zusätzlich berechnet):</b>						0%
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet (Chi-Quadrat Test) <sup>c</sup> RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet <sup>d</sup> 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 4-64: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Rasse (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/ TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p- Wert <sup>b</sup>	RD (%) [95%-KI] <sup>c</sup>
	N	n (%)	N	n (%)		
Kaukasisch	214	182 (85,0)	227	193 (85,0)	1,00 [0,59;1,70], p=0,9942	0,1 [-6,7;6,8]
Nicht-kaukasisch	134	111 (82,8)	125	94 (75,2)	1,59 [0,87;2,91], p=0,1324	7,6 [-2,4;17,7]
<b>p-Wert Interaktionstest<sup>d</sup>:</b>						0,26
<b>I<sup>2</sup> (zusätzlich berechnet):</b>						22,2%
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet (Chi-Quadrat Test) <sup>c</sup> RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet <sup>d</sup> 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 4-65: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Ausgangsviruslast (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/ TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p- Wert <sup>b</sup>	RD (%) [95%- KI] <sup>c</sup>
	N	n (%)	N	n (%)		
≤100.000 Kopien/mL	230	197 (85,7)	236	191 (80,9)	1,41 [0,86;2,30], p=0,1736	4,7 [-2,0;11,5]
>100.000 Kopien/mL	118	96 (81,4)	116	96 (82,8)	0,91 [0,47;1,77], p=0,7799	-1,4 [-11,2;8,4]
<b>p-Wert Interaktionstest<sup>d</sup>:</b>					0,30	
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					5,9%	
<sup>a</sup> Logistisches Regressionsmodell mit Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet (Chi-Quadrat Test) <sup>c</sup> RD und 95%-KI basierend auf Normalapproximation berechnet <sup>d</sup> 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 4-66: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/ TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p- Wert <sup>b</sup>	RD (%) [95%-KI] <sup>c</sup>
	N	n (%)	N	n (%)		
≤350 Zellen/μL	155	124 (80,0)	147	118 (80,3)	0,98 [0,56;1,73], p=0,9528	-0,3 [-9,4;8,8]
>350 Zellen/μL	193	169 (87,6)	205	169 (82,4)	1,50 [0,86;2,62], p=0,1550	5,1 [-2,0;12,2]
<b>p-Wert Interaktionstest<sup>d</sup>:</b>					0,30	
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					7,8%	
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet (Chi-Quadrat Test) <sup>c</sup> RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet. <sup>d</sup> 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 4-67: Subgruppenanalyse für virologisches Ansprechen (Snapshot), stratifiziert nach Therapieadhärenz (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/ TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>b</sup>	RD (%) [95%-KI] <sup>c</sup>
	N	n (%)	N	n (%)		
<95%	90	67 (74,4)	86	54 (62,8)	1,73 [0,91;3,29], p=0,0969	11,1 [-2,5;24,8]
≥95%	256	226 (88,3)	263	233 (88,6)	0,97 [0,57;1,66], p=0,9115	-0,3 [-5,9;5,3]
<b>p-Wert Interaktionstest<sup>d</sup>:</b>					0,18	
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					44,8%	
<sup>a</sup> Logistisches Regressionsmodell mit Baseline HIV-1-RNA Level, Subgruppe, Behandlung und Interaktion zwischen Behandlung und Subgruppe <sup>b</sup> zusätzlich berechnet (Chi-Quadrat Test) <sup>c</sup> RD und 95%-KI basierend auf Baseline HIV-1-RNA Schicht-adjustierten Mantel Haenszel Anteilen und Normalapproximation berechnet <sup>d</sup> 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 Rasse**

Tabelle 4-68: Subgruppenanalyse für virologisches Ansprechen (TLOVR), stratifiziert nach Rasse (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	214	171 (79,9)	227	184 (81,1)	0,93 [0,58;1,49], p=0,7605
Nicht-kaukasisch	134	105 (78,4)	125	88 (70,4)	1,52 [0,87;2,67], p=0,1432
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,19
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					42,4%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, RD = Risikodifferenz					



#### 4.3.1.3.2.2.1.2 Virologisches Ansprechen (TLOVR) stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-69: Subgruppenanalyse für virologisches Ansprechen (TLOVR), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤350 Zellen/μL	155	112 (72,3)	147	112 (76,2)	0,81 [0,49;1,37], p=0,4355
>350 Zellen/μL	193	164 (85,0)	205	160 (78,0)	1,59 [0,95;2,66], p=0,0774
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,07
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					69,1%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Veränderung der CD4-Zellzahl

##### 4.3.1.3.2.3.1 Veränderung der CD4-Zellzahl zu Woche 96

##### 4.3.1.3.2.3.1.1 Veränderung der CD4-Zellzahl stratifiziert nach Therapieadhärenz

Tabelle 4-70: Subgruppenanalyse für Veränderung der CD4-Zellzahl, stratifiziert nach Therapieadhärenz (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		MWD <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	MW der Veränderung (SD)	N	MW der Veränderung (SD)	
<95%	69	317 (217,3)	61	245 (168,6)	72,00 [5,52;138,48], p=0,0386
≥95%	238	288 (212,1)	241	280 (194,4)	8,00 [-28,45;44,45], p=0,6671
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,10
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					63,5%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test  KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio, SD = Standardabweichung					

**4.3.1.3.2.4 Gesamtrate unerwünschter Ereignisse****4.3.1.3.2.4.1 Gesamtrate unerwünschter Ereignisse zu Woche 96****4.3.1.3.2.4.1.1 Gesamtrate unerwünschter Ereignisse stratifiziert nach Alter**

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

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<40 Jahre	200	194 (97,0)	197	189 (95,9)	1,37 [0,47;4,02], p=0,5681
≥40 Jahre	148	143 (96,6)	155	153 (98,7)	0,37 [0,07;1,96], p=0,2441
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,1979
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					39,7%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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					

**4.3.1.3.2.5 Gesamtrate schwerwiegender unerwünschter Ereignisse****4.3.1.3.2.5.1 Gesamtrate schwerwiegender unerwünschter Ereignisse zu Woche 96****4.3.1.3.2.5.1.1 Gesamtrate schwerwiegender unerwünschter Ereignisse stratifiziert nach Rasse**

Tabelle 4-72: Subgruppenanalyse für Gesamtrate schwerwiegender unerwünschter Ereignisse, stratifiziert nach Rasse (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	214	34 (15,9)	227	16 (7,0)	2,49 [1,33;4,66], p=0,0043
Nicht-kaukasisch	134	22 (16,4)	125	17 (13,6)	1,25 [0,63;2,48], p=0,5268
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,14
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					53,0%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test  KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.6 Therapieabbrüche aufgrund unerwünschter Ereignisse****4.3.1.3.2.6.1 Therapieabbrüche aufgrund unerwünschter Ereignisse zu Woche 48****4.3.1.3.2.6.1.1 Therapieabbrüche aufgrund unerwünschter Ereignisse stratifiziert nach Rasse**

Tabelle 4-73: Subgruppenanalyse für Therapieabbrüche aufgrund unerwünschter Ereignisse, stratifiziert nach Rasse (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>Kaukasisch</b>					
S102	214	11 (5,1)	227	8 (3,5)	1,48 [0,58;3,76], p=0,4062
S104	33	0 (0,0)	18	0 (0,0)	n.b.
<b>Nicht-kaukasisch</b>					
S102	134	2 (1,5)	125	10 (8,0)	0,17 [0,04;0,81], p=0,0260
S104	15	0 (0,0)	5	1 (20,0)	0,10 [0,00;2,81], p=0,1742
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,009
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					85,4%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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					

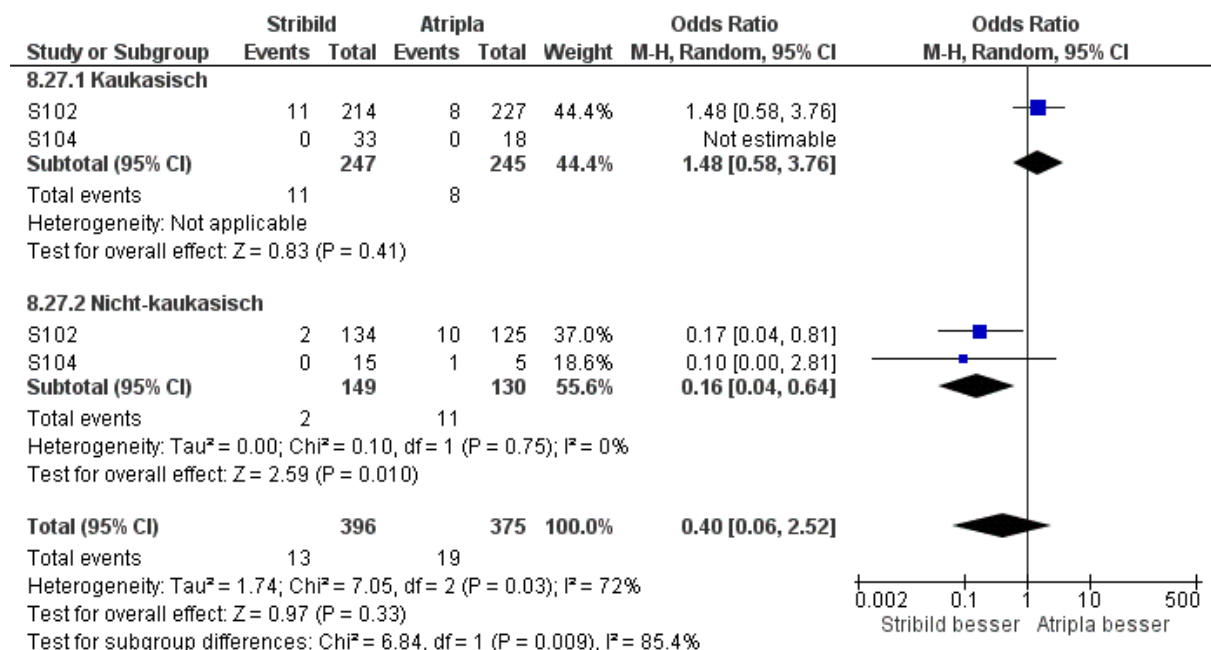


Abbildung 34: Meta-Analyse für Therapieabbrüche aufgrund unerwünschter Ereignisse, stratifiziert nach Rasse aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

### 4.3.1.3.2.6.2 Therapieabbrüche aufgrund unerwünschter Ereignisse zu Woche 96

#### 4.3.1.3.2.6.2.1 Therapieabbrüche aufgrund unerwünschter Ereignisse stratifiziert nach Rasse

Tabelle 4-74: Subgruppenanalyse für Therapieabbrüche aufgrund unerwünschter Ereignisse, stratifiziert nach Rasse (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	214	14 (6,5)	227	13 (5,7)	1,15 [0,53;2,51], p=0,7214
Nicht-kaukasisch	134	3 (2,2)	125	11 (8,8)	0,24 [0,06;0,87], p=0,0303
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,04
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					76,0%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

#### 4.3.1.3.2.7 Unerwünschte Ereignisse Grad 3-4

##### 4.3.1.3.2.7.1 Unerwünschte Ereignisse Grad 3-4 zu Woche 48

##### 4.3.1.3.2.7.1.1 Unerwünschte Ereignisse Grad 3-4 stratifiziert nach Alter

Tabelle 4-75: Subgruppenanalyse für Unerwünschte Ereignisse Grad 3-4, stratifiziert nach Alter (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>&lt;40 Jahre</b>					
S102	200	27 (13,5)	197	16 (8,1)	1,77 [0,92;3,39], p=0,0877
S104	31	0 (0,0)	16	0 (0,0)	n.b.
<b>≥40 Jahre</b>					
S102	148	18 (12,2)	155	23 (14,8)	0,79 [0,41;1,54], p=0,4966
S104	17	2 (11,8)	7	2 (28,6)	0,33 [0,04;3,03], p=0,3290
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,06
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					71,5%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> 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					

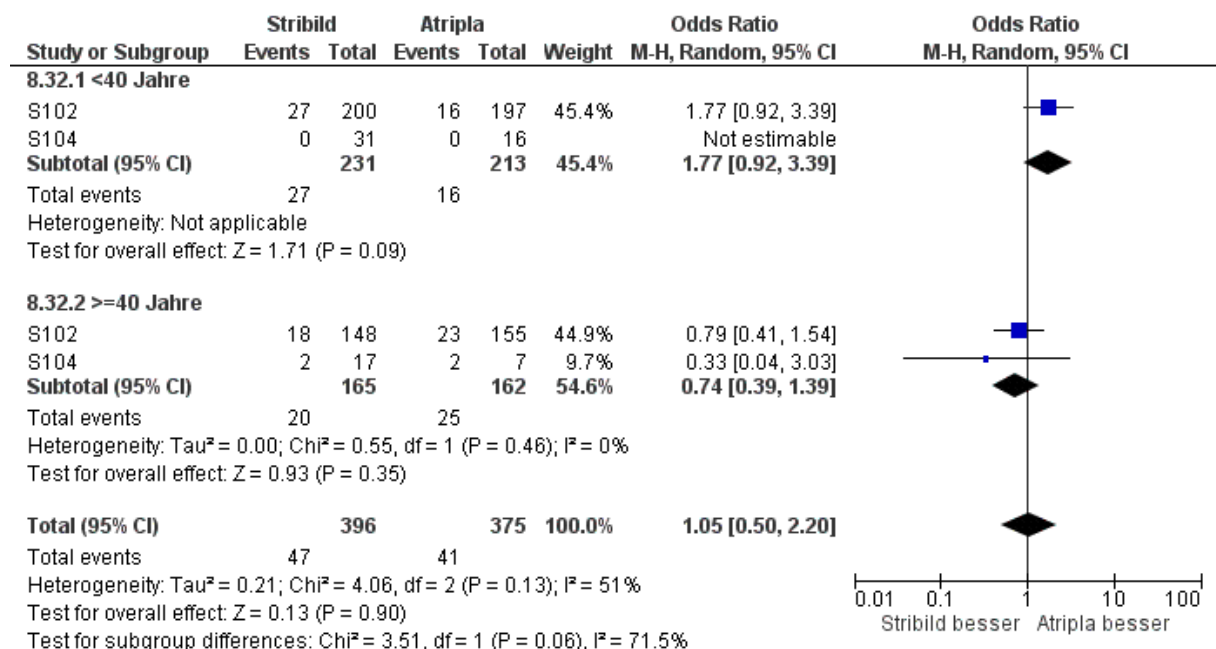


Abbildung 35: Meta-Analyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Alter, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

#### 4.3.1.3.2.7.1.2 Unerwünschte Ereignisse Grad 3-4 stratifiziert nach Rasse

Tabelle 4-76: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Rasse (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>Kaukasisch</b>					
S102	214	25 (11,7)	227	17 (7,5)	1,63 [0,86;3,12], p=0,1367
S104	33	1 (3,0)	18	1 (5,6)	0,53 [0,03;9,03], p=0,6617
<b>Nicht-kaukasisch</b>					
S102	134	20 (14,9)	125	22 (17,6)	0,82 [0,42;1,59], p=0,5599
S104	15	1 (6,7)	5	1 (20,0)	0,29 [0,01;5,66], p=0,4109
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,14
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					54,3%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

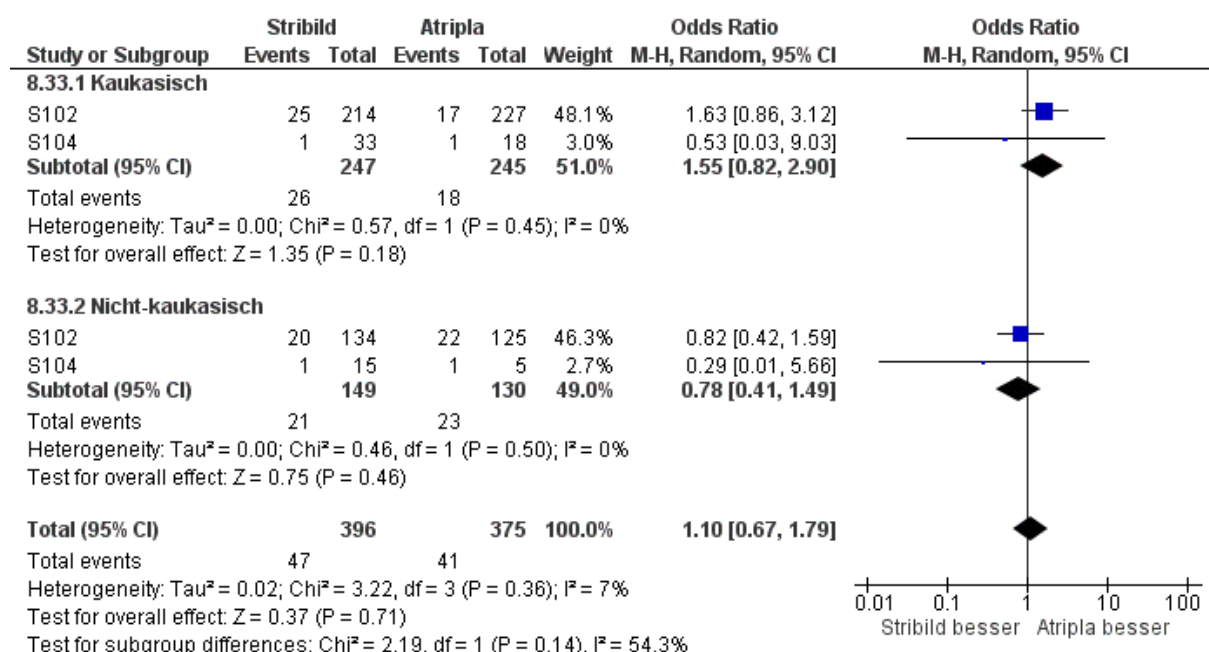


Abbildung 36: Meta-Analyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**4.3.1.3.2.7.2 Unerwünschte Ereignisse Grad 3-4 zu Woche 96****4.3.1.3.2.7.2.1 Unerwünschte Ereignisse Grad 3-4 stratifiziert nach Rasse**

Tabelle 4-77: Subgruppenanalyse für unerwünschte Ereignisse Grad 3-4, stratifiziert nach Rasse (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	214	34 (15,9)	227	24 (10,6)	1,60 [0,91;2,80], p=0,1009
Nicht-kaukasisch	134	27 (20,1)	125	27 (21,6)	0,92 [0,50;1,67], p=0,7740
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,18
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					43,4%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.8 Neurologische & psychiatrische Ereignisse****4.3.1.3.2.8.1 Neurologische & psychiatrische Ereignisse zu Woche 48****4.3.1.3.2.8.1.1 Neurologische & psychiatrische Ereignisse stratifiziert nach Therapieadhärenz**

Tabelle 4-78: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Therapieadhärenz (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>&lt;95%</b>					
S102	86	44 (51,2)	89	55 (61,8)	0,65 [0,35;1,18], p=0,1568
S104	12	6 (50,0)	5	3 (60,0)	0,67 [0,08;5,54], p=0,7074
<b>≥95%</b>					
S102	260	105 (40,4)	260	164 (63,1)	0,40 [0,28;0,56], p<0,00001
S104	34	14 (41,2)	17	10 (58,8)	0,49 [0,15;1,60], p=0,2372
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,16
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					48,3%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

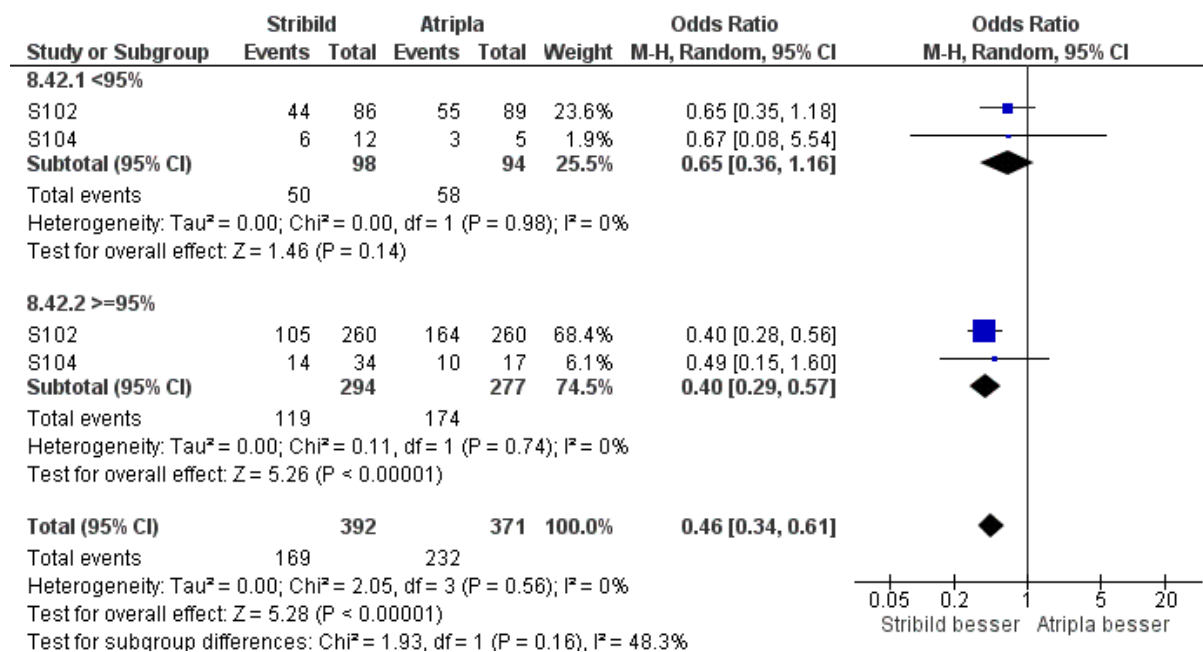


Abbildung 37: Meta-Analyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Therapieadhärenz, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

#### 4.3.1.3.2.8.2 Neurologische & psychiatrische Ereignisse zu Woche 60

##### 4.3.1.3.2.8.2.1 Neurologische & psychiatrische Ereignisse stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-79: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤350 Zellen/μL	24	12 (50,0)	8	3 (37,5)	1,67 [0,32;8,59], p=0,5415
>350 Zellen/μL	24	8 (33,3)	15	10 (66,7)	0,25 [0,06;0,98], p=0,0471
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,08
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					67,0%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					



### 4.3.1.3.2.8.3 Neurologische & psychiatrische Ereignisse zu Woche 96

#### 4.3.1.3.2.8.3.1 Neurologische & psychiatrische Ereignisse stratifiziert nach Alter

Tabelle 4-80: Subgruppenanalyse für neurologische & psychiatrische Ereignisse, stratifiziert nach Alter (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<40 Jahre	200	100 (50,0)	197	128 (65,0)	0,54 [0,36;0,81], p=0,0027
≥40 Jahre	148	62 (41,9)	155	104 (67,1)	0,35 [0,22;0,56], p<0,0001
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,18
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					44,2%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Hautausschläge

#### 4.3.1.3.2.9.1 Hautausschläge zu Woche 48

##### 4.3.1.3.2.9.1.1 Hautausschläge stratifiziert nach Alter

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

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>&lt;40 Jahre</b>					
S102	200	39 (19,5)	197	52 (26,4)	0,68 [0,42;1,08], p=0,1032
S104	31	8 (25,8)	16	3 (18,8)	1,51 [0,34;6,69], p=0,5897
<b>≥40 Jahre</b>					
S102	148	20 (13,5)	155	46 (29,7)	0,37 [0,21;0,66], p=0,0009
S104	17	1 (5,9)	7	3 (42,9)	0,08 [0,01;1,03], p=0,0528
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,12
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					58,3%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test  KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

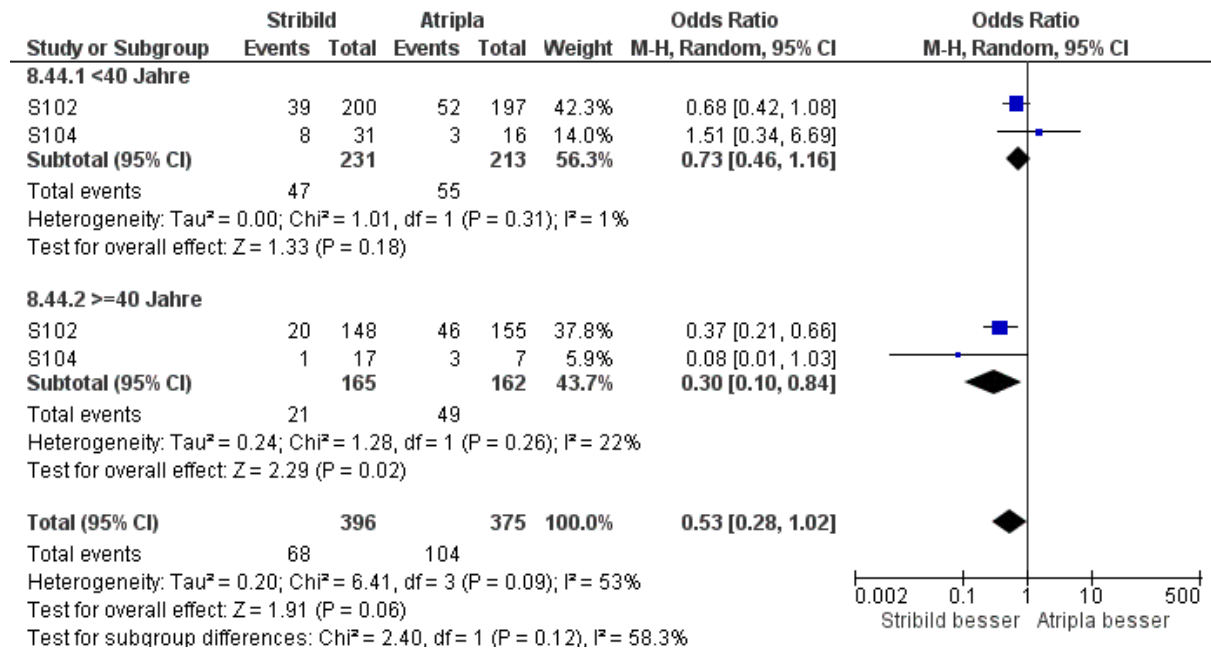


Abbildung 38: Meta-Analyse für Hautausschläge, stratifiziert nach Alter, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**4.3.1.3.2.9.1.2 Hautausschläge stratifiziert nach Rasse**

Tabelle 4-82: Subgruppenanalyse für Hautausschläge, stratifiziert nach Rasse (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>Kaukasisch</b>					
S102	214	37 (17,3)	227	77 (33,9)	0,41 [0,26;0,64], p=0,0001
S104	33	6 (18,2)	18	5 (27,8)	0,58 [0,15;2,25], p=0,4288
<b>Nicht-kaukasisch</b>					
S102	134	22 (16,4)	125	21 (16,8)	0,97 [0,51;1,87], p=0,9342
S104	15	3 (20,0)	5	1 (20,0)	1,00 [0,08;12,56], p=1,0000
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,03
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					78,4%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

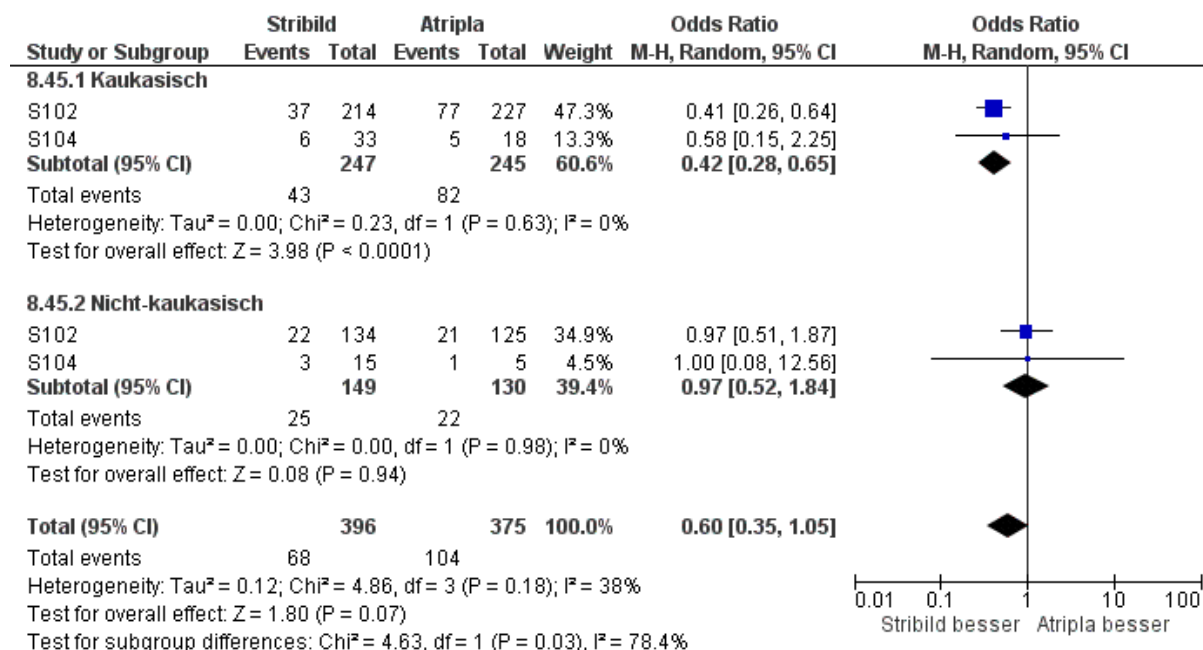


Abbildung 39: Meta-Analyse für Hautausschläge, stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

### 4.3.1.3.2.9.2 Hautausschläge zu Woche 60

#### 4.3.1.3.2.9.2.1 Hautausschläge stratifiziert nach Ausgangsviruslast

Tabelle 4-83: Subgruppenanalyse für Hautausschläge, stratifiziert nach Ausgangsviruslast (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	37	6 (16,2)	18	6 (33,3)	0,39 [0,10;1,44], p=0,1566
>100.000 Kopien/mL	11	3 (27,3)	5	0 (0,0)	4,53 [0,19;105,84], p=0,3475
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,16
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					49,8%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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					

**4.3.1.3.2.9.2.2 Hautausschläge stratifiziert nach Therapieadhärenz**

Tabelle 4-84: Subgruppenanalyse für Hautausschläge, stratifiziert nach Therapieadhärenz (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<95%	13	1 (7,7)	5	2 (40,0)	0,13 [0,01;1,89], p=0,1331
≥95%	33	8 (24,2)	17	4 (23,5)	1,04 [0,26;4,11], p=0,9554
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,17
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					46,3%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.9.3 Hautausschläge zu Woche 96****4.3.1.3.2.9.3.1 Hautausschläge stratifiziert nach Rasse**

Tabelle 4-85: Subgruppenanalyse für Hautausschläge, stratifiziert nach Rasse (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	214	47 (22,0)	227	81 (35,7)	0,51 [0,33;0,77], p=0,0016
Nicht-kaukasisch	134	27 (20,1)	125	27 (21,6)	0,92 [0,50;1,67], p=0,7740
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,11
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					59,9%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> 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 Renale Ereignisse

#### 4.3.1.3.2.10.1 Renale Ereignisse zu Woche 48

##### 4.3.1.3.2.10.1.1 Renale Ereignisse stratifiziert nach Ausgangsviruslast

Tabelle 4-86: Subgruppenanalyse für renale Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>≤100.000 Kopien/mL</b>					
S102	230	5 (2,2)	236	0 (0,0)	11,54 [0,63;209,84], p=0,0985
S104	37	0 (0,0)	18	0 (0,0)	n.b.
<b>&gt;100.000 Kopien/mL</b>					
S102	118	0 (0,0)	116	1 (0,9)	0,32 [0,01;8,06], p=0,4925
S104	15	0 (0,0)	5	0 (0,0)	n.b.
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,11
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					61,8%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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					

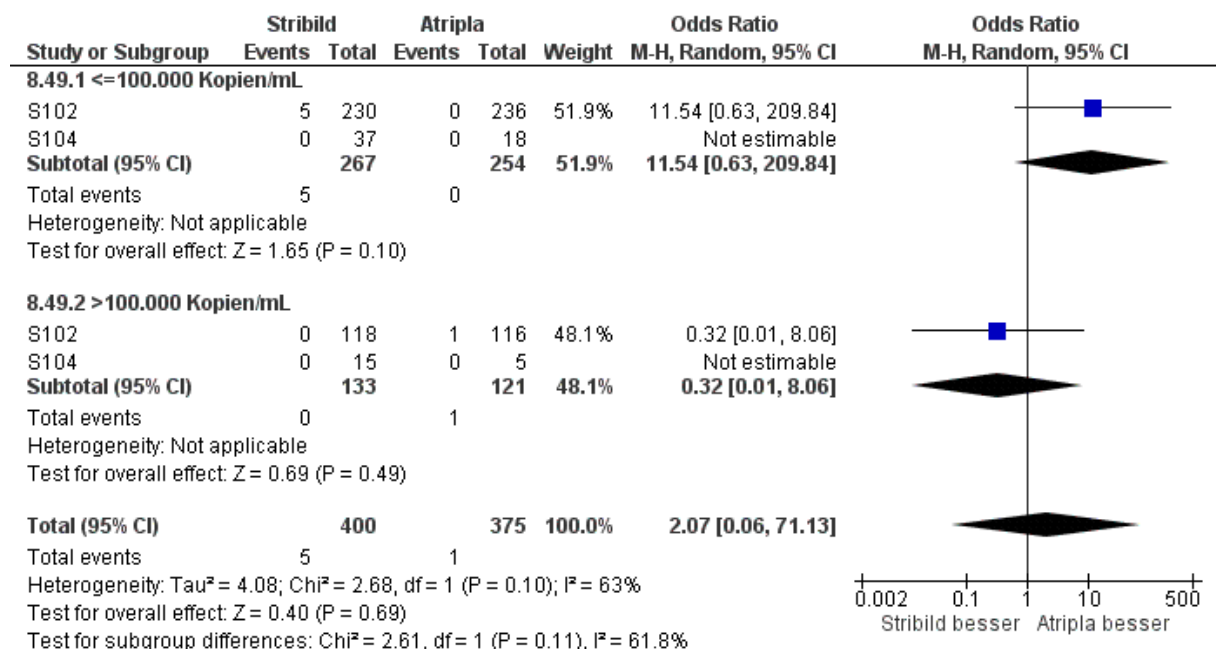


Abbildung 40: Meta-Analyse für renale Ereignisse, stratifiziert nach Ausgangsviruslast, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**4.3.1.3.2.10.2 Renale Ereignisse zu Woche 96****4.3.1.3.2.10.2.1 Renale Ereignisse stratifiziert nach Ausgangsviruslast**

Tabelle 4-87: Subgruppenanalyse für renale Ereignisse, stratifiziert nach Ausgangsviruslast (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	230	6 (2,6)	236	0 (0,0)	13,69 [0,77;244,51], p=0,0751
>100.000 Kopien/mL	118	1 (0,8)	116	1 (0,9)	0,98 [0,06;15,90], p=0,9903
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,1976
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					39,8%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Knochenfrakturen****4.3.1.3.2.11.1 Knochenfrakturen zu Woche 96****4.3.1.3.2.11.1.1 Knochenfrakturen stratifiziert nach CD4-Zellzahl zu Baseline**

Tabelle 4-88: Subgruppenanalyse für Knochenfrakturen, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤350 Zellen/μL	155	6 (3,9)	147	2 (1,4)	2,92 [0,58;14,70], p=0,1939
>350 Zellen/μL	193	4 (2,1)	205	6 (2,9)	0,70 [0,20;2,53], p=0,5881
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,18
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					45,5%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Erkrankungen des Gastrointestinaltrakts (SOC)

4.3.1.3.2.12.1 Erkrankungen des Gastrointestinaltrakts (SOC) zu Woche 48

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

Tabelle 4-89: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>≤350 Zellen/μL</b>					
S102	155	80 (51,6)	147	54 (36,7)	1,84 [1,16;2,91], p=0,0096
S104	24	11 (45,8)	8	2 (25,0)	2,54 [0,42;15,21], p=0,3078
<b>&gt;350 Zellen/μL</b>					
S102	193	108 (56,0)	205	107 (52,2)	1,16 [0,78;1,73], p=0,4516
S104	24	12 (50,0)	15	5 (33,3)	2,00 [0,52;7,63], p=0,3103
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,15
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					52,6%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

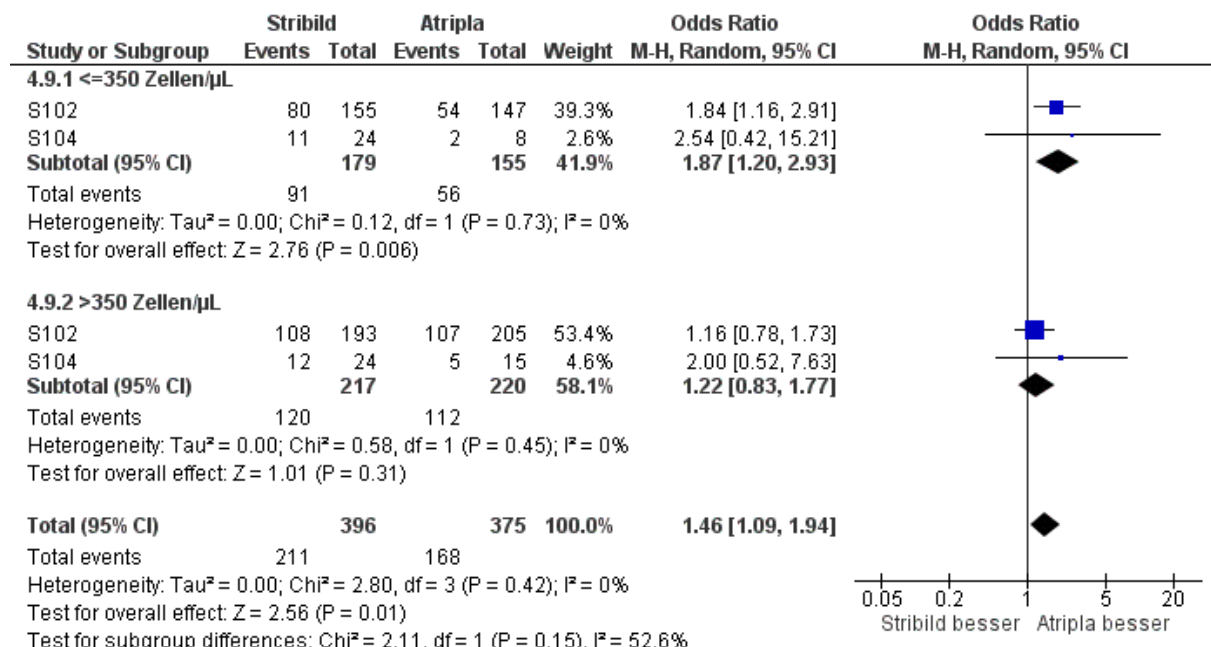


Abbildung 41: Meta-Analyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

#### 4.3.1.3.2.12.2 Erkrankungen des Gastrointestinaltrakts (SOC) zu Woche 60

##### 4.3.1.3.2.12.2.1 Erkrankungen des Gastrointestinaltrakts (SOC) stratifiziert nach Geschlecht

Tabelle 4-90: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Geschlecht (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Männlich	44	22 (50,0)	21	7 (33,3)	2,00 [0,68;5,91], p=0,2096
Weiblich	4	4 (100,0)	2	0 (0,0)	45,00 [0,67;3042,57], p=0,0766
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,16
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					49,2%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

#### 4.3.1.3.2.12.3 Erkrankungen des Gastrointestinaltrakts (SOC) zu Woche 96

##### 4.3.1.3.2.12.3.1 Erkrankungen des Gastrointestinaltrakts (SOC) stratifiziert nach Ausgangsviruslast

Tabelle 4-91: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Ausgangsviruslast (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	230	134 (58,3)	236	134 (56,8)	1,06 [0,74;1,53], p=0,7464
>100.000 Kopien/mL	118	77 (65,3)	116	54 (46,6)	2,16 [1,27;3,65], p=0,0042
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,03
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					78,6%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					



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

Tabelle 4-92: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤350 Zellen/μL	155	93 (60,0)	147	68 (46,3)	1,74 [1,10;2,75], p=0,0171
>350 Zellen/μL	193	118 (61,1)	205	120 (58,5)	1,11 [0,75;1,66], p=0,5966
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,15
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					51,9%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

#### 4.3.1.3.2.12.3.3 Erkrankungen des Gastrointestinaltrakts (SOC) stratifiziert nach Therapieadhärenz

Tabelle 4-93: Subgruppenanalyse für Erkrankungen des Gastrointestinaltrakts (SOC), stratifiziert nach Therapieadhärenz (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<95%	90	60 (66,7)	86	42 (48,8)	2,10 [1,14;3,85], p=0,0173
≥95%	256	151 (59,0)	263	146 (55,5)	1,15 [0,81;1,63], p=0,4243
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,09
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					64,2%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> 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 Diarrhoe

## 4.3.1.3.2.13.1 Diarrhoe zu Woche 48

## 4.3.1.3.2.13.1.1 Diarrhoe stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-94: Subgruppenanalyse für Diarrhoe, stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>≤350 Zellen/μL</b>					
S102	155	38 (24,5)	147	23 (15,6)	1,75 [0,98;3,12], p=0,0567
S104	24	4 (16,7)	8	1 (12,5)	1,40 [0,13;14,74], p=0,7794
<b>&gt;350 Zellen/μL</b>					
S102	193	42 (21,8)	205	43 (21,0)	1,05 [0,65;1,69], p=0,8483
S104	24	6 (25,0)	15	3 (20,0)	1,33 [0,28;6,39], p=0,7189
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,19
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					40,9%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

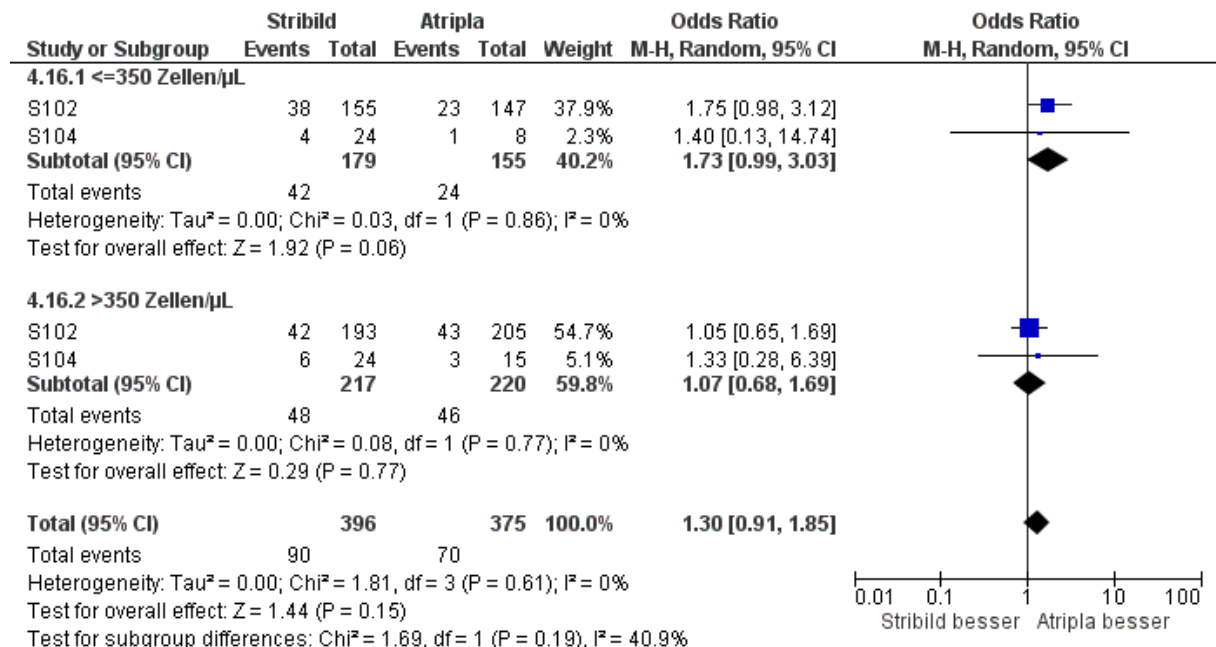


Abbildung 42: Meta-Analyse für Diarrhoe, stratifiziert nach CD4-Zellzahl zu Baseline, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**4.3.1.3.2.13.2 Diarrhoe zu Woche 60****4.3.1.3.2.13.2.1 Diarrhoe stratifiziert nach Ausgangsviruslast**

Tabelle 4-95: Subgruppenanalyse für Diarrhoe, stratifiziert nach Ausgangsviruslast (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	37	11 (29,7)	18	2 (11,1)	3,38 [0,66;17,28], p=0,1427
>100.000 Kopien/mL	11	2 (18,2)	5	2 (40,0)	0,33 [0,03;3,51], p=0,3607
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,11
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					60,2%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Diarrhoe stratifiziert nach Therapieadhärenz**

Tabelle 4-96: Subgruppenanalyse für Diarrhoe, stratifiziert nach Therapieadhärenz (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<95%	13	7 (53,8)	5	0 (0,0)	12,69 [0,58;276,08], p=0,1059
≥95%	33	6 (18,2)	17	4 (23,5)	0,72 [0,17;3,01], p=0,6551
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,10
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					63,5%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.13.3 Diarrhoe zu Woche 96****4.3.1.3.2.13.3.1 Diarrhoe stratifiziert nach Therapieadhärenz**

Tabelle 4-97: Subgruppenanalyse für Diarrhoe, stratifiziert nach Therapieadhärenz (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<95%	90	26 (28,9)	86	17 (19,8)	1,65 [0,82;3,32], p=0,1612
≥95%	256	61 (23,8)	263	66 (25,1)	0,93 [0,63;1,39], p=0,7372
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,17
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					47,7%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Übelkeit****4.3.1.3.2.14.1 Übelkeit zu Woche 48****4.3.1.3.2.14.1.1 Übelkeit stratifiziert nach Ausgangsviruslast**

Tabelle 4-98: Subgruppenanalyse für Übelkeit, stratifiziert nach Ausgangsviruslast (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>≤100.000 Kopien/mL</b>					
S102	230	44 (19,1)	236	35 (14,8)	1,36 [0,84;2,21], p=0,2172
S104	n.z.				
<b>&gt;100.000 Kopien/mL</b>					
S102	118	28 (23,7)	116	13 (11,2)	2,46 [1,20;5,04], p=0,0135
S104	n.z.				
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,18
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					45,0%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.z. = nicht zutreffend, OR = Odds Ratio					

**4.3.1.3.2.14.2 Übelkeit zu Woche 96****4.3.1.3.2.14.2.1 Übelkeit stratifiziert nach Ausgangsviruslast**

Tabelle 4-99: Subgruppenanalyse für Übelkeit, stratifiziert nach Ausgangsviruslast (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	230	46 (20,0)	236	39 (16,5)	1,26 [0,79;2,02], p=0,3321
>100.000 Kopien/mL	118	30 (25,4)	116	14 (12,1)	2,48 [1,24;4,98], p=0,0104
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,11
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					59,8%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)****4.3.1.3.2.15.1 Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) zu Woche 48****4.3.1.3.2.15.1.1 Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) stratifiziert nach CD4-Zellzahl zu Baseline**

Tabelle 4-100: Subgruppenanalyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>≤350 Zellen/μL</b>					
S102	155	29 (18,7)	147	50 (34,0)	0,45 [0,26;0,76], p=0,0028
S104	24	6 (25,0)	8	3 (37,5)	0,56 [0,10;3,05], p=0,4989
<b>&gt;350 Zellen/μL</b>					
S102	193	39 (20,2)	205	49 (23,9)	0,81 [0,50;1,30], p=0,3751
S104	24	7 (29,2)	15	4 (26,7)	1,13 [0,27;4,80], p=0,8660
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,08
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					67,4%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

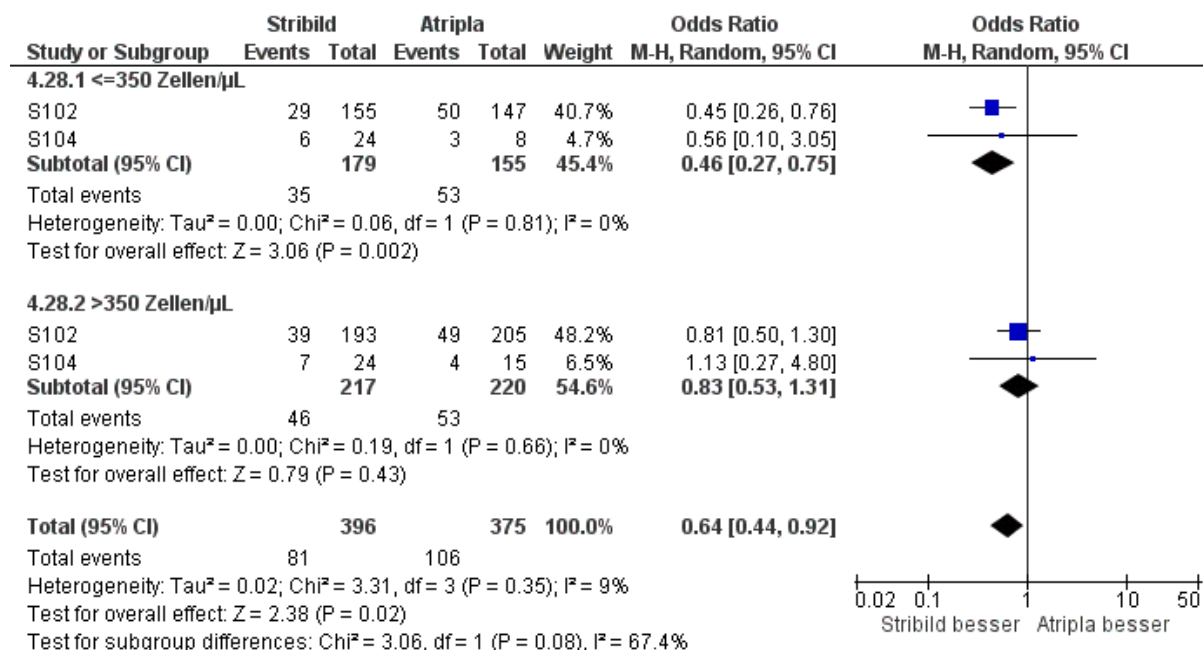


Abbildung 43: Meta-Analyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), stratifiziert nach CD4-Zellzahl zu Baseline, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

### 4.3.1.3.2.15.2 Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) zu Woche 96

#### 4.3.1.3.2.15.2.1 Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) stratifiziert nach Geschlecht

Tabelle 4-101: Subgruppenanalyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), stratifiziert nach Geschlecht (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Männlich	307	71 (23,1)	316	102 (32,3)	0,63 [0,44;0,90], p=0,0111
Weiblich	41	14 (34,1)	36	7 (19,4)	2,15 [0,75;6,13], p=0,1527
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,03
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					78,8%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) stratifiziert nach CD4-Zellzahl zu Baseline

Tabelle 4-102: Subgruppenanalyse für allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC), stratifiziert nach CD4-Zellzahl zu Baseline (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤350 Zellen/μL	155	35 (22,6)	147	55 (37,4)	0,49 [0,29;0,81], p=0,0052
>350 Zellen/μL	193	50 (25,9)	205	54 (26,3)	0,98 [0,62;1,53], p=0,9214
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,04
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					75,6%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 96

##### 4.3.1.3.2.16.1.1 Fatigue stratifiziert nach Geschlecht

Tabelle 4-103: Subgruppenanalyse für Fatigue, stratifiziert nach Geschlecht (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Männlich	307	36 (11,7)	316	50 (15,8)	0,71 [0,45;1,12], p=0,1396
Weiblich	41	9 (22,0)	36	2 (5,6)	4,78 [0,96;23,83], p=0,0563
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,02
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					80,1%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Therapieadhärenz**

Tabelle 4-104: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Therapieadhärenz (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>&lt;95%</b>					
S102	86	28 (32,6)	89	31 (34,8)	0,90 [0,48;1,69], p=0,7505
S104	12	6 (50,0)	5	2 (40,0)	1,50 [0,18;12,46], p=0,7074
<b>≥95%</b>					
S102	260	67 (25,8)	260	112 (43,1)	0,46 [0,32;0,66], p<0,0001
S104	34	9 (26,5)	17	6 (35,3)	0,66 [0,19;2,31], p=0,5157
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,0532
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					73,2%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

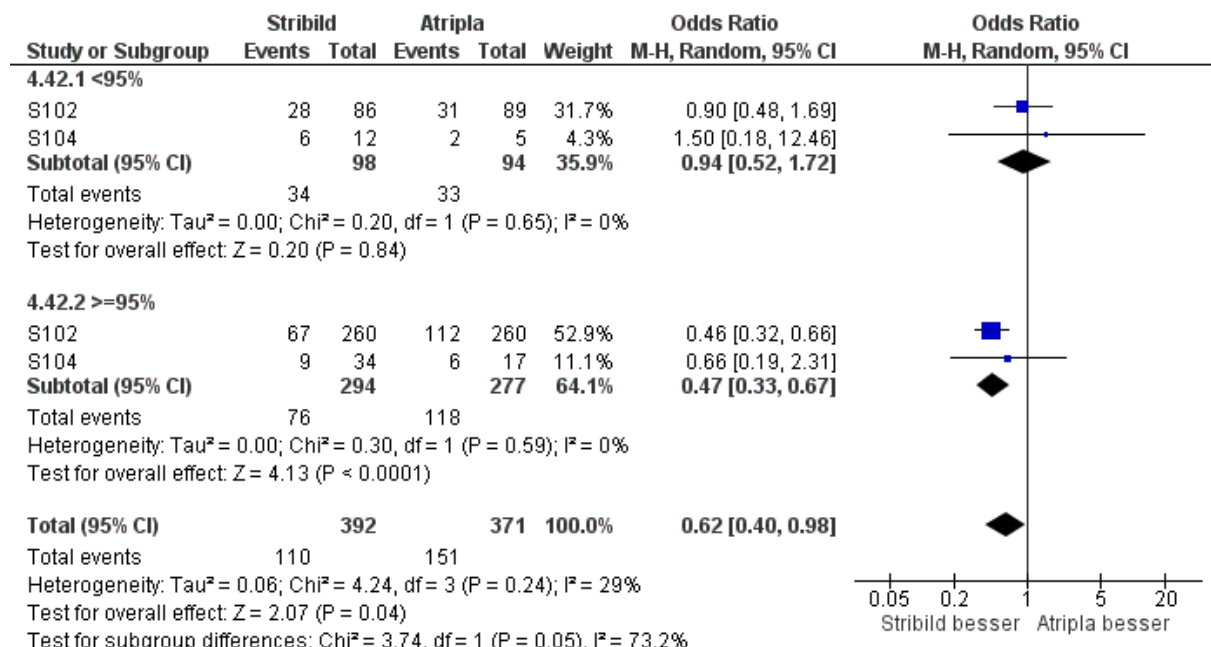


Abbildung 44: Meta-Analyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Therapieadhärenz, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)



**4.3.1.3.2.17.2 Erkrankungen des Nervensystems (SOC) zu Woche 60****4.3.1.3.2.17.2.1 Erkrankungen des Nervensystems (SOC) stratifiziert nach Rasse**

Tabelle 4-105: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Rasse (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	33	9 (27,3)	18	4 (22,2)	1,31 [0,34;5,06], p=0,6929
Nicht-kaukasisch	15	6 (40,0)	5	4 (80,0)	0,17 [0,01;1,88], p=0,1472
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,14
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					53,0%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.17.3 Erkrankungen des Nervensystems (SOC) zu Woche 96****4.3.1.3.2.17.3.1 Erkrankungen des Nervensystems (SOC) stratifiziert nach Therapieadhärenz**

Tabelle 4-106: Subgruppenanalyse für Erkrankungen des Nervensystems (SOC), stratifiziert nach Therapieadhärenz (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<95%	90	35 (38,9)	86	37 (43,0)	0,84 [0,46;1,54], p=0,5772
≥95%	256	77 (30,1)	263	119 (45,2)	0,52 [0,36;0,75], p=0,0004
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,18
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					44,8%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Geschlecht

Tabelle 4-107: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Geschlecht (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>Männlich</b>					
S102	307	40 (13,0)	316	25 (7,9)	1,74 [1,03;2,95], p=0,0385
S104	44	7 (15,9)	21	3 (14,3)	1,14 [0,26;4,91], p=0,8654
<b>Weiblich</b>					
S102	41	9 (22,0)	36	9 (25,0)	0,84 [0,29;2,43], p=0,7526
S104	4	0 (0,0)	2	1 (50,0)	0,11 [0,00;4,48], p=0,2439
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,18
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					43,3%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

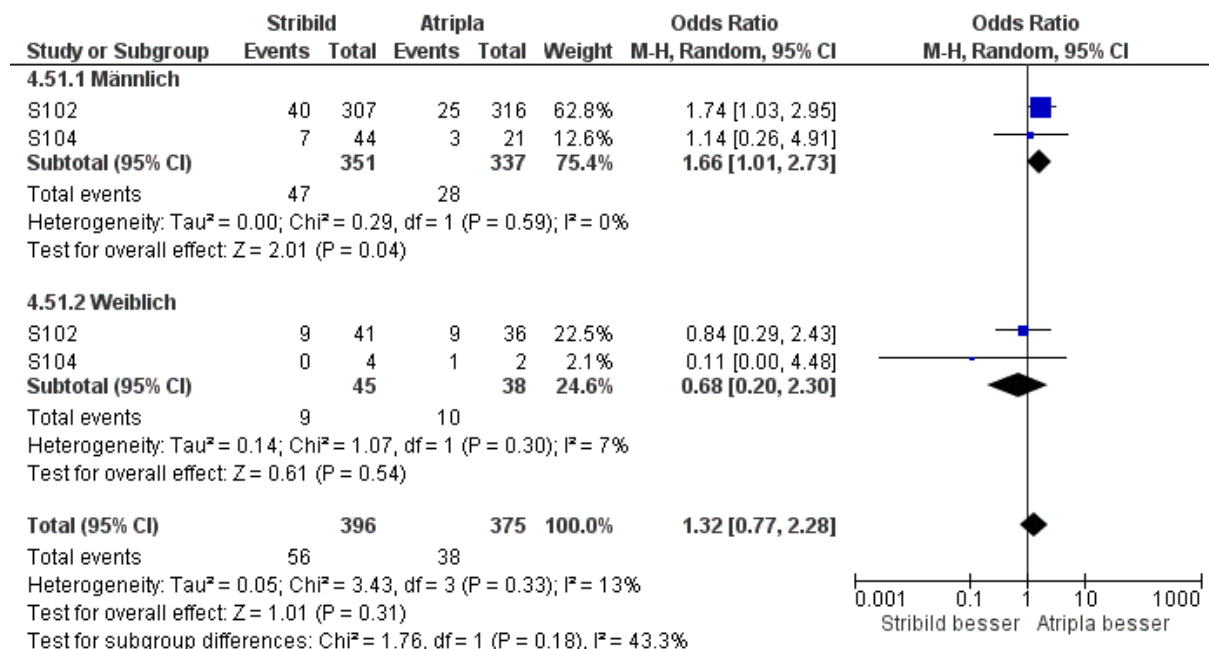


Abbildung 45: Meta-Analyse für Kopfschmerzen, stratifiziert nach Geschlecht, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**4.3.1.3.2.18.1.2 Kopfschmerzen stratifiziert nach Rasse**

Tabelle 4-108: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Rasse (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>Kaukasisch</b>					
S102	214	31 (14,5)	227	18 (7,9)	1,97 [1,06;3,63], p=0,0308
S104	33	5 (15,2)	18	2 (11,1)	1,43 [0,25;8,23], p=0,6897
<b>Nicht-kaukasisch</b>					
S102	134	18 (13,4)	125	16 (12,8)	1,06 [0,51;2,18], p=0,8802
S104	15	2 (13,3)	5	2 (40,0)	0,23 [0,02;2,37], p=0,2169
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,19
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					43,0%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

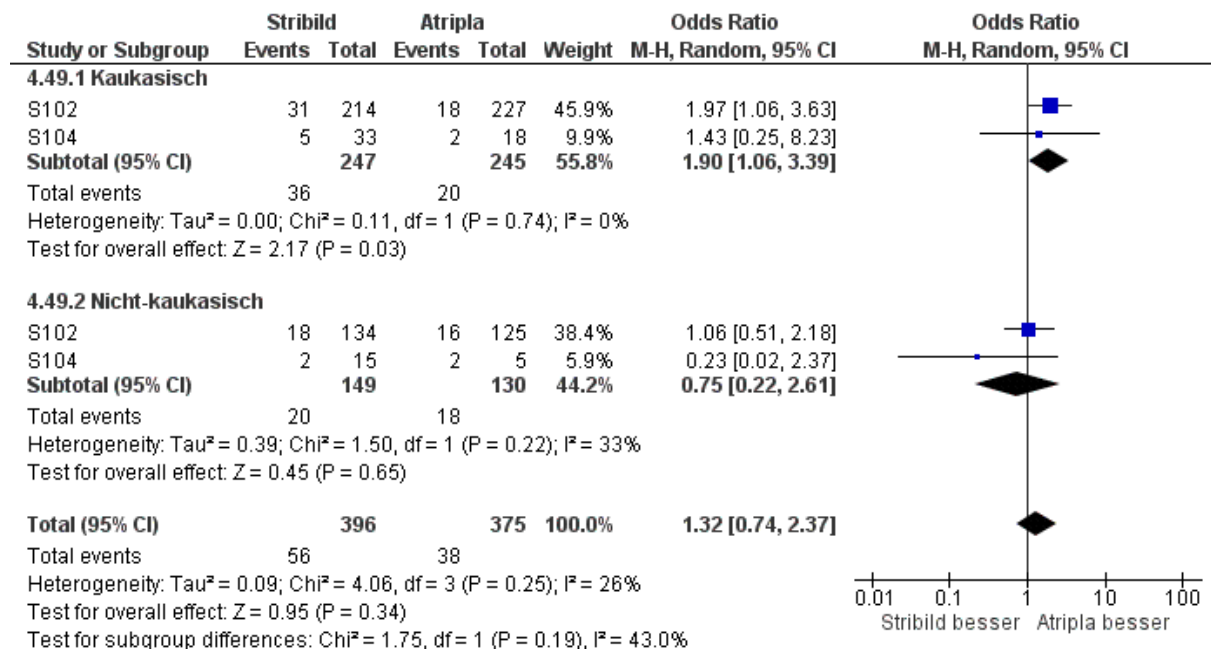


Abbildung 46: Meta-Analyse für Kopfschmerzen, stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**4.3.1.3.2.18.1.3 Kopfschmerzen stratifiziert nach Therapieadhärenz**

Tabelle 4-109: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Therapieadhärenz (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>&lt;95%</b>					
S102	86	12 (14,0)	89	4 (4,5)	3,45 [1,07;11,14], p=0,0388
S104	12	2 (16,7)	5	1 (20,0)	0,80 [0,06;11,50], p=0,8697
<b>≥95%</b>					
S102	260	37 (14,2)	260	30 (11,5)	1,27 [0,76;2,13], p=0,3603
S104	34	5 (14,7)	17	3 (17,6)	0,80 [0,17;3,86], p=0,7857
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,18
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					43,9%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

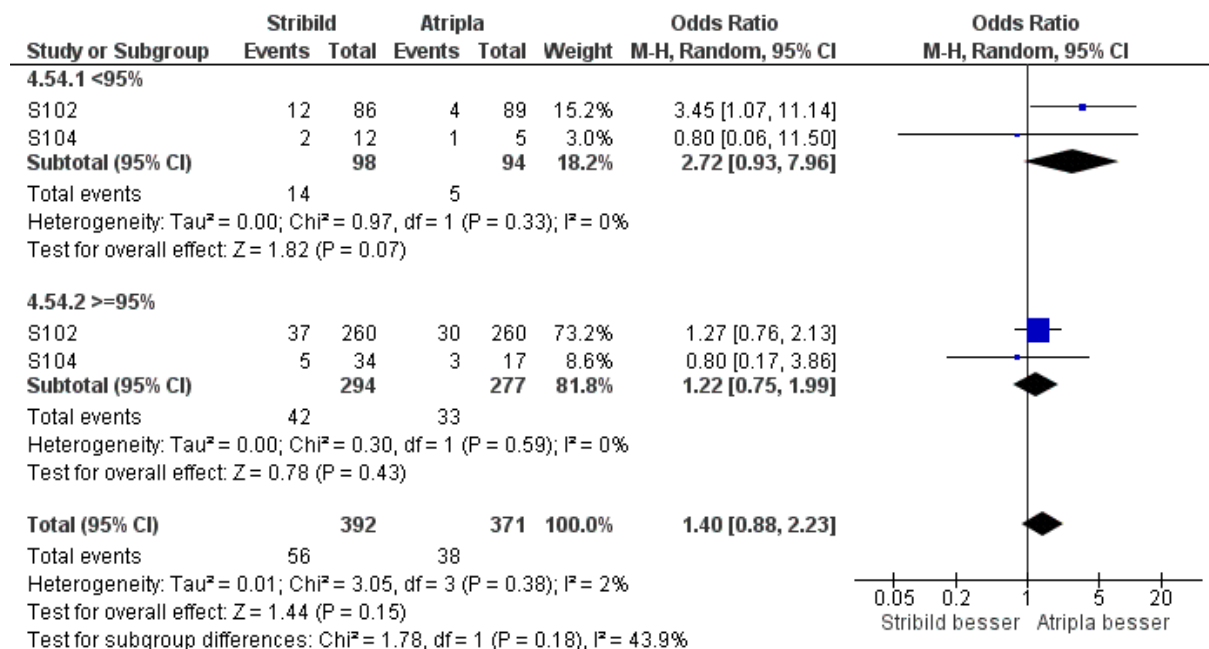


Abbildung 47: Meta-Analyse für Kopfschmerzen, stratifiziert nach Therapieadhärenz, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**4.3.1.3.2.18.2 Kopfschmerzen zu Woche 60****4.3.1.3.2.18.2.1 Kopfschmerzen stratifiziert nach Alter**

Tabelle 4-110: Subgruppenanalyse für Kopfschmerzen, stratifiziert nach Alter (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<40 Jahre	31	6 (19,4)	16	2 (12,5)	1,68 [0,30;9,47], p=0,5564
≥40 Jahre	17	1 (5,9)	7	2 (28,6)	0,16 [0,01;2,11], p=0,1620
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,14
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					55,0%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 60****4.3.1.3.2.19.1.1 Psychiatrische Erkrankungen (SOC) stratifiziert nach Rasse**

Tabelle 4-111: Subgruppenanalyse für psychiatrische Erkrankungen (SOC), stratifiziert nach Rasse (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	33	8 (24,2)	18	9 (50,0)	0,32 [0,09;1,08], p=0,0671
Nicht-kaucasisch	15	8 (53,3)	5	2 (40,0)	1,71 [0,22;13,41], p=0,6075
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,17
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					47,2%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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

#### 4.3.1.3.2.20.1 Abnorme Träume zu Woche 48

##### 4.3.1.3.2.20.1.1 Abnorme Träume stratifiziert nach Rasse

Tabelle 4-112: Subgruppenanalyse für abnorme Träume, stratifiziert nach Rasse (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>Kaukasisch</b>					
S102	214	30 (14,0)	227	67 (29,5)	0,39 [0,24;0,63], p=0,0001
S104	33	2 (6,1)	18	7 (38,9)	0,10 [0,02;0,56], p=0,0089
<b>Nicht-kaukasisch</b>					
S102	134	23 (17,2)	125	28 (22,4)	0,72 [0,39;1,33], p=0,2908
S104	15	3 (20,0)	5	1 (20,0)	1,00 [0,08;12,56], p=1,0000
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,13
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					56,0%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

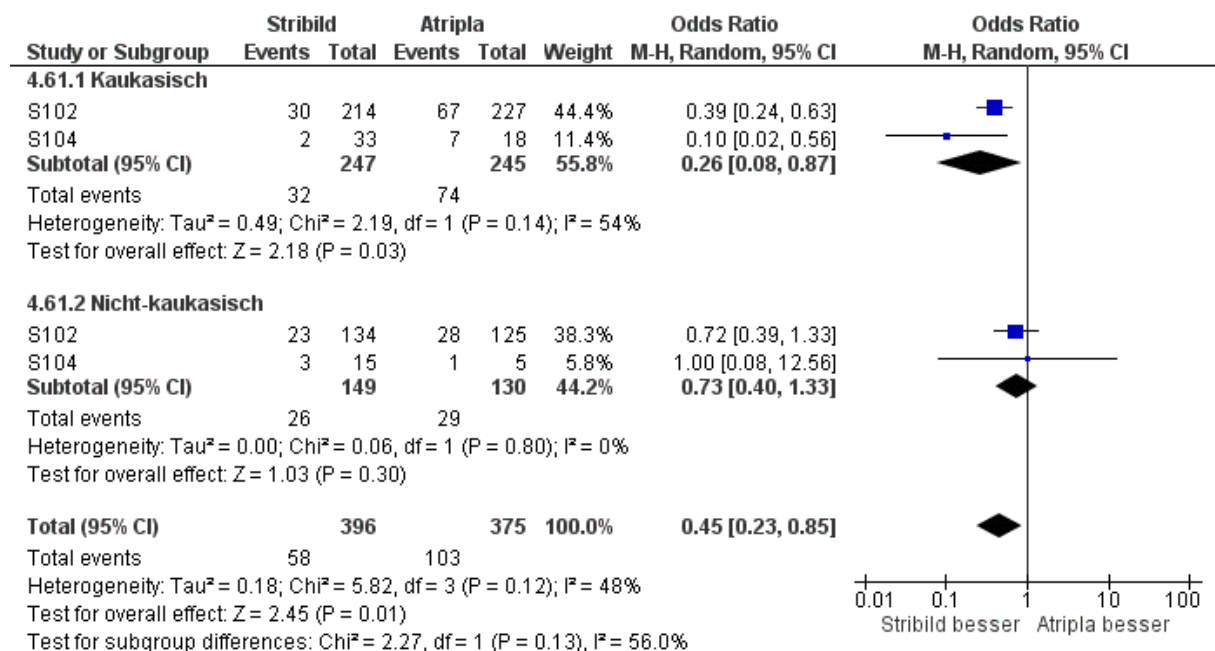


Abbildung 48: Meta-Analyse für abnorme Träume, stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

**4.3.1.3.2.20.2 Abnorme Träume zu Woche 60****4.3.1.3.2.20.2.1 Abnorme Träume stratifiziert nach Alter**

Tabelle 4-113: Subgruppenanalyse für abnorme Träume, stratifiziert nach Alter (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<40 Jahre	31	2 (6,5)	16	7 (43,8)	0,09 [0,02;0,51], p=0,0064
≥40 Jahre	17	3 (17,6)	7	1 (14,3)	1,29 [0,11;15,00], p=0,8411
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,08
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					67,0%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.20.2.2 Abnorme Träume stratifiziert nach Geschlecht**

Tabelle 4-114: Subgruppenanalyse für abnorme Träume, stratifiziert nach Geschlecht (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Männlich	44	4 (9,1)	21	8 (38,1)	0,16 [0,04;0,63], p=0,0085
Weiblich	4	1 (25,0)	2	0 (0,0)	2,14 [0,06;77,54], p=0,6772
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,19
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					42,4%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.20.2.3 Abnorme Träume stratifiziert nach Rasse**

Tabelle 4-115: Subgruppenanalyse für abnorme Träume, stratifiziert nach Rasse (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	33	2 (6,1)	18	7 (38,9)	0,10 [0,02;0,56], p=0,0089
Nicht-kaukasisch	15	3 (20,0)	5	1 (20,0)	1,00 [0,08;12,56], p=1,0000
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,14
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					53,6%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.20.3 Abnorme Träume zu Woche 96****4.3.1.3.2.20.3.1 Abnorme Träume stratifiziert nach Rasse**

Tabelle 4-116: Subgruppenanalyse für abnorme Träume, stratifiziert nach Rasse (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	214	30 (14,0)	227	68 (30,0)	0,38 [0,24;0,62], p=0,0001
Nicht-kaukasisch	134	23 (17,2)	125	29 (23,2)	0,69 [0,37;1,26], p=0,2270
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,14
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					54,4%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					



**4.3.1.3.2.21 Depression****4.3.1.3.2.21.1 Depression zu Woche 48****4.3.1.3.2.21.1.1 Depression stratifiziert nach Alter**

Tabelle 4-117: Subgruppenanalyse für Depression, stratifiziert nach Alter (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>&lt;40 Jahre</b>					
S102	200	21 (10,5)	197	18 (9,1)	1,17 [0,60;2,26], p=0,6485
S104				n.z.	
<b>≥40 Jahre</b>					
S102	148	12 (8,1)	155	21 (13,5)	0,56 [0,27;1,19], p=0,1324
S104				n.z.	
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,15
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					51,0%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, n.z. = nicht zutreffend, OR = Odds Ratio					

**4.3.1.3.2.21.2 Depression zu Woche 96****4.3.1.3.2.21.2.1 Depression stratifiziert nach Alter**

Tabelle 4-118: Subgruppenanalyse für Depression, stratifiziert nach Alter (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<40 Jahre	200	28 (14,0)	197	25 (12,7)	1,12 [0,63;2,00], p=0,7014
≥40 Jahre	148	14 (9,5)	155	24 (15,5)	0,57 [0,28;1,15], p=0,1167
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,15
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					52,7%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

**4.3.1.3.2.21.2.2 Depression stratifiziert nach Geschlecht**

Tabelle 4-119: Subgruppenanalyse für Depression, stratifiziert nach Geschlecht (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Männlich	307	34 (11,1)	316	46 (14,6)	0,73 [0,46;1,17], p=0,1952
Weiblich	41	8 (19,5)	36	3 (8,3)	2,67 [0,65;10,94], p=0,1733
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,09
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					65,5%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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 Erkrankungen der Haut und des Unterhautzellgewebes (SOC)****4.3.1.3.2.22.1 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) zu Woche 48****4.3.1.3.2.22.1.1 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) stratifiziert nach Rasse**

Tabelle 4-120: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Rasse (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>Kaukasisch</b>					
S102	214	51 (23,8)	227	93 (41,0)	0,45 [0,30;0,68], p=0,0001
S104	33	10 (30,3)	18	7 (38,9)	0,68 [0,21;2,28], p=0,5351
<b>Nicht-kaukasisch</b>					
S102	134	36 (26,9)	125	36 (28,8)	0,91 [0,53;1,56], p=0,7285
S104	15	5 (33,3)	5	1 (20,0)	2,00 [0,17;22,95], p=0,5777
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,04
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					76,6%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

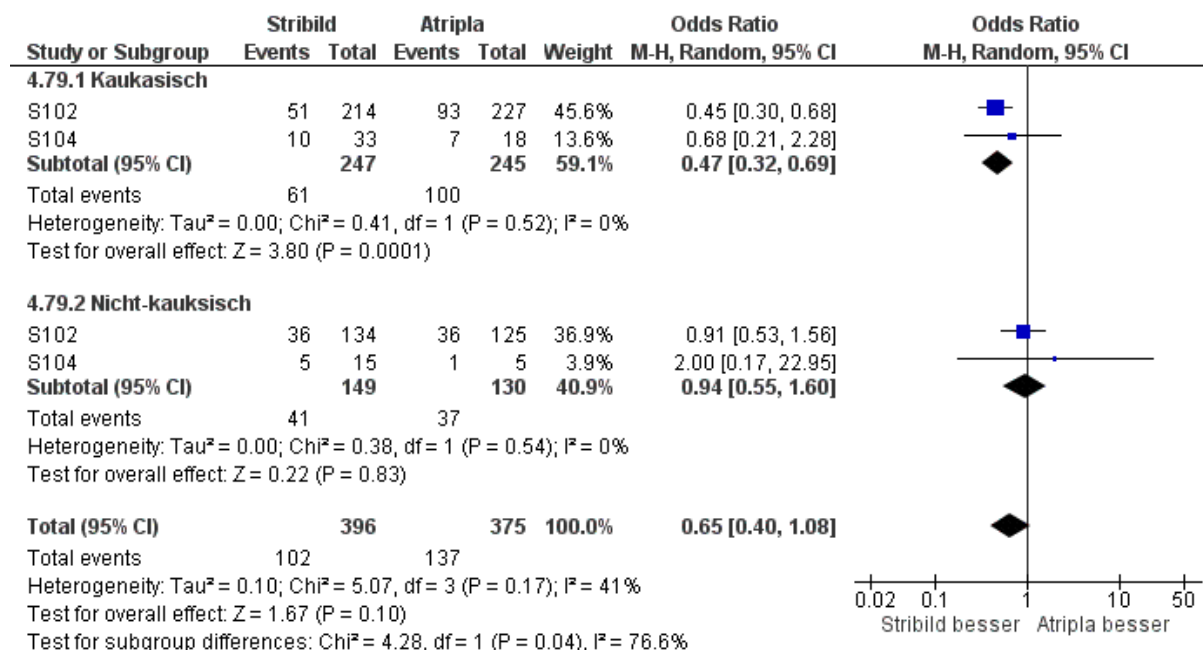


Abbildung 49: Meta-Analyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Rasse, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

#### 4.3.1.3.2.22.1.2 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) stratifiziert nach Therapieadhärenz

Tabelle 4-121: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Therapieadhärenz (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>&lt;95%</b>					
S102	86	25 (29,1)	89	25 (28,1)	1,05 [0,54;2,02], p=0,8859
S104	12	3 (25,0)	5	2 (40,0)	0,50 [0,05;4,58], p=0,5397
<b>≥95%</b>					
S102	260	62 (23,8)	260	104 (40,0)	0,47 [0,32;0,69], p=0,0001
S104	34	12 (35,3)	17	6 (35,3)	1,00 [0,30;3,38], p=1,0000
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,17
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					46,0%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

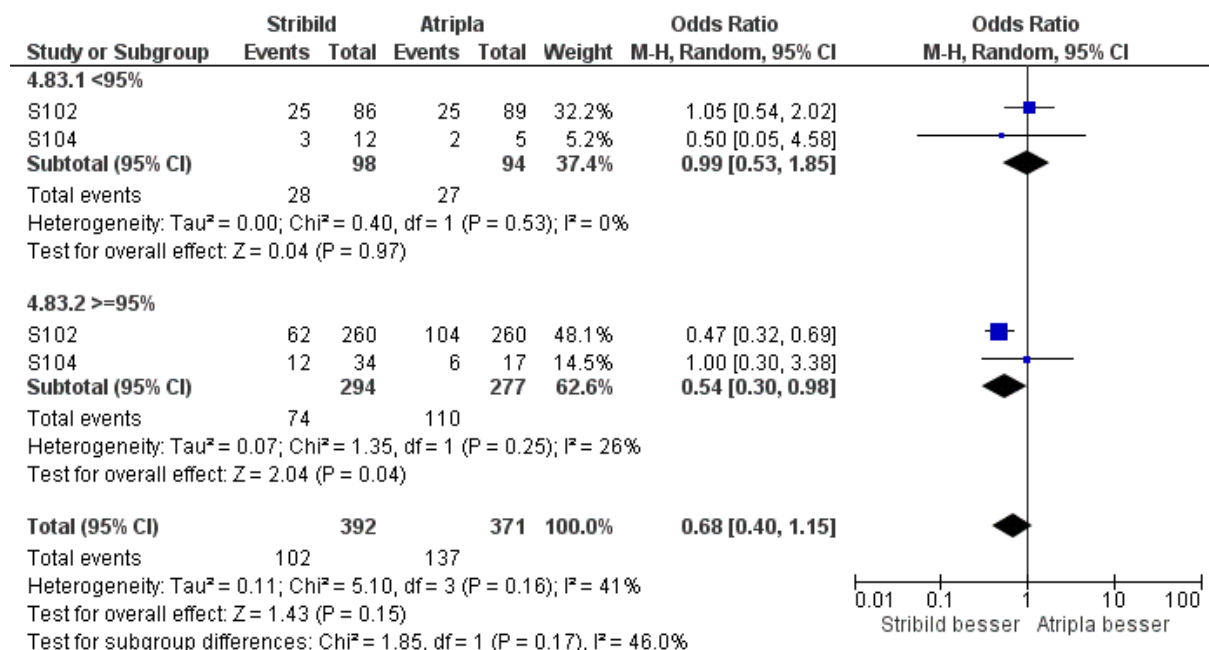


Abbildung 50: Meta-Analyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Therapieadhärenz, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

#### 4.3.1.3.2.22.2 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) zu Woche 60

##### 4.3.1.3.2.22.2.1 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) stratifiziert nach Geschlecht

Tabelle 4-122: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Geschlecht (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Männlich	44	12 (27,3)	21	8 (38,1)	0,61 [0,20;1,84], p=0,3786
Weiblich	4	3 (75,0)	2	0 (0,0)	11,67 [0,32;422,14], p=0,1797
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,12
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					57,9%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> 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 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) stratifiziert nach Ausgangsviruslast

Tabelle 4-123: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Ausgangsviruslast (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	37	11 (29,7)	18	8 (44,4)	0,53 [0,16;1,70], p=0,2845
>100.000 Kopien/mL	11	4 (36,4)	5	0 (0,0)	6,60 [0,29;149,77], p=0,2361
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,14
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					54,6%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

#### 4.3.1.3.2.22.3 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) zu Woche 96

##### 4.3.1.3.2.22.3.1 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) stratifiziert nach Rasse

Tabelle 4-124: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Rasse (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
Kaukasisch	214	66 (30,8)	227	104 (45,8)	0,53 [0,36;0,78], p=0,0013
Nicht-kaucasisch	134	45 (33,6)	125	43 (34,4)	0,96 [0,58;1,61], p=0,8896
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,07
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					70,2%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

#### 4.3.1.3.2.22.3.2 Erkrankungen der Haut und des Unterhautzellgewebes (SOC) stratifiziert nach Therapieadhärenz

Tabelle 4-125: Subgruppenanalyse für Erkrankungen der Haut und des Unterhautzellgewebes (SOC), stratifiziert nach Therapieadhärenz (Woche 96)

Studie S102 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<95%	90	33 (36,7)	86	30 (34,9)	1,08 [0,58;2,00], p=0,8052
≥95%	256	78 (30,5)	263	117 (44,5)	0,55 [0,38;0,78], p=0,0010
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,06
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					71,4%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> 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 Hautausschlag

##### 4.3.1.3.2.23.1 Hautausschlag zu Woche 48

##### 4.3.1.3.2.23.1.1 Hautausschlag stratifiziert nach Alter

Tabelle 4-126: Subgruppenanalyse für Hautausschlag, stratifiziert nach Alter (Woche 48)

Subgruppe Studie	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
<b>&lt;40 Jahre</b>					
S102	200	15 (7,5)	197	23 (11,7)	0,61 [0,31;1,21], p=0,1605
S104	31	4 (12,9)	16	2 (12,5)	1,04 [0,17;6,37], p=0,9687
<b>≥40 Jahre</b>					
S102	148	7 (4,7)	155	20 (12,9)	0,34 [0,14;0,82], p=0,0164
S104	17	1 (5,9)	7	2 (28,6)	0,16 [0,01;2,11], p=0,1620
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,17
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					48,1%
<sup>a</sup> zusätzlich berechnet					
<sup>b</sup> Chi-Quadrat-Test					
KI = Konfidenzintervall, N: Anzahl der Patienten in der Auswertung, n: Anzahl der Patienten mit Ereignis, OR = Odds Ratio					

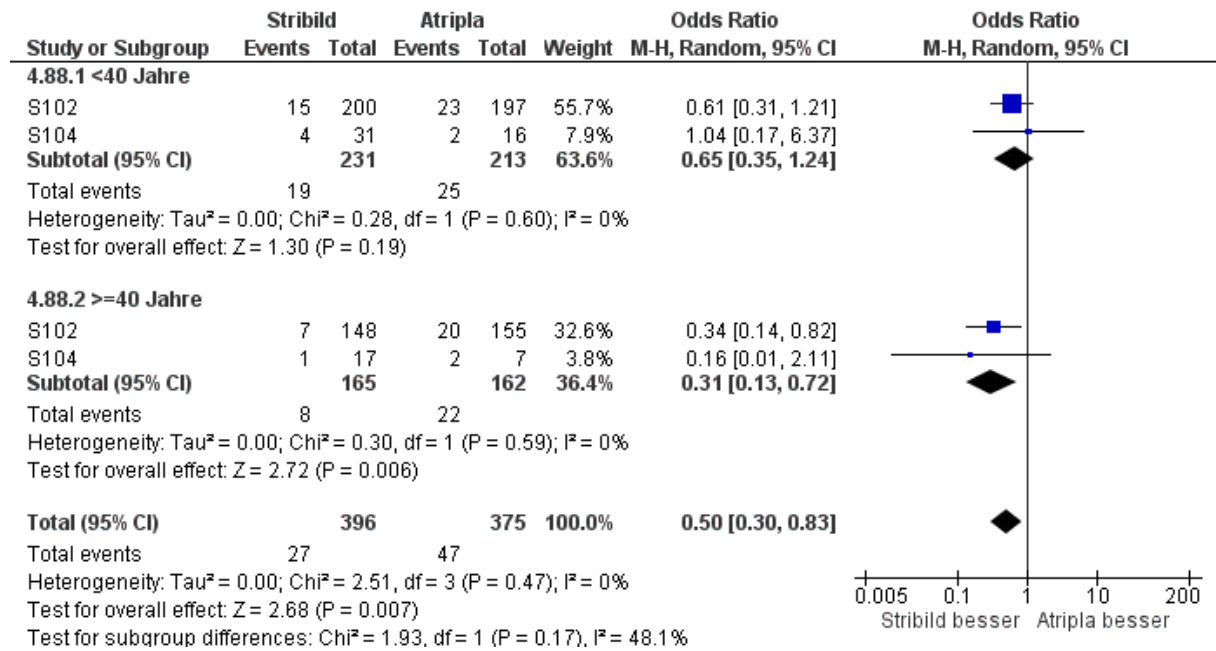


Abbildung 51: Meta-Analyse für Hautausschlag, stratifiziert nach Alter, aus RCT; EVG/COBI/FTC/TDF versus EFV/FTC/TDF (Woche 48, REM, OR)

### 4.3.1.3.2.23.2 Hautausschlag zu Woche 60

#### 4.3.1.3.2.23.2.1 Hautausschlag stratifiziert nach Ausgangsviruslast

Tabelle 4-127: Subgruppenanalyse für Hautausschlag, stratifiziert nach Ausgangsviruslast (Woche 60)

Studie S104 Subgruppe	EVG/COBI/FTC/TDF		EFV/FTC/TDF		OR <sup>a</sup> [95%-KI] <sup>a</sup> , p-Wert <sup>a,b</sup>
	N	n (%)	N	n (%)	
≤100.000 Kopien/mL	37	2 (5,4)	18	4 (22,2)	0,20 [0,03;1,22], p=0,0809
>100.000 Kopien/mL	11	3 (27,3)	5	0 (0,0)	4,53 [0,19;105,84], p=0,3475
<b>p-Wert Interaktionstest<sup>a</sup>:</b>					0,09
<b>I<sup>2</sup> (zusätzlich berechnet):</b>					64,7%
<sup>a</sup> zusätzlich berechnet <sup>b</sup> 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.*

Im vorliegenden Dossier zur Nutzenbewertung wird der Nutzen bzw. Zusatznutzen von Elvitegravir/Cobicistat/Emtricitabin/Tenofoviridisoproxil (EVG/COBI/FTC/TDF) anhand zweier randomisierter, aktivkontrollierter Studien gezeigt, welche EVG/COBI/FTC/TDF in der zugelassenen Dosierung mit der zweckmäßigen Vergleichstherapie, Efavirenz/Emtricitabin/Tenofoviridisoproxil (EFV/FTC/TDF), vergleichen. In beiden Studien bildet die Studienpopulation, die aus antiretroviral nicht-vorbehandelten erwachsenen Patienten mit einer HIV-1-Infektion besteht, einen Teil der Zulassungspopulation. Die Zulassungspopulation besteht aus nicht-vorbehandelten, HIV-1-infizierten Patienten ( $\geq 18$  Jahre) **und** 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. Da bei der letztgenannten Gruppe keine Angabe zum Behandlungsstatus gemacht wurde, umfasst diese Population auch Patienten unter Therapie (vorbehandelte Patienten) – ausgeschlossen werden lediglich 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. Da jedoch die Patientenpopulation der nicht-vorbehandelten Patienten ( $\geq 18$  Jahre) 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, umfasst, werden in Modul 4B nur **vorbehandelte** Patienten ( $\geq 18$  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. Dort finden sich Untersuchungen zu dieser Patientenpopulation mit der zweckmäßigen Vergleichstherapie Efavirenz/ Emtricitabin/ Tenofoviridisoproxil (EFV/FTC/TDF).

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 unerwünschte 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]).

Zur Berechnung der Meta-Analyse wurde standardmäßig das REM verwendet.



## Virologisches Ansprechen

### *Missing=Failure, Missing=Excluded und Missing/ART-Switch=Failure*

In Studie S104 war das virologische Ansprechen zu Woche 24, erfasst mittels dreier verschiedener Algorithmen und unter Verwendung des ITT-Sets, primärer Endpunkt. Mittels Missing=Failure-Imputation besteht kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,29 [0,28;5,94],  $p=0,72$ ). Die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze konnte nicht gezeigt werden (RD [95%-KI]: 2,8 [-14,5;20,1]). Gleiches gilt unter Verwendung des Missing=Excluded-Algorithmus, wo zu Woche 24 weder ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,07 [0,09;12,56],  $p=0,90$ ) noch die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze gezeigt werden kann (RD [95%-KI]: 0,7 [-13,4;14,8]). Mittels des Missing/ART-Switch=Failure-Algorithmus kann jedoch die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze gezeigt werden. Die RD [95%-KI] ist dabei 7,2 [-11,7;26,0]. Ein statistisch signifikanter Unterschied besteht zwischen den Behandlungen jedoch nicht (OR [95%-KI]: 1,81 [0,44;7,50],  $p=0,39$ ).

Unter Verwendung der PP-Population zeigen sich ähnliche Ergebnisse. Zu keinem der verwendeten Algorithmen kann ein statistisch signifikanter Unterschied zwischen den Behandlungen oder die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zu 12%-Grenze gezeigt werden. Das OR [95%-KI] ist jeweils 1,13 [0,10;13,25] mit  $p=0,88$  und die RD [95%-KI] ergibt sich jeweils zu 0,8 [-13,7;15,3]. Allerdings war die Studie mit einer geplanten Fallzahl von 75 Patienten, einem Zuteilungsverhältnis von 2:1 und einer daraus resultierenden Power von 26%, um Nichtunterlegenheit zur 12%-Grenze zu zeigen, nicht geeignet, um Unterschiede in der Wirksamkeit zu zeigen. Zu Woche 48 kann jedoch unter Verwendung des TLOVR- und Snapshot-Algorithmus die Nichtunterlegenheit zur 12%-Grenze gezeigt werden (siehe unten).

An dieser Stelle sei auf die Empfehlung der EMA verwiesen, die die Möglichkeit einer validen Erfassung des Endpunkts virologisches Ansprechen erst nach 48 Wochen Behandlung sieht [28].

### *Snapshot-Algorithmus*

#### *ITT-Population*

Der Snapshot-Algorithmus zu Woche 24 zeigt in Studie S104 weder einen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,29 [0,28;5,94],  $p=0,73$ ), noch eine signifikante Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF hinsichtlich des virologischen Ansprechens (Snapshot) (RD [95%-KI]: 2,7 [-15,4;20,9]). Zu Woche 48 zeigt sich hinsichtlich des **virologischen Ansprechens (Snapshot)** in der Meta-Analyse der beiden eingeschlossenen Studien kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI] aus Meta-Analyse: 1,40 [0,93;2,11],  $p=0,11$ ). Die Ergebnisse sind unter Verwendung eines FEM robust. In beiden

Studien kann die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze gezeigt werden (RD [95,2%-KI] Studie S102: 3,6 [-1,6;8,8]; RD [95%-KI] Studie S104: 9,2 [-9,9;28,3]). Zu Woche 96 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,21 [0,81;1,79],  $p=0,35$ ), die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze bleibt auch nach 96 Wochen Behandlung bestehen (RD [95%-KI]: 2,7 [-2,9;8,3]).

Es besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,18$ ;  $I^2=44,8\%$ ). 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.

### *PP-Population*

In der PP-Population besteht zu Woche 48 kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 0,83 [0,40;1,74],  $p=0,62$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF besteht jedoch auch hier in beiden Studien zur 12%-Grenze (RD [95,2%-KI] Studie S102: -1,0 [-4,4;2,4]; RD [95%-KI] Studie S104: 2,9 [-11,6;17,4]). Zu Woche 96 ist der Unterschied zwischen den Behandlungen statistisch nicht signifikant (OR [95%-KI]: 1,23 [0,50;3,01],  $p=0,65$ ). Die Nichtunterlegenheit von EVG/COBI/FTC/TDF zur 12%-Grenze gegenüber EFV/FTC/TDF kann mit einer RD [95%-KI] von 0,7 [-2,4;3,7] gezeigt werden.

### *TLOVR-Algorithmus*

Es besteht zu Woche 48 in der ITT-Population kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen hinsichtlich des **virologischen Ansprechens (TLOVR)**. Das OR [95%-KI] aus der Meta-Analyse beträgt 1,29 [0,86;1,91] mit  $p=0,21$ . Eine Sensitivitätsanalyse, unter Verwendung des FEM, zeigt robuste Ergebnisse. EVG/COBI/FTC/TDF ist gegenüber EFV/FTC/TDF in beiden Studien nichtunterlegen zur 12%-Grenze (RD [95,2%] Studie S102: 2,7 [-2,6;8,1]; RD [95%-KI]: 9,1 [-10,1;28,3]). Auch zu Woche 96 zeigt sich EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF nichtunterlegen zur 12%-Grenze (RD [95%-KI]: 2,1 [-4,0;8,2]); es gibt keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,13 [0,79;1,62],  $p=0,50$ ).

Zu Woche 96 bestehen Hinweise auf Effektmodifikation durch Rasse ( $p=0,19$ ;  $I^2=42,4\%$ ) und CD4-Zellzahl zu Baseline ( $p=0,07$ ;  $I^2=69,1\%$ ). 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.

### **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] aus der Meta-Analyse ergibt sich zu 0,88 [0,62;1,25] mit  $p=0,47$ . Auch an Woche 96 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (HR [95%-KI]: 0,87 [0,63;1,19],  $p=0,28$ ).

Subgruppenanalysen liefern keine Hinweise oder Belege auf Effektmodifikation.

### Veränderung der CD4-Zellzahl

Zu Woche 48 besteht hinsichtlich der **Veränderung der CD4-Zellzahl** in der ITT-Population 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],  $p=0,003$ ). Die CD4-Zellzahl stieg unter einer Behandlung mit EVG/COBI/FTC/TDF somit signifikant stärker an als unter EFV/FTC/TDF. Eine Sensitivitätsanalyse, unter Verwendung des FEM, zeigt robuste Ergebnisse. Die standardisierte Mittelwertdifferenz (Hedges'  $g$ ) [95%-KI] von 0,22 [0,07;0,37] mit  $p=0,003$  zeigt jedoch, dass der Unterschied zwar statistisch signifikant, allerdings nicht klinisch relevant (Untergrenze des Konfidenzintervalls  $>0,20$ ) ist. In der PP-Population ist der Unterschied zwischen den Behandlungen ebenfalls statistisch signifikant (MWD [95%-KI] aus Meta-Analyse: 39,45 [14,51;64,40],  $p=0,002$ ), allerdings auch nicht klinisch relevant (Hedges'  $g$  [95%-KI] aus Meta-Analyse: 0,24 [0,09;0,40],  $p=0,002$ ). Zu Woche 60 (nur Studie S104) ist der Unterschied zwischen den Behandlungen nicht statistisch signifikant. Die MWD [95%-KI] ergibt sich zu 23,83 [-0,76;48,43] mit  $p=0,06$ . Auch hier zeigt eine Sensitivitätsanalyse, unter Verwendung des FEM, robuste Ergebnisse. Zu Woche 96 besteht in der ITT-Population kein statistisch signifikanter Unterschied zwischen den Behandlungen (MWD [95%-KI]: 22 [-10;54],  $p=0,19$ ). Und auch in der PP-Population ist der Unterschied zwischen den Behandlungen statistisch nicht signifikant (MWD [95%-KI]: 23,00 [-9,96;55,96],  $p=0,1718$ ).

Es besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,10$ ;  $I^2=63,5\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhärenz waren, beträgt die MWD [95%-KI] 72,00 [5,52;138,48] mit  $p=0,0386$ , was einen statistisch signifikanten Unterschied zugunsten von EVG/COBI/FTC/TDF anzeigt, und in der Gruppe der zu  $\geq 95\%$  adhärenz Patienten liegt die MWD [95%-KI] bei 8,00 [-28,45;44,45] mit  $p=0,6671$ . Bei hoher Therapieadhärenz steigt die CD4-Zellzahl unter einer Behandlung mit EVG/COBI/FTC/TDF stärker an als unter einer Behandlung mit EFV/FTC/TDF.

### Mortalität

In Studie S102 verstarben bis Woche 48 ein Patient im EVG/COBI/FTC/TDF-Arm und zwei Patienten im EFV/FTC/TDF-Arm. In Studie S104 verstarb in den ersten 48 Wochen der Studie kein Patient. Es besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen hinsichtlich der **Mortalitätsrate** (OR [95%-KI]: 0,50 [0,05;5,59]  $p=0,5769$ ). Auch bis Woche 60 verstarb in Studie S104 kein Patient. Zu Woche 96 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen, da keine weiteren Todesfälle seit Woche 48 auftraten. Das OR [95%-KI] beträgt 0,50 [0,05;5,59],  $p=0,5769$ .

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 Behandlungen (OR [95%-KI] aus Meta-Analyse: 0,92 [0,51;1,68],  $p=0,80$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Auch zu Woche 60 kann kein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden (nur Studie S104; OR [95%-KI]: 1,05 [0,18;6,18],  $p=0,9590$ ). Zu Woche 96 sind die Behandlungen hinsichtlich der Gesamtrate unerwünschter Ereignisse ebenso vergleichbar (OR [95%-KI]: 0,90 [0,38;2,14],  $p=0,8041$ ).

Zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,1979$ ;  $I^2=39,7\%$ ). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

### Gesamtrate schwerwiegender unerwünschter Ereignisse

**Schwerwiegende unerwünschte Ereignisse** traten zu Woche 48 signifikant seltener unter EFV/FTC/TDF auf (OR [95%-KI] aus Meta-Analyse: 1,77 [1,06;2,97],  $p=0,03$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Dieser Unterschied zwischen den Behandlungen besteht zu Woche 60 nicht mehr (nur Studie S104; OR [95%-KI]: 0,96 [0,08;11,12],  $p=0,9717$ ). Zu Woche 96 ist der Unterschied zwischen den Behandlungen wieder statistisch signifikant zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,85 [1,17;2,93],  $p=0,0083$ ).

Es besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,14$ ;  $I^2=53,0\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 2,49 [1,33;4,66] mit  $p=0,0043$ , was einen statistisch signifikanten Unterschied zwischen den beiden Behandlungen zuungunsten von EVG/COBI/FTC/TDF anzeigt; in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 1,25 [0,63;2,48] mit  $p=0,5268$ .

### Therapieabbrüche aufgrund unerwünschter Ereignisse

Der Anteil an Patienten die die **Behandlung aufgrund unerwünschter Ereignisse abbrach**, unterschied sich zu Woche 48 zwischen den Behandlungen statistisch nicht signifikant (OR [95%-KI] aus Meta-Analyse: 0,67 [0,33;1,36],  $p=0,27$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Auch zu Woche 60 (nur Studie S104; OR [95%-KI]: 0,15 [0,01;3,95],  $p=0,2587$ ) und Woche 96 (OR [95%-KI]: 0,70 [0,37;1,33],  $p=0,2782$ ) besteht hinsichtlich Therapieabbrüchen aufgrund unerwünschter Ereignisse kein statistisch signifikanter Unterschied zwischen den Behandlungen.

Es besteht zu Woche 48 ein Beleg für Effektmodifikation durch Rasse ( $p=0,009$ ;  $I^2=85,4\%$ ). In der Gruppe der kaukasischen Patienten besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,48 [0,58;3,76],  $p=0,41$ ). In der Gruppe der nicht-kaukasischen Patienten besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,16 [0,04;0,64],

p=0,01). Nicht-kaukasische Patienten brechen somit unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener die Behandlung aufgrund unerwünschter Ereignisse ab als unter einer Behandlung mit EFV/FTC/TDF. Auch zu Woche 96 besteht ein Beleg für Effektmodifikation durch Rasse (p=0,04;  $I^2=76\%$ ). Hier ergibt sich in der Gruppe der Kaukasier ein OR [95%-KI] von 1,15 [0,53;2,51] mit p=0,7214 und in der Gruppe der nicht-kaukasischen Patienten beträgt das OR [95%-KI] 0,24 [0,06;0,87] mit p=0,0303. D.h. in der Gruppe der nicht-kaukasischen Patienten besteht ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen zugunsten von EVG/COBI/FTC/TDF.

### **Unerwünschte Ereignisse Grad 3-4**

Zu keinem der hier betrachteten Zeitpunkte besteht hinsichtlich **unerwünschter Ereignisse Grad 3-4** ein statistisch signifikanter Unterschied zwischen den Behandlungen. Zu Woche 48 beträgt das OR [95%-KI] aus der Meta-Analyse 1,14 [0,73;1,78] mit p=0,57 (eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse), zu Woche 60 (nur Studie S104) ist das OR [95%-KI] 0,46 [0,06;3,46] mit p=0,4483 und zu Woche 96 ergibt sich das OR [95%-KI] zu 1,25 [0,84;1,88], p=0,2733.

Es bestehen zu Woche 48 Hinweise auf Effektmodifikation durch Alter (p=0,06;  $I^2=71,5\%$ ) und Rasse (p=0,14;  $I^2=54,3\%$ ). Außerdem zeigt eine Subgruppenanalyse zu Woche 96 einen Hinweis auf Effektmodifikation durch Rasse (p=0,18;  $I^2=43,4\%$ ). 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**

Bis Woche 48 traten unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant weniger **neurologische & psychiatrische Ereignisse** auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,46 [0,34;0,61], p<0,00001). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Unter Ausschluss der Studie S104, dessen Verzerrungspotenzial für diesen Endpunkt nicht sicher ausgeschlossen werden kann, sind die Ergebnisse robust (OR [95%-KI]: 0,45 [0,33;0,61], p<0,001). Zu Woche 60 (nur S104) traten zwar auch numerisch deutlich weniger neurologische & psychiatrische Ereignisse unter EVG/COBI/FTC/TDF auf, allerdings erreicht dieser Unterschied, vermutlich aufgrund der geringen Patientenzahl und der daraus resultierenden niedrigen Power in Studie S104, keine statistische Signifikanz (OR [95%-KI]: 0,55 [0,20;1,50], p=0,31). Zu Woche 96 zeigt der Unterschied zwischen den Behandlungen hinsichtlich den neurologischen & psychiatrischen Ereignissen statistische Signifikanz zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,45 [0,33;0,61], p<0,001).

Subgruppenanalysen zu Woche 48 zeigen einen Hinweis auf Effektmodifikation durch Therapieadhärenz (p=0,16;  $I^2=48,3\%$ ). In der Gruppe der Patienten, die zu <95% adhären waren beträgt das OR [95%-KI] 0,65 [0,36;1,16] mit p=0,14, und für Patienten, die ≥95% adhären waren, ergibt sich das OR [95%-KI] zu 0,40 [0,29;0,57] mit p<0,00001. Wenn die

Medikation also zu mindestens 95% adhärent eingenommen wird, treten neurologische & psychiatrische Ereignisse unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF. Dieser Hinweis besteht auch unter Ausschluss der für diesen Endpunkt als nicht sicher niedrig verzerrt bewerteten Studie S104 ( $p=0,17$ ;  $I^2=47,4\%$ ). Zu Woche 60 (nur S104) besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,08$ ;  $I^2=67,0\%$ ). In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  aufwiesen, beträgt das OR [95%-KI] 1,67 [0,32;8,59] mit  $p=0,5415$ , und Patienten, die zu Baseline eine CD4-Zellzahl  $>350$  Zellen/ $\mu\text{L}$  aufwiesen, berechnet sich das OR [95%-KI] zu 0,25 [0,06;0,98] mit  $p=0,0471$ . Es besteht somit für Patienten, die zu Baseline eine höhere CD4-Zellzahl zu Baseline aufweisen, ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Außerdem besteht ein Hinweis auf Effektmodifikation zu Woche 96 durch Alter ( $p=0,18$ ;  $I^2=44,2\%$ ). Da dies jedoch nur eine quantitative Interaktion ist, wirkt sie sich nicht auf das Gesamtfazit aus (OR [95%-KI]  $<40$  Jahre: 0,54 [0,36;0,81],  $p=0,0027$ ; OR [95%-KI]  $\geq 40$  Jahre: 0,35 [0,22;0,56],  $p<0,0001$ ).

### Hautausschläge

Unter einer Behandlung mit EVG/COBI/FTC/TDF traten bis Woche 48 statistisch signifikant weniger **Hautausschläge** auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,54 [0,38;0,76],  $p=0,0005$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Unter Ausschluss der Studie S104, dessen Verzerrungspotenzial für diesen Endpunkt nicht sicher ausgeschlossen werden kann, sind die Ergebnisse robust (OR [95%-KI]: 0,53 [0,37;0,76],  $p<0,001$ ). Auch zu Woche 60 (nur S104) traten numerisch weniger Hautausschläge in der EVG/COBI/FTC/TDF-Gruppe auf, allerdings ist der Unterschied zwischen den Behandlungen, vermutlich aufgrund der geringen Patientenzahl und der daraus resultierenden niedrigen Power in Studie S104, statistisch nicht signifikant (OR [95%-KI]: 0,65 [0,20;2,13],  $p=0,54$ ). Zu Woche 96 besteht hinsichtlich Hautausschlägen wieder ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,61 [0,43;0,86],  $p=0,006$ ).

Zu Woche 48 besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,12$ ;  $I^2=58,3\%$ ). In der Gruppe der  $<40$ -jährigen Patienten beträgt das OR [95%-KI] 0,73 [0,46;1,16] mit  $p=0,18$ . Für  $\geq 40$ -jährige Patienten berechnet sich das OR [95%-KI] zu 0,30 [0,10;0,84] mit  $p=0,02$ . Unter den älteren Patienten besteht somit ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Dieser Hinweis besteht auch unter Ausschluss der für diesen Endpunkt als nicht sicher niedrig verzerrt bewerteten Studie S104 ( $p=0,12$ ;  $I^2=59,4\%$ ). Ein Beleg für Effektmodifikation besteht durch Rasse ( $p=0,03$ ;  $I^2=78,4\%$ ). In der kaukasischen Patientengruppe beträgt das OR [95%-KI] 0,42 [0,28;0,65] mit  $p<0,0001$  und für die nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,97 [0,52;1,84] mit  $p=0,94$ . Es besteht hier also in der kaukasischen Patientengruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Dieser Beleg besteht auch unter Ausschluss der für diesen Endpunkt als nicht sicher niedrig verzerrt bewerteten Studie S104 ( $p=0,03$ ;  $I^2=78,4\%$ ). Zu Woche 60

(nur S104) zeigen sich Hinweise auf Effektmodifikation durch Ausgangsviruslast ( $p=0,16$ ;  $I^2=49,8\%$ ) und Therapieadhärenz ( $p=0,17$ ;  $I^2=46,3\%$ ). 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. Es besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,11$ ;  $I^2=59,9\%$ ). In der Gruppe der Patienten kaukasischer Herkunft beträgt das OR [95%-KI] 0,51 [0,33;0,77] mit  $p=0,0016$ , was einen statistisch signifikanten Unterschied zwischen den beiden Behandlungen zugunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der nicht-kaukasischen Patienten besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,92 [0,50;1,67],  $p=0,7740$ ).

### Renale Ereignisse

Der Unterschied zwischen den Behandlungen hinsichtlich **renalere Ereignisse** ist zu Woche 48 statistisch nicht signifikant (OR [95%-KI]: 5,12 [0,59;44,02],  $p=0,12$ ). In Studie S104 traten bis Woche 60 keine renalen Ereignisse auf. Zu Woche 96 zeigt sich mit einem Peto Odds Ratio [95%-KI] von 4,60 [1,14;18,54] und  $p=0,037$  ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF.

Subgruppenanalysen zu Woche 48 zeigen einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,11$ ;  $I^2=61,8\%$ ). Außerdem zeigt eine Subgruppenanalyse zu Woche 96 einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,1976$ ;  $I^2=39,8\%$ ). 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.

### Knochenfrakturen

Es besteht hinsichtlich **Knochenfrakturen** zu keinem der hier betrachteten Zeitpunkte ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen. Zu Woche 48 beträgt das OR [95%-KI] aus der Meta-Analyse 1,06 [0,36;3,10] mit  $p=0,92$  (eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse), zu Woche 60 ist das OR [95%-KI] 1,48 [0,06;37,84] mit  $p=0,8111$  und zu Woche 96 beträgt das OR [95%-KI] 1,27 [0,50;3,26],  $p=0,64$ .

Eine Subgruppenanalyse zu Woche 48 ergibt einen Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,18$ ;  $I^2=45,5\%$ ). Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

### Erkrankungen des Gastrointestinaltrakts (SOC)

Es besteht zu Woche 48 hinsichtlich der **Erkrankungen des Gastrointestinaltrakts** ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,44 [1,08;1,91],  $p=0,01$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Zu Woche 60 (nur S104) ist der

Unterschied zwischen den Behandlungen allerdings nicht mehr statistisch signifikant (OR [95%-KI]: 2,70 [0,94;7,75],  $p=0,0647$ ). Und auch zu Woche 96 besteht hinsichtlich Erkrankungen des Gastrointestinaltrakts kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,34 [1,00;1,81],  $p=0,0538$ ).

Zu Woche 48 besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,15$ ;  $I^2=52,6\%$ ). In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  aufwiesen, beträgt das OR [95%-KI] 1,87 [1,20;2,93] mit  $p=0,006$ , und in der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl  $>350$  Zellen/ $\mu\text{L}$  aufwiesen, ergibt sich das OR [95%-KI] zu 1,22 [0,83;1,77] mit  $p=0,31$ . Das heißt, dass sich der statistisch signifikante Unterschied zuungunsten von EVG/COBI/FTC/TDF auf die Gruppe der Patienten beschränkt, die zu Baseline eine CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  aufwiesen. Es besteht zu Woche 60 (nur S104) ein Hinweis auf Effektmodifikation durch Geschlecht ( $p=0,16$ ;  $I^2=49,2\%$ ). 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 gibt es einen Beleg für Effektmodifikation durch Ausgangsviruslast ( $p=0,03$ ;  $I^2=78,6\%$ ). In der Subgruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 1,06 [0,74;1,53] mit  $p=0,7464$ . In der Subgruppe der Patienten, die zu Baseline eine Viruslast  $>100.000$  Kopien/mL aufwiesen, besteht ein statistisch signifikanter Unterschied, und das OR [95%-KI] beträgt 2,16 [1,27;3,65] mit  $p=0,0042$ . Des Weiteren besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,15$ ;  $I^2=51,9\%$ ). In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline  $\leq 350$  Zellen/ $\mu\text{L}$  beträgt das OR [95%-KI] 1,74 [1,10;2,75] mit  $p=0,0171$ , was einen statistisch signifikanten Unterschied zuungunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der Patienten mit CD4-Zellzahl zu Baseline  $>350$  Zellen/ $\mu\text{L}$  ergibt sich das OR [95%-KI] zu 1,11 [0,75;1,66] mit  $p=0,5966$ , was keinen statistisch signifikanten Unterschied anzeigt. Auch hinsichtlich der Therapieadhärenz kann ein Hinweis auf Effektmodifikation gefunden werden ( $p=0,09$ ;  $I^2=64,2\%$ ). In der Gruppe der Patienten mit einer Therapieadhärenz von  $<95\%$  beträgt das OR [95%-KI] 2,10 [1,14;3,85] mit  $p=0,0173$  und zeigt einen statistisch signifikanten Unterschied an. In der Gruppe der Patienten, die zu  $\geq 95\%$  adhärenz waren, liegt das OR [95%-KI] bei 1,15 [0,81;1,63] mit  $p=0,4243$ .

### **Diarrhoe**

**Diarrhoe** trat zu allen hier betrachteten Zeitpunkten vergleichbar oft in beiden Behandlungen auf. Zu Woche 48 beträgt das OR [95%-KI] aus der Meta-Analyse 1,29 [0,91;1,83] mit  $p=0,16$ . Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Zu Woche 60 berechnet sich das OR [95%-KI] zu 1,76 [0,50;6,17] mit  $p=0,3742$  und zu Woche 96 ergibt sich das OR [95%-KI] zu 1,08 [0,76;1,53] mit  $p=0,6613$ .

Es gibt zu Woche 48 einen Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,19$ ;  $I^2=40,9\%$ ). Außerdem besteht zu Woche 60 ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,11$ ;  $I^2=60,2\%$ ) und Therapieadhärenz ( $p=0,10$ ;  $I^2=63,5\%$ ). Des



Weiteren zeigt sich zu Woche 96 ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,17$ ;  $I^2=47,7\%$ ). 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.

### **Übelkeit**

Zu Woche 48 besteht hinsichtlich von **Übelkeit** ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 1,61 [1,09;2,38],  $p=0,02$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Dieser Unterschied besteht auch noch nach 96 Wochen Behandlung (OR [95%-KI]: 1,58 [1,07;2,32],  $p=0,0213$ ).

Es gibt zu Woche 48 einen Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,18$ ;  $I^2=45,0\%$ ). In der Gruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 1,36 [0,84;2,21] mit  $p=0,2172$ , und in der Gruppe der Patienten mit Ausgangsviruslast  $>100.000$  Kopien/mL ergibt sich das OR [95%-KI] zu 2,46 [1,20;5,04] mit  $p=0,0135$ . Zu Woche 96 besteht ebenso ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,11$ ;  $I^2=59,8\%$ ). In der Gruppe der Patienten, die zu Baseline eine Viruslast  $\leq 100.000$  Kopien/mL aufwiesen, beträgt das OR [95%-KI] 1,26 [0,79;2,02] mit  $p=0,3321$ , und in der Gruppe der Patienten mit Ausgangsviruslast  $>100.000$  Kopien/mL beträgt das OR [95%-KI] 2,48 [1,24;4,98] mit  $p=0,0104$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen anzeigt. Der statistisch signifikante Unterschied zuungunsten von EVG/COBI/FTC/TDF, der für die Gesamtpopulation gefunden wurde, beschränkt sich somit an beiden Zeitpunkten auf die Gruppe der Patienten, die zu Baseline eine Viruslast  $>100.000$  Kopien/mL aufwiesen.

### **Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)**

Zu Woche 48 traten **allgemeine Erkrankungen und Beschwerden am Verabreichungsort** statistisch signifikant seltener unter einer Behandlung mit EVG/COBI/FTC/TDF auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,64 [0,46;0,89],  $p=0,009$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Trotz numerisch geringerer Ereignisrate im EVG/COBI/FTC/TDF-Arm besteht zu Woche 60 (nur S104) kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,76 [0,25;2,29],  $p=0,6288$ ). Und auch zu Woche 96 besteht zwischen den Behandlungen kein statistisch signifikanter Unterschied (OR [95%-KI]: 0,72 [0,52;1,01],  $p=0,0537$ ), obwohl numerisch deutlich weniger allgemeine Erkrankungen und Beschwerden am Verabreichungsort unter einer Behandlung mit EVG/COBI/FTC/TDF auftraten als unter einer Behandlung mit EFV/FTC/TDF.

Subgruppenanalysen zu Woche 48 zeigen einen Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,08$ ;  $I^2=67,4\%$ ). In der Gruppe der Patienten, die zu Baseline eine CD4-Zellzahl  $\leq 350$  Zellen/ $\mu\text{L}$  aufwiesen, beträgt das OR [95%-KI] 0,46 [0,27;0,75] mit  $p=0,002$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen

zugunsten von EVG/COBI/FTC/TDF anzeigt, und in der Gruppe der Patienten mit höherer CD4-Zellzahl zu Baseline ergibt sich das OR [95%-KI] zu 0,83 [0,53;1,31]. Zu Woche 96 besteht ein Beleg für Effektmodifikation durch Geschlecht ( $p=0,03$ ;  $I^2=78,8\%$ ). Unter den Männern traten allgemeine Erkrankungen und Beschwerden am Verabreichungsort statistisch signifikant seltener im EVG/COBI/FTC/TDF-Arm auf als im EFV/FTC/TDF-Arm (OR [95%-KI]: 0,63 [0,44;0,90],  $p=0,0111$ ). In der Gruppe der Frauen kann kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen beobachtet werden (OR [95%-KI]: 2,15 [0,75;6,13],  $p=0,1527$ ). Auch für CD4-Zellzahl zu Baseline kann ein Beleg für Effektmodifikation gefunden werden ( $p=0,04$ ;  $I^2=75,6\%$ ). Patienten mit  $\leq 350$  Zellen/ $\mu\text{L}$  zu Baseline weisen unter eine Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant weniger allgemeine Erkrankungen und Beschwerden am Verabreichungsort auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,49 [0,29;0,81],  $p=0,0052$ ). In der Gruppe der Patienten, die zu Baseline  $>350$  Zellen/ $\mu\text{L}$  aufwiesen, besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,98 [0,62;1,53],  $p=0,9214$ ).

### **Fatigue**

Hinsichtlich des Auftretens von **Fatigue** besteht zu keinem der hier betrachteten Zeitpunkte ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen. Zu Woche 48 beträgt das OR [95%-KI] aus der Meta-Analyse 0,89 [0,58;1,37] mit  $p=0,60$  (eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse), zu Woche 60 (nur S104) ergibt sich das OR [95%-KI] zu 0,95 [0,25;3,55] mit  $p=0,9392$ , und zu Woche 96 beträgt das OR [95%-KI] 0,86 [0,56;1,32] mit  $p=0,4810$ .

Zu Woche 96 besteht ein Beleg für Effektmodifikation durch Geschlecht ( $p=0,02$ ;  $I^2=80,1\%$ ). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirken sich diese Hinweise auf Effektmodifikation nicht auf das Gesamtfazit aus.

### **Erkrankungen des Nervensystems (SOC)**

**Erkrankungen des Nervensystems** traten bis Woche 48 unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,55 [0,41;0,75],  $p=0,0001$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Zu Woche 60 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,85 [0,30;2,44],  $p=0,7661$ ). Zu Woche 96 ist der Unterschied zwischen den Behandlungen hinsichtlich von Erkrankungen des Nervensystems wieder statistisch signifikant (OR [95%-KI]: 0,58 [0,42;0,78],  $p=0,0004$ ).

Zu Woche 48 besteht ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,0532$ ;  $I^2=73,2\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhären waren, ergibt sich das OR [95%-KI] zu 0,94 [0,52;1,72] mit  $p=0,84$  und für Patienten, die eine Therapieadhärenz  $\geq 95\%$  aufwiesen, beträgt das OR [95%-KI] 0,47 [0,33;0,67] mit  $p<0,0001$ . Es besteht somit in der Gruppe der Patienten, die die Studienmedikation  $\geq 95\%$  adhären einnahmen, ein statistisch

signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Dort traten Erkrankungen des Nervensystems statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF. Zu Woche 60 besteht ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,14$ ;  $I^2=53,0\%$ ). Da aber in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich diese Interaktion nicht auf das Gesamtfazit aus. Zu Woche 96 ergibt sich ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,18$ ;  $I^2=44,8\%$ ). In der Gruppe der Patienten, die eine Therapieadhärenz  $<95\%$  aufwiesen, ergibt sich das OR [95%-KI] zu 0,84 [0,46;1,54] mit  $p=0,5772$ . Und in der Gruppe der Patienten, die zu  $\geq 95\%$  adhärenz waren, beträgt das OR [95%-KI] 0,52 [0,36;0,75] mit  $p=0,0004$ , was einen statistisch signifikanten Unterschied in der Vermeidung von Erkrankungen des Nervensystems zugunsten von EVG/COBI/FTC/TDF anzeigt.

### **Schwindel**

Zu Woche 48 besteht hinsichtlich des Auftretens von **Schwindel** ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,21 [0,13;0,34],  $p<0,00001$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Zu Woche 60 (nur S104) zeigt sich lediglich ein numerischer Vorteil von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF, der allerdings nicht statistisch signifikant ist (OR [95%-KI]: 0,06 [0,00;1,22],  $p=0,0674$ ). Zu Woche 96 besteht wieder ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,22 [0,14;0,36],  $p<0,00001$ ).

Subgruppenanalysen liefern keine Hinweise oder Belege auf Effektmodifikation.

### **Kopfschmerzen**

Zu keinem der hier betrachteten Zeitpunkte besteht hinsichtlich des Auftretens von **Kopfschmerzen** ein statistisch signifikanter Unterschied zwischen den Behandlungen. Zu Woche 48 beträgt das OR [95%-KI] aus der Meta-Analyse 1,43 [0,92;2,22] mit  $p=0,11$  (eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse), zu Woche 60 ergibt sich das OR [95%-KI] zu 0,81 [0,21;3,11] mit  $p=0,7599$  und zu Woche 96 ist das OR [95%-KI] 1,50 [0,97;2,31] mit  $p=0,0702$ .

Es besteht zu Woche 48 ein Hinweis auf Effektmodifikation durch Geschlecht ( $p=0,18$ ;  $I^2=43,3\%$ ). In der Gruppe der männlichen Patienten beträgt das OR [95%-KI] 1,66 [1,01;2,73] mit  $p=0,04$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der weiblichen Patienten ergibt sich das OR [95%-KI] zu 0,68 [0,20;2,30] mit  $p=0,54$ . Ein weiterer Hinweis auf Effektmodifikation besteht durch Rasse ( $p=0,19$ ;  $I^2=43,0\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 1,90 [1,06;3,39] mit  $p=0,03$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF anzeigt, und in der Gruppe der nicht-kaukasischen Patienten ergibt

sich das OR [95%-KI] zu 0,75 [0,22;2,61] mit  $p=0,65$ . Außerdem besteht ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,18$ ;  $I^2=43,9\%$ ). 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 60 (nur S104) besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,14$ ;  $I^2=55,0\%$ ). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

### **Psychiatrische Erkrankungen (SOC)**

Zu Woche 48 traten **psychiatrische Erkrankungen** statistisch signifikant seltener unter einer Behandlung mit EVG/COBI/FTC/TDF auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,59 [0,44;0,79],  $p=0,0003$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Trotz numerisch geringerer Ereignisrate im EVG/COBI/FTC/TDF-Arm besteht zu Woche 60 (nur S104) kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,55 [0,20;1,50],  $p=0,2417$ ). Zu Woche 96 besteht hinsichtlich dem Auftreten psychiatrischer Erkrankungen jedoch wieder ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,64 [0,47;0,86],  $p=0,0030$ ).

Subgruppenanalysen zu Woche 60 zeigen einen Hinweis auf Effektmodifikation durch Rasse ( $p=0,17$ ;  $I^2=47,2\%$ ). Da in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

### **Abnorme Träume**

Im EVG/COBI/FTC/TDF-Arm traten bis Woche 48 statistisch signifikant seltener **abnorme Träume** auf als im EFV/FTC/TDF-Arm (OR [95%-KI] aus Meta-Analyse: 0,41 [0,22;0,78],  $p=0,006$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt nur leichte Abweichungen (OR [95%-KI] FEM: 0,46 [0,32;0,65],  $p<0,0001$ ). Der statistisch signifikante Unterschied zugunsten von EVG/COBI/FTC/TDF besteht auch nach 60 Wochen Behandlung (OR [95%-KI]: 0,22 [0,06;0,77],  $p=0,0180$ , nur S104), und auch nach 96 Wochen Behandlung besteht hinsichtlich abnormer Träume ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,47 [0,32;0,69],  $p=0,0001$ ).

Subgruppenanalysen zu Woche 48 zeigen einen Hinweis auf Effektmodifikation durch Rasse ( $p=0,13$ ;  $I^2=56,0\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 0,26 [0,08;0,87] mit  $p=0,03$ , und in der Gruppe der nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,73 [0,40;1,33] mit  $p=0,30$ . Somit treten bei kaukasischen Patienten unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener abnorme Träume auf als unter einer Behandlung mit EFV/FTC/TDF. Außerdem besteht zu Woche 60 (nur S104) ein Hinweis auf Effektmodifikation durch Alter ( $p=0,08$ ;  $I^2=67,0\%$ ). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,09 [0,02;0,51] mit  $p=0,0064$ , was

einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF anzeigt. In der Gruppe der  $\geq 40$ -jährigen Patienten ergibt sich das OR [95%-KI] zu 1,29 [0,11;15,00] mit  $p=0,8411$ . Ein weiterer Hinweis auf Effektmodifikation besteht durch Geschlecht ( $p=0,19$ ;  $I^2=42,4\%$ ). In der Gruppe der männlichen Patienten beträgt das OR [95%-KI] 0,16 [0,04;0,63] mit  $p=0,0085$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF zeigt. Für die Frauen beträgt das OR [95%-KI] 2,14 [0,06;77,54] mit  $p=0,6772$ . Außerdem besteht ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,14$ ;  $I^2=53,6\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 0,10 [0,02;0,56] mit  $p=0,0089$ . In der Gruppe der nicht-kaukasischen Patienten ergibt sich das OR [95%-KI] zu 1,00 [0,08;12,56] mit  $p=1,0000$ . Somit treten abnorme Träume bei kaukasischen Patienten unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF. Es besteht zu Woche 96 ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,14$ ;  $I^2=54,4\%$ ). In der Gruppe der kaukasischen Patienten wurden im EVG/COBI/FTC/TDF-Arm statistisch signifikant weniger abnorme Träume beobachtet als im EFV/FTC/TDF-Arm (OR [95%-KI]: 0,38 [0,24;0,62],  $p=0,0001$ ), während unter den nicht-kaukasischen Patienten kein statistisch signifikanter Unterschied zwischen den beiden Behandlungen beobachtet werden kann (OR [95%-KI]: 0,69 [0,37;1,26],  $p=0,2270$ ).

### **Schlaflosigkeit**

Zu Woche 48 besteht hinsichtlich des Auftretens von **Schlaflosigkeit** ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,62 [0,39;0,98],  $p=0,04$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Zu Woche 60 (nur S104) besteht dieser Unterschied zwischen den Behandlungen nicht (OR [95%-KI]: 1,50 [0,28;8,08],  $p=0,6370$ ). Die Meta-Analyse zeigt zu Woche 96 ebenso wie in Woche 48 einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,63 [0,41;0,99],  $p=0,0428$ ).

Subgruppenanalysen liefern keine Hinweise oder Belege auf Effektmodifikation.

### **Depression**

Weder zu Woche 48 (OR [95%-KI] aus Meta-Analyse: 0,85 [0,53;1,36],  $p=0,49$ ; eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse) noch zu Woche 96 (OR [95%-KI]: 0,85 [0,55;1,32],  $p=0,4668$ ) besteht hinsichtlich des Auftretens von **Depressionen** ein statistisch signifikanter Unterschied zwischen den Behandlungen.

Zu Woche 48 besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,15$ ;  $I^2=51,0\%$ ). Außerdem liegen zu Woche 96 Hinweise auf Effektmodifikation durch Alter ( $p=0,15$ ;  $I^2=52,7\%$ ) und Geschlecht ( $p=0,09$ ;  $I^2=65,5\%$ ) vor. 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.

### **Angst**

Es besteht hinsichtlich des Auftretens von **Angst** weder zu Woche 48 (OR [95%-KI] aus Meta-Analyse: 0,76 [0,22;2,67],  $p=0,67$ ; OR [95%-KI] FEM: 0,97 [0,49;1,90],  $p=0,67$ ) noch zu Woche 60 (OR [95%-KI]: 0,29 [0,04;1,87],  $p=0,1930$ ) ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen.

Subgruppenanalysen liefern keine Hinweise oder Belege auf Effektmodifikation.

### **Erkrankungen der Haut und des Unterhautzellgewebes (SOC)**

**Erkrankungen der Haut und des Unterhautzellgewebes** traten zu Woche 48 unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auf als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,60 [0,44;0,81],  $p=0,001$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Zu Woche 60 (nur S104) besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,85 [0,30;2,44],  $p=0,7661$ ). Die Analyse zu Woche 96 (S102) zeigt ebenso wie in Woche 48 einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,65 [0,48;0,89],  $p=0,0070$ ).

Zu Woche 48 besteht ein Beleg für Effektmodifikation durch Rasse ( $p=0,04$ ;  $I^2=76,6\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] 0,47 [0,32;0,69] mit  $p=0,0001$ , was einen statistisch signifikanten Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF zeigt, und in der Gruppe der nicht kaukasischen Patienten ergibt sich das OR [95%-KI] zu 0,94 [0,55;1,60] mit  $p=0,83$ . Ein Hinweis auf Effektmodifikation besteht zu Woche 48 durch Therapieadhärenz ( $p=0,17$ ;  $I^2=46,0\%$ ). In der Gruppe der Patienten, die zu  $<95\%$  adhärenz waren, beträgt das OR [95%-KI] 0,99 [0,53;1,85] mit  $p=0,97$ , und in der Gruppe der zu  $\geq 95\%$  adhärenz Patienten ergibt sich das OR [95%-KI] zu 0,54 [0,30;0,98] mit  $p=0,04$ . Somit besteht in der Gruppe der hochadhärenz Patienten ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF. Zu Woche 60 (nur S104) bestehen Hinweise auf Effektmodifikation durch Geschlecht ( $p=0,12$ ;  $I^2=57,9\%$ ) und durch Ausgangsviruslast ( $p=0,14$ ;  $I^2=54,6\%$ ). 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 Rasse ( $p=0,07$ ;  $I^2=70,2\%$ ). In der Gruppe der kaukasischen Patienten besteht ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,53 [0,36;0,78],  $p=0,0013$ ), der in der Gruppe der nicht-kaukasischen Patienten nicht mehr beobachtet werden kann (OR [95%-KI]: 0,96 [0,58;1,61],  $p=0,8896$ ). Des Weiteren besteht ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,06$ ;  $I^2=71,4\%$ ). Bei Patienten, die zu Woche 96 eine Therapieadhärenz  $<95\%$  aufwiesen, besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,08 [0,58;2,00],  $p=0,8052$ ). In der Gruppe der Patienten, die zu  $\geq 95\%$  adhärenz waren, wurden statistisch signifikant weniger Erkrankungen der Haut und des Unterhautzellgewebes im

EVG/COBI/FTC/TDF-Arm beobachtet als im EFV/FTC/TDF-Arm (OR [95%-KI]: 0,55 [0,38;0,78],  $p=0,0010$ ).

### **Hautausschlag**

Zu Woche 48 besteht hinsichtlich des Auftretens von Hautausschlag ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI] aus Meta-Analyse: 0,49 [0,30;0,81],  $p=0,006$ ). Eine Sensitivitätsanalyse unter Verwendung eines FEM zeigt robuste Ergebnisse. Der Unterschied zu Woche 60 (nur S104) ist nicht statistisch signifikant (OR [95%-KI]: 0,55 [0,13;2,29],  $p=0,4130$ ). Die Analyse zu Woche 96 (S102) zeigt ebenso wie zu Woche 48, dass Hautausschläge unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant seltener auftreten als unter einer Behandlung mit EFV/FTC/TDF (OR [95%-KI]: 0,50 [0,30;0,82],  $p=0,0066$ ).

Zu Woche 48 gibt es einen Hinweis auf Effektmodifikation durch Alter ( $p=0,17$ ;  $I^2=48,1\%$ ). In der Gruppe der unter 40-jährigen Patienten beträgt das OR [95%-KI] 0,65 [0,35;1,24] mit  $p=0,19$  und in der Gruppe der mindestens 40-jährigen Patienten ergibt sich das OR [95%-KI] 0,31 [0,13;0,72] mit  $p=0,006$ . Somit besteht in der Gruppe der  $\geq 40$ -jährigen Patienten ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Außerdem zeigt sich zu Woche 60 (nur S104) ein Hinweis auf Effektmodifikation durch Ausgangsviruslast ( $p=0,09$ ;  $I^2=64,7\%$ ). Da jedoch in keiner Subgruppe ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht, wirkt sich dieser Hinweis auf Effektmodifikation nicht auf das Gesamtfazit aus.

Generell sind die Subgruppenanalysen, stratifiziert nach Therapieadhärenz, kritisch zu betrachten. 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 nicht praxisrelevant auf die klinische Praxis und den tatsächlichen Einsatz von EVG/COBI/FTC/TDF übertragbar, da die Patienten aufgrund des double-dummy Designs der hier eingeschlossenen Studien nicht nur eine Tablette täglich einnehmen mussten, wie für EVG/COBI/FTC/TDF und EFV/FTC/TDF in der Praxis vorgesehen, sondern zusätzlich eine Placebo-Tablette, verteilt über zwei Dosierungen am Tag. Verschiedene klinische Daten zeigen jedoch einen Unterschied zwischen einmal und zweimal täglicher Gabe [71, 72], was bedeutet, dass die Adhärenz-Daten in diesen Studien nicht auf die Praxis übertragbar sind. 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.

#### 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*

Nicht zutreffend.

#### 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.*

Nicht zutreffend.



#### 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 4-128: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

##### 4.3.2.1.3.1 <Endpunkt xxx> – 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 4-129: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Referenzen Studien	Intervention	<Vergleichs- therapie 1>	<Vergleichs- therapie 2>	<Vergleichs- therapie 3>
1	<Studie 1>	•		•	•
2	<Studie 2>	•		•	
	<Studie 3>	•		•	
1	<Studie 4>		•	•	•
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

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

Tabelle 4-130: Operationalisierung von &lt;Endpunkt xxx&gt;

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-131: Bewertung des Verzerrungspotenzials für &lt;Endpunkt xxx&gt; in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

*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 4-132: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)
<Studie 1>	

Nicht zutreffend.

*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.*

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

#### **4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT**

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.*

Nicht zutreffend.

#### 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 4-133: 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 4-134: Operationalisierung von &lt;Endpunkt xxx&gt;

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 4-135: Verzerrungsaspekte für &lt;Endpunkt xxx&gt; – 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

Nicht zutreffend.

**4.3.2.3.2 Charakteristika der weiteren Untersuchungen**

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungspotenzial.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht zutreffend.

**4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen****4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen**

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

Tabelle 4-136: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

*Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.*

Nicht zutreffend.

*Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung.*

Nicht zutreffend.

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

#### **4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen**

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.*

Nicht zutreffend.

#### **4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen**

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

*Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.*

Nicht zutreffend.



#### **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.*

Alle in die Nutzenbewertung eingeschlossenen Studien sind randomisierte, kontrollierte Studien und entsprechen demnach der Evidenzstufe Ib. Die Studien enthalten direkte Vergleiche des zu bewertenden Arzneimittels EVG/COBI/FTC/TDF mit der zweckmäßigen Vergleichstherapie EFV/FTC/TDF. Die Vollständigkeit des Studienpools wurde durch eine umfassende systematische Studienregistersuche bestätigt (siehe Abschnitt 4.2.3.3 und Anhang 4-B).

Die Studien wurden in einem double-dummy Parallelgruppendesign durchgeführt und wurden auf Studienebene jeweils mit einem niedrigen Verzerrungspotenzial bewertet. Auf Endpunktebene bestand in beiden Studien für alle Endpunkte ein niedriges Verzerrungspotenzial. In Studie S104 waren die Endpunkte virologisches Ansprechen (TLOVR) und Zeit bis Verlust des virologischen Ansprechens nicht präspezifiziert und wurden hier lediglich aus Vergleichbarkeitsgründen mit Studie S102 nachberechnet. Da dies unabhängig von den Daten geschah, wurde nicht von einem erhöhten Verzerrungspotenzial ausgegangen. Außerdem 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. Für wichtige unerwünschte Ereignisse (neurologische & psychiatrische Ereignisse und Hautausschläge) und für unerwünschte Ereignisse von Interesse (renale Ereignisse und Knochenfrakturen) konnte das Verzerrungspotenzial in Studie S104 aufgrund fehlender Präspezifikation nicht sicher ausgeschlossen werden. Dies wird in der Ableitung des Zusatznutzens berücksichtigt.

Da zwei randomisierte und kontrollierte Studien zur Bewertung des Zusatznutzens herangezogen wurden, sind grundsätzlich Belege für einen Zusatznutzen ableitbar.

Alle in die Nutzenbewertung eingeschlossenen Endpunkte sind valide und patientenrelevante Endpunkte, die auch u.a. vom IQWiG und der EMA als solche anerkannt wurden. Die Verwendung von Surrogatendpunkten wurde in Abschnitt 4.5.4 begründet.

##### **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.*

Hier wird der Nutzen bzw. Zusatznutzen des zu bewertenden Arzneimittels EVG/COBI/FTC/TDF im Vergleich zu der vom G-BA festgelegten zweckmäßigen Vergleichstherapie EFV/FTC/TDF bewertet. Dazu wurden zwei randomisierte aktiv-kontrollierte Studien, S102 und S104, in die Nutzenbewertung eingeschlossen, welche jeweils EVG/COBI/FTC/TDF mit der zweckmäßigen Vergleichstherapie EFV/FTC/TDF vergleichen. 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, Woche 60 und Woche 96 herangezogen. Dabei gilt zu beachten, dass in den beiden Studien teils unterschiedliche Zeitpunkte zur Analyse herangezogen wurden:

- Zu Woche 48 waren Daten aus beiden Studien S102 und S104 verfügbar.
- Zu Woche 60 liegen ausschließlich Daten aus der Studie S104 vor, da die kontrollierte Phase der Studie S104 nur bis Woche 60 andauerte. Für Studie S104 werden, neben der Veränderung der CD4-Zellzahl, die Ergebnisse zu Woche 60 nur für Sicherheitsendpunkte dargestellt.
- Zu Woche 96 liegen ausschließlich Daten aus der Studie S102 vor.

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. Die Ableitung des Zusatznutzens stützt sich hauptsächlich auf die Daten zu Woche 48 und Woche 96. 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. Dies trifft auch auf Nutzenaussagen in Subgruppen zu. Fett gedruckte Effektschätzer resultieren aus einer Meta-Analyse der beiden Studien S102 und S104.

Tabelle 4-137: Ausmaß des Zusatznutzen von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF auf Endpunktebene

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
<p>Virologisches Ansprechen Snapshot</p> <p>TLOVR</p>	<p><u>Woche 48:</u> <b>OR: 1,40 [0,93;2,11], p=0,12</b> <b>88,1% vs. 84,0%</b></p> <p><u>Woche 96:</u> OR: 1,21 [0,81;1,79], p=0,35 84,2% vs. 81,5%</p> <p><u>Woche 48:</u> <b>OR: 1,29 [0,86;1,91], p=0,21</b> <b>86,6% vs. 83,2%</b></p> <p><u>Woche 96:</u> OR: 1,13 [0,79;1,62], p=0,50 79,3% vs. 77,3%</p>	<p>Kein Zusatznutzen oder Schaden ableitbar (Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF wurde gezeigt)</p> <p>Kein Zusatznutzen oder Schaden ableitbar</p>
Zeit bis zum Verlust des virologischen Ansprechens	<p><u>Woche 48:</u> <b>HR: 0,88 [0,62;1,25], p=0,47</b></p> <p><u>Woche 96:</u> HR: 0,87 [0,63;1,19], p=0,28</p>	Kein Zusatznutzen oder Schaden ableitbar
Veränderung der CD4-Zellzahl	<p><u>Woche 48:</u> <b>MWD: 36,26 [12,42;60,10], p=0,003</b> <b>Hedges' g: 0,22 [0,07;0,37], p=0,003</b></p> <p><u>Woche 60:</u> MWD: 23,83 [-0,76;48,43], p=0,06</p> <p><u>Woche 96:</u> Adj. MWD: 22 [-10;54], p=0,18</p> <p><i>Interaktionstest: p=0,10</i> &lt;95% Therapieadhärenz: MWD: 72,00 [5,52;138,48],</p>	<p>Endpunktkategorie: AIDS-definierende Erkrankungen/Tod</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Beleg</p>

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	p=0,0386 ≥95% Therapieadhärenz: MWD: 8,00 [-28,45;44,45], p=0,6671	
Unerwünschte Ereignisse		
Mortalität	<u>Woche 48:</u> OR: 0,50 [0,05;5,59] p=0,5769 0,3% vs. 0,6%  <u>Woche 60:</u> OR: n.b.  <u>Woche 96:</u> OR: 0,50 [0,05;5,59], p=0,5769 0,3% vs. 0,6%	Kein Zusatznutzen oder Schaden ableitbar
Gesamtrate UEs	<u>Woche 48:</u> <b>OR: 0,92 [0,51;1,68], p=0,80</b> <b>93,7% vs. 94,4%</b>  <u>Woche 60:</u> OR: 1,05 [0,18;6,18], p=0,9590 91,7% vs. 91,3%  <u>Woche 96:</u> OR: 0,90 [0,38;2,14], p=0,8041 96,8% vs. 97,2%	Kein Zusatznutzen oder Schaden ableitbar
Gesamtrate SUEs	<u>Woche 48:</u> <b>OR: 1,77 [1,06;2,97], p=0,03</b> <b>10,9% vs. 6,7%</b>  <u>Woche 60:</u> OR: 0,96 [0,08;11,12], p=0,9717 4,2% vs. 4,3%  <u>Woche 96:</u> OR: 1,85 [1,17;2,93], p=0,0083 16,1% vs. 9,4%	Endpunktkategorie: Schwerwiegende Nebenwirkungen  <u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden <sup>a</sup> Wahrscheinlichkeit: Beleg  <u>Kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden:

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	<p><i>Interaktionstest: p=0,14</i></p> <p>Kaukasisch: OR: 2,49 [1,33;4,66], p=0,0043</p> <p>Nicht-kaukasisch: OR: 1,25 [0,63;2,48], p=0,5268</p>	<p>Geringer Schaden<sup>a</sup> Wahrscheinlichkeit: Hinweis</p> <p><u>Nicht-kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden<sup>a</sup> Wahrscheinlichkeit: Anhaltspunkt</p>
Therapieabbrüche aufgrund UEs	<p><u>Woche 48:</u> <b>OR: 0,67 [0,33;1,36], p=0,27</b> <b>3,3% vs. 5,1%</b></p> <p><i>Interaktionstest: p=0,009</i></p> <p>Kaukasisch: OR: 1,48 [0,58;3,76], p=0,4062</p> <p>Nicht-kaukasisch: <b>OR: 0,16 [0,04;0,64], p=0,010</b></p> <p><u>Woche 60:</u> OR: 0,15 [0,01;3,95], p=0,2587 0,0% vs. 4,3%</p> <p><u>Woche 96:</u> OR: 0,70 [0,37;1,33], p=0,2782 4,9% vs. 6,8%</p> <p><i>Interaktionstest: p=0,04</i></p> <p>Kaukasisch: OR: 1,15 [0,53;2,51], p=0,7214</p> <p>Nicht-kaukasisch: OR: 0,24 [0,06;0,87], p=0,0303</p>	<p>Endpunktkategorie: Nicht- schwerwiegende Nebenwirkungen</p> <p><u>Nicht-kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen<sup>b</sup> Wahrscheinlichkeit: Beleg</p>
UEs Grad 3-4	<p><u>Woche 48:</u> <b>OR: 1,14 [0,73;1,78], p=0,57</b> <b>11,9% vs. 10,9%</b></p> <p><u>Woche 60:</u> OR: 0,46 [0,06;3,46], p=0,4483</p>	<p>Kein Zusatznutzen oder Schaden ableitbar</p>

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	4,2% vs. 8,7%  <u>Woche 96:</u> OR: 1,25 [0,84;1,88], p=0,2733 17,5% vs. 14,5%	
Neurologische & psychiatrische Ereignisse	<u>Woche 48:</u> <b>OR: 0,46 [0,34;0,61], p&lt;0,00001</b> <b>42,7% vs. 62,1%</b>  <i>Interaktionstest: p=0,16</i> <95% Therapieadhärenz: <b>OR: 0,65 [0,36;1,16], p=0,14</b> ≥95% Therapieadhärenz: <b>OR: 0,40 [0,29;0,57],</b> <b>p=&lt;0,00001</b>  <u>Woche 60:</u> OR: 0,55 [0,20;1,50], p=0,31 41,7% vs. 56,5%  <i>Interaktionstest: p=0,08</i> ≤350 Zellen/μL: OR: 1,67 [0,32;8,59], p=0,5415 >350 Zellen/μL: OR: 0,25 [0,06;0,98], p=0,0471  <u>Woche 96:</u> OR: 0,45 [0,33;0,61], p<0,001 46,6% vs. 65,9%  <i>Interaktionstest: p=0,18</i> <40 Jahre: OR: 0,54 [0,36;0,81], p=0,0027 ≥40 Jahre: OR: 0,35 [0,22;0,56], p<0,0001	Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen  <u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Beleg
Hautausschläge	<u>Woche 48:</u> <b>OR: 0,54 [0,38;0,76], p=0,0005</b> <b>17,2% vs. 27,7%</b>	Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	<p><i>Interaktionstest: p=0,12</i> &lt;40 Jahre: <b>OR: 0,73 [0,46;1,16], p=0,18</b> ≥40 Jahre: <b>OR: 0,30 [0,10;0,84], p=0,02</b></p> <p><i>Interaktionstest: p=0,03</i> Kaukasisch: <b>OR: 0,42 [0,28;0,65], p&lt;0,0001</b> Nicht-kaukasisch: <b>OR: 0,97 [0,52;1,84], p=0,94</b></p> <p><u>Woche 60:</u> OR: 0,65 [0,20;2,13], p=0,54 18,8% vs. 26,1%</p> <p><u>Woche 96:</u> OR: 0,61 [0,43;0,86], p=0,006 21,3% vs. 30,7%</p> <p><i>Interaktionstest: p=0,11</i> Kaukasisch: OR: 0,51 [0,33;0,77], p=0,0016 Nicht-kaukasisch: OR: 0,92 [0,50;1,67], p=0,7740</p>	<p><u>Kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Beleg</p> <p><u>Nicht-kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis</p>
Renale Ereignisse	<p><u>Woche 48:</u> OR: 5,12 [0,59;44,02], p=0,12 1,4% vs. 0,3%</p> <p><u>Woche 60:</u> OR: n.b.</p> <p><u>Woche 96:</u> POR: 4,60 [1,14;18,54], p=0,037 2,0% vs. 0,3%</p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden Wahrscheinlichkeit: Hinweis</p>
Knochenfrakturen	<p><u>Woche 48:</u> <b>OR: 1,06 [0,36;3,10], p=0,92</b> <b>1,8% vs. 1,6%</b></p>	Kein Zusatznutzen oder Schaden ableitbar



Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	<p><u>Woche 60:</u> OR: 1,48 [0,06;37,84], p=0,8111 2,1% vs. 0,0%</p> <p><u>Woche 96:</u> OR: 1,27 [0,50;3,26], p=0,64 2,9% vs. 2,3%</p>	
Erkrankungen des Gastrointestinaltrakts (SOC)	<p><u>Woche 48:</u> <b>OR: 1,44 [1,08;1,91], p=0,01</b> <b>53,3% vs. 44,8%</b></p> <p><i>Interaktionstest: p=0,15</i> ≤350 Zellen/μL: <b>OR: 1,87 [1,20;2,93], p=0,006</b> &gt;350 Zellen/μL: <b>OR: 1,22 [0,83;1,77], p=0,31</b></p> <p><u>Woche 60:</u> OR: 2,70 [0,94;7,75], p=0,0647 54,2% vs. 30,4%</p> <p><u>Woche 96:</u> OR: 1,34 [1,00;1,81], p=0,0538 60,6% vs. 53,4%</p> <p><i>Interaktionstest: p=0,03</i> ≤100.000 Kopien/mL: OR: 1,06 [0,74;1,53], p=0,7464 &gt;100.000 Kopien/mL: OR: 2,16 [1,27;3,65], p=0,0042</p> <p><i>Interaktionstest: p=0,15</i> ≤350 Zellen/μL: OR: 1,74 [1,10;2,75], p=0,0171 &gt;350 Zellen/μL: OR: 1,11 [0,75;1,66], p=0,5966</p> <p><i>Interaktionstest: p=0,09</i></p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden Wahrscheinlichkeit: Hinweis</p> <p><u>≤350 Zellen/μL:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden Wahrscheinlichkeit: Beleg</p> <p><u>&gt;350 Zellen/μL:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden Wahrscheinlichkeit: Hinweis</p>

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	<95% Therapieadhärenz: OR: 2,10 [1,14;3,85], p=0,0173 ≥95% Therapieadhärenz: OR: 1,15 [0,81;1,63], p=0,4243	
<i>Diarrhoe</i>	<u>Woche 48:</u> <b>OR: 1,29 [0,91;1,83], p=0,16</b> <b>22,7% vs. 18,7%</b>  <u>Woche 60:</u> OR: 1,76 [0,50,6,17], p=0,3742 27,1% vs. 17,4%  <u>Woche 96:</u> OR: 1,08 [0,76;1,53], p=0,6613 25,0% vs. 23,6%	Kein Zusatznutzen oder Schaden ableitbar
<i>Übelkeit</i>	<u>Woche 48:</u> <b>OR: 1,61 [1,09;2,38], p=0,02</b> <b>19,2% vs. 13,3%</b>  <i>Interaktionstest: p=0,18</i> ≤100.000 Kopien/mL: OR: 1,36 [0,84;2,21], p=0,2172 >100.000 Kopien/mL: OR: 2,46 [1,20;5,04], p=0,0135  <u>Woche 96:</u> OR: 1,58 [1,07;2,32], p=0,0213 21,8% vs. 15,1%  <i>Interaktionstest: p=0,11</i> ≤100.000 Kopien/mL: OR: 1,26 [0,79;2,02], p=0,3321 >100.000 Kopien/mL: OR: 2,48 [1,24;4,98], p=0,0104	Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen  <u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden Wahrscheinlichkeit: Beleg  <u>&gt;100.000 Kopien/mL:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden Wahrscheinlichkeit: Beleg  <u>≤100.000 Kopien/mL:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Schaden Wahrscheinlichkeit: Hinweis

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
Allgemeine Erkrankungen und Beschwerden am Verbreichungsort (SOC)	<p><u>Woche 48:</u> <b>OR: 0,64 [0,46;0,89], p=0,0009</b> <b>20,5% vs. 28,3%</b></p> <p><i>Interaktionstest: p=0,08</i> ≤350 Zellen/μL: <b>OR: 0,46 [0,27;0,75], p=0,002</b> &gt;350 Zellen/μL: <b>OR: 0,83 [0,53;1,31], p=0,43</b></p> <p><u>Woche 60:</u> OR: 0,76 [0,25;2,29], p=0,6288 25,0% vs. 30,4%</p> <p><u>Woche 96:</u> OR: 0,72 [0,52;1,01], p=0,0537 24,4% vs. 31,0%</p> <p><i>Interaktionstest: p=0,03</i> Männlich: OR: 0,63 [0,44;0,90], p=0,0111 Weiblich: OR: 2,15 [0,75;6,13], p=0,1527</p> <p><i>Interaktionstest: p=0,04</i> ≤350 Zellen/μL: OR: 0,49 [0,29;0,81], p=0,0052 &gt;350 Zellen/μL: OR: 0,98 [0,62;1,53], p=0,9214</p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Hinweis</p> <p><u>≤350 Zellen/μL:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Beleg</p> <p><u>&gt;350 Zellen/μL:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Hinweis</p>
Fatigue	<p><u>Woche 48:</u> <b>OR: 0,89 [0,58;1,37], p=0,60</b> <b>12,1% vs. 13,1%</b></p> <p><u>Woche 60:</u> OR: 0,95 [0,25;3,55], p=0,9392 16,7% vs. 17,4%</p> <p><u>Woche 96:</u> OR: 0,86 [0,56;1,32], p=0,4810</p>	Kein Zusatznutzen oder Schaden ableitbar

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	12,9% vs. 14,8%	
Erkrankungen des Nervensystems (SOC)	<p><u>Woche 48:</u> <b>OR: 0,55 [0,41;0,75], p=0,0001</b> 27,8% vs. 41,1%</p> <p><i>Interaktionstest: p=0,0532</i> &lt;95% Therapieadhärenz: <b>OR: 0,94 [0,52;1,72], p=0,84</b> ≥95% Therapieadhärenz: <b>OR: 0,47 [0,33;0,67], p&lt;0,0001</b></p> <p><u>Woche 60:</u> OR: 0,85 [0,30;2,44], p=0,7661 31,3% vs. 34,8%</p> <p><u>Woche 96:</u> OR: 0,58 [0,42;0,78], p=0,0004 32,2% vs. 45,2%</p> <p><i>Interaktionstest: p=0,18</i> &lt;95% Therapieadhärenz: OR: 0,84 [0,46;1,54], p=0,5772 ≥95% Therapieadhärenz: OR: 0,52 [0,36;0,75], p=0,0004</p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Beleg</p>
Schwindel	<p><u>Woche 48:</u> <b>OR: 0,21 [0,13;0,34], p&lt;0,00001</b> 5,8% vs. 23,7%</p> <p><u>Woche 60:</u> OR: 0,06 [0,00;1,22], p=0,0674 0,0% vs. 13,0%</p> <p><u>Woche 96:</u> OR: 0,22 [0,14;0,36], p&lt;0,00001 7,2% vs. 25,9%</p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Beleg</p>
Kopfschmerzen	<p><u>Woche 48:</u> <b>OR: 1,43 [0,92;2,22], p=0,11</b> 14,1% vs. 10,1%</p>	Kein Zusatznutzen oder Schaden ableitbar

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	<p><i>Interaktionstest: p=0,18</i> Männlich: <b>OR: 1,66 [1,01;2,73], p=0,04</b> Weiblich: <b>OR: 0,68 [0,20;2,30], p=0,54</b></p> <p><i>Interaktionstest: p=0,19</i> Kaukasisch: <b>OR: 1,90 [1,06;3,39], p=0,03</b> Nicht-kaukasisch: <b>OR: 0,75 [0,22;2,61], p=0,65</b></p> <p><u>Woche 60:</u> OR: 0,81 [0,21;3,11], p=0,7599 14,6% vs. 17,4%</p> <p><u>Woche 96:</u> OR: 1,50 [0,97,2,31], p=0,0702 16,1% vs. 11,4%</p>	
Psychiatrische Erkrankungen (SOC)	<p><u>Woche 48:</u> <b>OR: 0,59 [0,44;0,79], p=0,0003</b> <b>33,6% vs. 46,4%</b></p> <p><u>Woche 60:</u> OR: 0,55 [0,20;1,50], p=0,2417 33,3% vs. 47,8%</p> <p><u>Woche 96:</u> OR: 0,64 [0,47;0,86], p=0,0030 39,7% vs. 50,9%</p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Beleg</p>
Abnorme Träume	<p><u>Woche 48:</u> <b>OR: 0,41 [0,22;0,78], p=0,006</b> <b>14,6% vs. 27,5%</b></p> <p><i>Interaktionstest: p=0,13</i> Kaukasisch: <b>OR: 0,26 [0,08;0,87], p=0,03</b></p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen</p>

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	<p>Nicht-kaukasisch: <b>OR: 0,73 [0,40;1,33], p=0,30</b></p> <p><u>Woche 60:</u> OR: 0,22 [0,06;0,77], p=0,0180 10,4% vs. 34,8%</p> <p><i>Interaktionstest: p=0,08</i> &lt;40 Jahre: OR: 0,09 [0,02;0,51], p=0,0064 ≥40 Jahre: OR: 1,29 [0,11;15,00], p=0,8411</p> <p><i>Interaktionstest: p=0,19</i> Männlich: OR: 0,16 [0,04;0,63], p=0,0085 Weiblich: OR: 2,14 [0,06;77,54], p=0,6772</p> <p><i>Interaktionstest: p=0,14</i> Kaukasisch: OR: 0,10 [0,02;0,56], p=0,0089 Nicht-kaukasisch: OR: 1,00 [0,08;12,56], p=1,0000</p> <p><u>Woche 96:</u> OR: 0,47 [0,32;0,69], p=0,0001 15,2% vs. 27,6%</p> <p><i>Interaktionstest: p=0,14</i> Kaukasisch: OR: 0,38 [0,24;0,62], p=0,0001 Nicht-kaukasisch: OR: 0,69 [0,37;1,26], p=0,2270</p>	<p>Wahrscheinlichkeit: Beleg</p> <p><u>Kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Beleg</p> <p><u>Nicht-kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis</p>
Schlaflosigkeit	<p><u>Woche 48:</u> <b>OR: 0,62 [0,39;0,98], p=0,04</b> <b>8,8% vs. 13,6%</b></p> <p><u>Woche 60:</u></p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u></p>

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	OR: 1,50 [0,28;8,08], p=0,6370 12,5% vs. 8,7%  <u>Woche 96:</u> OR: 0,63 [0,41;0,99], p=0,0428 10,9% vs. 16,2%	Ausmaß des Zusatznutzen/ Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Beleg
<i>Depression</i>	<u>Woche 48:</u> <b>OR: 0,85 [0,53;1,36], p=0,49</b> 9,3% vs. 10,9%  <u>Woche 96:</u> OR: 0,85 [0,55;1,32], p=0,4668 12,1% vs. 13,9%	Kein Zusatznutzen oder Schaden ableitbar
<i>Angst</i>	<u>Woche 48:</u> <b>OR: 0,76 [0,22;2,67], p=0,67</b> 4,5% vs. 4,5%  <u>Woche 60:</u> OR: 0,29 [0,04;1,87], p=0,1930 4,2% vs. 13,0%	Kein Zusatznutzen oder Schaden ableitbar
Erkrankungen der Haut und des Unterhautzellgewebes (SOC)	<u>Woche 48:</u> <b>OR: 0,60 [0,44;0,81], p=0,001</b> 25,8% vs. 36,5%  <i>Interaktionstest: p=0,04</i> Kaukasisch: <b>OR: 0,47 [0,32;0,69], p=0,0001</b> Nicht-kaukasisch: <b>OR: 0,94 [0,55;1,60], p=0,83</b>  <i>Interaktionstest: p=0,17</i> <95% Therapieadhärenz: <b>OR: 0,99 [0,53;1,85], p=0,97</b> ≥95% Therapieadhärenz: <b>OR: 0,54 [0,30;0,98], p=0,04</b>  <u>Woche 60:</u> OR: 0,85 [0,30;2,44], p=0,7661	Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen  <u>Kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Beleg  <u>Nicht-kaukasische Patienten:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis

Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
	<p>31,3% vs. 34,8%</p> <p><u>Woche 96:</u> OR: 0,65 [0,48;0,89], p=0,0070 31,9% vs. 41,8%</p> <p><i>Interaktionstest: p=0,07</i> Kaukasisch: OR: 0,53 [0,36;0,78], p=0,0013 Nicht-kaukasisch: OR: 0,96 [0,58;1,61], p=0,8896</p> <p><i>Interaktionstest: p=0,06</i> &lt;95% Therapieadhärenz: OR: 1,08 [0,58;2,00], p=0,8052 ≥95% Therapieadhärenz: OR: 0,55 [0,38;0,78], p=0,0010</p>	
Hautausschlag	<p><u>Woche 48:</u> <b>OR: 0,49 [0,30;0,81], p=0,006</b> <b>6,8% vs. 12,5%</b></p> <p><i>Interaktionstest: p=0,17</i> &lt;40 Jahre: <b>OR: 0,65 [0,35;1,24], p=0,19</b> ≥40 Jahre: <b>OR: 0,31 [0,13;0,72], p=0,006</b></p> <p><u>Woche 60:</u> OR: 0,55 [0,13;2,29], p=0,4130 10,4% vs. 17,4%</p> <p><u>Woche 96:</u> OR: 0,50 [0,30;0,82], p=0,0066 7,5% vs. 13,9%</p>	<p>Endpunktkategorie: Nicht-schwerwiegende Nebenwirkungen</p> <p><u>Gesamtpopulation:</u> Ausmaß des Zusatznutzen/ Schaden: Geringer Zusatznutzen Wahrscheinlichkeit: Beleg</p> <p><u>Patienten &lt;40 Jahre:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit: Hinweis</p> <p><u>Patienten ≥40 Jahre:</u> Ausmaß des Zusatznutzen/ Schaden: Beträchtlicher Zusatznutzen Wahrscheinlichkeit:</p>



Endpunkt	Effektschätzer [95%-KI], p-Wert Ereignisrate EVG/COBI/FTC/TDF vs. EFV/FTC/TDF	Endpunktkategorie/ Population für die Zusatznutzen abgeleitet wird/ Ausmaß des Zusatznutzen/ Wahrscheinlichkeit des Zusatznutzen
		Anhaltspunkt
<p><sup>a</sup> Rein formal wird hier ein Beleg für einen geringen Schaden in der Gesamtpopulation, bzw. ein Hinweis auf einen geringen Schaden in der Gruppe der kaukasischen Patienten und ein Anhaltspunkt auf einen geringen Schaden in der Gruppe der nicht-kaukasischen Patienten abgeleitet - medizinisch kann durch den fehlenden Zusammenhang zwischen Intervention und Nebenwirkung jedoch kein Schaden abgeleitet werden (Details siehe Gesamtfazit zum Zusatznutzen).</p> <p><sup>b</sup> Der in der nicht-kaukasischen Subgruppe gefundene Unterschied bei den Therapieabbrüchen aufgrund von UEs zugunsten von EVG/COBI/FTC/TDF lässt sich insbesondere auch auf Basis der vorliegenden Evidenz (eher höhere Nebenwirkungsraten unter Nicht-Kaukasieren) nicht schlüssig begründen und wird für die Nutzenableitung nicht weiter berücksichtigt</p> <p>Fett gedruckte Effektschätzer basieren auf einer Meta-Analyse der Studien S102 und S104</p> <p>KI = Konfidenzintervall, MWD = Mittelwertdifferenz, n.b. = nicht berechenbar, OR = Odds Ratio, POR = Peto Odds Ratio, UE = unerwünschtes Ereignis, SUE = schwerwiegendes unerwünschtes Ereignis</p>		

## 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]: 1,40 [0,93;2,11], p=0,12). Allerdings zeigen beide Studien die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze. Dies trifft sowohl auf die ITT- als auch die PP-Population zu. Sensitivitätsanalysen, unter Verwendung eines FEM, zeigen jeweils die Robustheit der Ergebnisse. Es besteht zu Woche 96 kein statistischer Unterschied zwischen den Behandlungen (OR [95%-KI]: 1,21 [0,81;1,79], p=0,35), jedoch kann auch hier, in der ITT- und der PP-Population, die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/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]: 1,29 [0,86;1,91], p=0,21; robuste Ergebnisse unter Verwendung eines FEM). Die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze wurde jedoch in beiden Studien gezeigt. Auch zu Woche 96 besteht die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/FTC/TDF zur 12%-Grenze; ein statistisch signifikanter Unterschied zwischen den Behandlungen besteht jedoch nicht (OR [95%-KI]: 1,23 [0,50;3,01], p=0,65).

Zusammenfassend lässt sich feststellen, dass hinsichtlich des virologischen Ansprechens der Nutzen von EVG/COBI/FTC/TDF äquivalent zu dem von EFV/FTC/TDF ist. Zusätzlich dazu konnte die Nichtunterlegenheit von EVG/COBI/FTC/TDF gegenüber EFV/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]: 0,88 [0,62;1,25], p=0,47) noch zu Woche 96 (HR [95%-KI]: 0,87 [0,63;1,19], p=0,28) 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 EFV/FTC/TDF. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### ***Veränderung der CD4-Zellzahl***

Bis Woche 48 stieg die CD4-Zellzahl unter einer Behandlung mit EVG/COBI/FTC/TDF statistisch signifikant stärker an als unter einer Behandlung mit EFV/FTC/TDF (MWD [95%-KI]: 36,26 [12,42;60,10], p=0,003). Eine Sensitivitätsanalyse, unter Verwendung eines FEM zeigt robuste Ergebnisse. Bei Berechnung der standardisierten Mittelwertdifferenz (Hedges' g) zeigt sich jedoch, dass dieser signifikante Unterschied zwischen den Behandlungen, gemäß IQWiG Methoden 4.0 [47], klinisch nicht relevant ist (Hedges' g [95%-KI]: 0,22 [0,07;0,37]). Weder zu Woche 60 (MWD [95%-KI]: 23,83 [-0,76;48,43], p=0,06) noch zu Woche 96 (MWD [95%-KI]: 22 [-10;54], p=0,19) kann ein statistisch signifikanter Unterschied zwischen den Behandlungen beobachtet werden.

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 ART: Diese sollte spätestens bei einer CD4-Zellzahl von 350 CD4-Zellen/ $\mu$ L begonnen werden [31]; ab einer CD4-Zellzahl von <200 CD4-Zellen/ $\mu$ L muss mit zunehmender Dauer der Immunsuppression vermehrt mit dem Auftreten von Aids-definierenden Erkrankungen gerechnet werden [21]. 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 in den ersten 48 Wochen ist somit direkt patientenrelevant. Hier besteht somit ein geringer Zusatznutzen für EVG/COBI/FTC/TDF in der Gesamtpopulation, der, obwohl der statistisch signifikante Unterschied zwar zu Woche 48, aber nicht mehr zu Woche 60 bzw. 96 besteht, mit der Wahrscheinlichkeit eines Belegs bewertet wird – der Einschätzung liegt die beschriebene deutlich höhere Patientenrelevanz des Endpunktes in den ersten 48 Wochen zugrunde.

### ***Mortalität***

Es besteht zu Woche 48 und 96 hinsichtlich der Mortalität kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI] jeweils: 0,50 [0,05;5,59] p=0,5769; robuste Ergebnisse unter Verwendung eines FEM). Zu Woche 60 (nur Studie S104) kann kein

Effektschätzer berechnet werden, da kein Patient in Studie S104 bis Woche 60 verstarb. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### ***Gesamtrate unerwünschter Ereignisse***

Zu Woche 48 besteht zwischen den beiden Behandlungen kein statistisch signifikanter Unterschied hinsichtlich der Gesamtrate unerwünschter Ereignisse (OR [95%-KI]: 0,92 [0,51;1,68],  $p=0,80$ ; robuste Ergebnisse unter Verwendung eines FEM). Auch zu Woche 60 (OR [95%-KI]: 1,05 [0,18;6,18],  $p=0,9590$ ) und Woche 96 (OR [95%-KI]: 0,90 [0,38;2,14],  $p=0,8041$ ) unterscheiden sich die beiden Behandlungen nicht statistisch signifikant. Hinsichtlich des Auftretens unerwünschter Ereignisse weist EVG/COBI/FTC/TDF somit einen ähnlichen Nutzen auf wie EFV/FTC/TDF. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### ***Gesamtrate schwerwiegender unerwünschter Ereignisse***

Hinsichtlich dem Auftreten schwerwiegender unerwünschter Ereignisse besteht zu Woche 48 ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,77 [1,06;2,97],  $p=0,03$ ; robuste Ergebnisse unter Verwendung eines FEM). Eine Sensitivitätsanalyse, unter Verwendung eines FEM zeigt robuste Ergebnisse. Unter einer Behandlung mit EVG/COBI/FTC/TDF traten bei 10,9% der Patienten und unter EFV/FTC/TDF bei 6,7% der Patienten schwerwiegende unerwünschte Ereignisse auf. Die am häufigsten beobachteten schwerwiegenden unerwünschten Ereignisse unter EVG/COBI/FTC/TDF waren aus der SOC Infektionen und parasitäre Erkrankungen. Bis Woche 48 wiesen von den bis zu diesem Zeitpunkt insgesamt 43 EVG/COBI/FTC/TDF-Patienten (S102: 41 (11,8%), S104: 2 (4,2%); EFV/FTC/TDF: S102: 6 (1,7%), S104: 1 (4,3%)) mit schwerwiegenden unerwünschten Ereignissen insgesamt allein 22 Patienten (S102: 20 (5,7%), S104: 2 (4,2%)) ein schwerwiegendes unerwünschtes Ereignis aus dem Bereich der Infektionen und parasitären Erkrankungen auf. Und auch zu Woche 96 wies ein Großteil der 56 EVG/COBI/FTC/TDF-Patienten mit schwerwiegenden unerwünschten Ereignissen ein Ereignis aus dem Bereich Infektionen und parasitäre Erkrankungen (nur Studie S102, EVG/COBI/FTC/TDF: 26 (7,5%), EFV/FTC/TDF: 10 (2,8%); OR [95%-KI]: 1,85 [1,17;2,93],  $p=0,0083$ ). In Studie S104 hingegen zeigte sich ebenso wie in Studie S103 [11] kein statistisch signifikanter Unterschied mehr zwischen den Behandlungen zu Woche 60 (OR [95%-KI]: 0,96 [0,08;11,12],  $p=0,9717$ ).

Ein erhöhter Anteil an Infektionen unter einer Behandlung im Gegensatz zur Vergleichstherapie, würde im HIV-Bereich zumeist ein nicht ausreichend gestärktes Immunsystem nahelegen. Da jedoch die Erhöhung der CD4-Zellzahlen unter EVG/COBI/FTC/TDF vergleichbar bzw. zu Woche 48 sogar signifikant stärker war als unter EFV/FTC/TDF, kann zumindest ein Versagen der Behandlung hinsichtlich Wirksamkeit und ein durch die HIV-1-Infektion geschwächtes Immunsystem als Grund ausgeschlossen werden. Darüber hinaus betraf der Unterschied dabei insbesondere die Infektionen Cellulitis und Appendicitis: Hier zeigte sich in Studie S102 ein deutlich größerer Anteil an EVG/COBI/FTC/TDF-Patienten als EFV/FTC/TDF-Patienten (Woche 48 Cellulitis: 4 (1,1%)

vs. 0, Appendicitis: 3 (0,9%) vs. 1 (0,3%); Woche 96: Cellulitis: 6 (1,7%) vs. 0, Appendicitis: 3 (0,9%) vs. 1 (0,3%)). Bei der Cellulitis handelt es sich um eine verursachte häufige, bakterielle Infektion von Dermis und Unterhautgewebe, primär verursacht durch Streptokokken und Staphylokokkus Aureus. Die Risikofaktoren für eine Cellulitis sind klar umschrieben und umfassen neben Hautläsionen (Risse, Schnitte oder Bisse) und kutanen Pilzinfektionen, vor allem Diabetes mellitus, ein geschwächtes Immunsystem beispielsweise durch HIV oder nach Chemotherapie, Lymphödeme und intravenösen Drogengebrauch. Laut den Fallberichten („patient narratives“ [12-17]) der EVG/COBI/FTC/TDF-Patienten mit schwerwiegender Cellulitis wies jeder dieser Patienten Risikofaktoren für Cellulitis auf. Es wurden zudem keine Untersuchungen zu Unbalanziertheiten bezüglich solcher Risikofaktoren durchgeführt und somit kann eine Unbalanziertheit der Risikofaktoren zuungunsten von EVG/COBI/FTC/TDF nicht ausgeschlossen werden. Ein Zusammenhang mit EVG/COBI/FTC/TDF ist für die Cellulitis ebenso wenig nachvollziehbar wie für die Appendicitis; alle aufgetretenen Fälle wurden vom jeweiligen Prüfarzt auch als nicht in Zusammenhang mit der Studienmedikation stehend beurteilt [18-20]. Des Weiteren ist hier anzumerken, dass in keiner anderen EVG/COBI/FTC/TDF-Studie (S103 und S104) eine besondere Anhäufung schwerwiegender Infektionen unter EVG/COBI/FTC/TDF auftrat. Ein Zusammenhang zwischen EVG/COBI/FTC/TDF und schweren Infektionen ist somit medizinisch nicht plausibel und das ungewöhnlich häufige Auftreten dieser schwerwiegenden Infektionen in Studie S102 kann hier nicht in einen kausalen Zusammenhang mit der Medikation gebracht werden. Rechnet man die Cellulitis- und Appendicitis-Fälle aus beiden Studien-Armen heraus, geht der statistisch signifikante Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF verloren.

Zu Woche 96 besteht ein Hinweis auf Effektmodifikation durch Rasse ( $p=0,14$ ;  $I^2=53,0\%$ ). Demnach besteht der statistisch signifikante Unterschied zuungunsten von EVG/COBI/FTC/TDF lediglich in der Gruppe der kaukasischen Patienten (OR [95%-KI]: 2,49 [1,33;4,66],  $p=0,0043$ ).

Zusammenfassend lässt sich hier feststellen, dass hinsichtlich des Auftretens schwerwiegender unerwünschter Ereignisse zwar ein signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF besteht, dieser jedoch vor allem durch ein in Studie S102 ungewöhnlich häufiges Auftreten schwerwiegender Infektionen unter EVG/COBI/FTC/TDF getrieben wurde. In zwei anderen Studien mit EVG/COBI/FTC/TDF (S103 und S104) traten schwerwiegende Infektionen unter EVG/COBI/FTC/TDF nicht häufiger auf als unter der Vergleichstherapie. Auch ein medizinischer Zusammenhang zwischen dem zu bewertenden Arzneimittel und schweren Infektionserkrankungen ist nicht plausibel, zumal diese auch von den Studienärzten als nicht in Zusammenhang stehend beurteilt wurden. Rein formal wird hier ein Beleg für einen geringen Schaden in der Gesamtpopulation, bzw. ein Hinweis auf einen geringen Schaden in der Gruppe der kaukasischen Patienten und ein Anhaltspunkt auf einen geringen Schaden in der Gruppe der nicht-kaukasischen Patienten abgeleitet - medizinisch kann durch den fehlenden

Zusammenhang zwischen Intervention und Nebenwirkung jedoch kein Schaden abgeleitet werden.

### ***Therapieabbrüche aufgrund unerwünschter Ereignisse***

Zu Woche 48 besteht hinsichtlich des Auftretens von Therapieabbrüchen aufgrund unerwünschter Ereignisse kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,67 [0,33;1,36],  $p=0,27$ ; robuste Ergebnisse unter Verwendung eines FEM). Und auch zu Woche 60 (OR [95%-KI]: 0,15 [0,01;3,95],  $p=0,2587$ ) und Woche 96 (OR [95%-KI]: 0,70 [0,37;1,33],  $p=0,2782$ ) kann zwischen den Behandlungen kein statistisch signifikanter Unterschied beobachtet werden.

Es besteht zu Woche 48 ( $p=0,009$ ;  $I^2=85,4\%$ ) und Woche 96 ( $p=0,04$ ;  $I^2=76\%$ ) jeweils ein Beleg für eine Effektmodifikation durch Rasse. Demnach besteht zu beiden Zeitpunkten in der Population der nicht-kaucasischen Patienten ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Zu Woche 48 ergibt sich das OR [95%-KI] zu 0,16 [0,04;0,64] mit  $p=0,010$  und zu Woche 96 beträgt das OR [95%-KI] 0,24 [0,06;0,87] mit  $p=0,0303$ .

Formal ließe sich hier ein Beleg für einen beträchtlichen Zusatznutzen in der nicht-kaucasischen Patientengruppe ableiten. Der in der nicht-kaucasischen Subgruppe gefundene Unterschied bei den Therapieabbrüchen aufgrund von UEs zugunsten von EVG/COBI/FTC/TDF lässt sich allerdings insbesondere auch auf Basis der vorliegenden Evidenz (eher höhere Nebenwirkungsraten unter Nicht-Kaukasiern) nicht schlüssig begründen und wird für die Nutzenableitung nicht weiter berücksichtigt.

### ***Unerwünschte Ereignisse Grad 3-4***

Hinsichtlich des Auftretens unerwünschter Ereignisse Grad 3-4 besteht zu keinem der hier betrachteten Zeitpunkte ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: Woche 48: 1,14 [0,73;1,78],  $p=0,57$  [robuste Ergebnisse unter Verwendung eines FEM]; Woche 60: 0,46 [0,06;3,46],  $p=0,4483$ ; Woche 96: 1,25 [0,84;1,88],  $p=0,2733$ ). Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### ***Neurologische & psychiatrische Ereignisse***

Hinsichtlich der Vermeidung neurologischer & psychiatrischer Ereignisse besteht sowohl zu Woche 48 (OR [95%-KI]: 0,46 [0,34;0,61],  $p<0,00001$ ; robuste Ergebnisse unter Verwendung eines FEM) als auch zu Woche 96 (OR [95%-KI]: 0,45 [0,33;0,61],  $p<0,001$ ) ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. Auch unter Ausschluss der als erhöht verzerrt bewerteten Studie S104 besteht dieser signifikante Unterschied. Zu Woche 60 konnte ein statistisch signifikanter Unterschied zwischen den Behandlungen trotz eines numerischen Vorteils von EVG/COBI/FTC/TDF nicht gezeigt werden. Dies kann durch die geringe Patientenzahl und der damit verbundenen geringen Power der Studie S104 bedingt sein. Unter einer Behandlung

mit EVG/COBI/FTC/TDF treten somit deutlich seltener neurologische & psychiatrische Ereignisse auf als unter EFV/FTC/TDF.

Es besteht zu Woche 48 ein Hinweis auf eine Effektmodifikation durch Therapieadhärenz ( $p=0,16$ ;  $I^2=48,3\%$ ). Demnach besteht der signifikante Unterschied in der Vermeidung der neurologischen & psychiatrischen Ereignisse zwischen den Behandlungen nur dann, wenn die Medikation  $\geq 95\%$  adhärenz eingenommen wird. Da Subgruppenanalysen, stratifiziert nach Therapieadhärenz, allgemein als nicht aussagekräftig bewertet werden (siehe Abschnitt 4.3.1.3.2), ergibt sich hieraus keine Konsequenz für die Ableitung eines Zusatznutzens. Zu Woche 60 besteht ein Hinweis auf Effektmodifikation durch CD4-Zellzahl zu Baseline ( $p=0,08$ ;  $I^2=67,0\%$ ). Da dieser Hinweis nur zu Woche 60 besteht, zu der nur die für diesen Endpunkt erhöht verzerrt bewertete Studie S104 Daten berichtet, wird dieser Hinweis auf Effektmodifikation als nicht relevant für die Ableitung eines Zusatznutzens bewertet. Der zu Woche 96 bestehende Hinweis auf Effektmodifikation ( $p=0,18$ ;  $I^2=44,2\%$ ) durch Alter ist quantitativer Natur und wirkt sich somit nicht auf die Bewertung des Zusatznutzens aus.

Zusammenfassend kann hier ein beträchtlicher Zusatznutzen von EVG/COBI/FTC/TDF hinsichtlich der Vermeidung neurologischer & psychiatrischer Ereignisse für die Gesamtpopulation abgeleitet werden. Da der signifikante Vorteil zu Woche 48 auf zwei Studien basiert und auch zu Woche 96 noch besteht, wird die Wahrscheinlichkeit des Zusatznutzens als Beleg eingeschätzt.

### **Hautausschläge**

Es besteht zu Woche 48 hinsichtlich des Auftretens von Hautausschlägen ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,54 [0,38;0,76],  $p=0,0005$ ; robuste Ergebnisse unter Verwendung eines FEM). Dieser signifikante Unterschied zugunsten von EVG/COBI/FTC/TDF besteht auch noch nach 96 Wochen Behandlung (OR [95%-KI]: 0,61 [0,43;0,86],  $p=0,006$ ). Studie S104 zeigt zu Woche 60 keinen statistisch signifikanten Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,65 [0,20;2,13],  $p=0,54$ ).

Zu Woche 48 besteht ein Hinweis auf Effektmodifikation durch Alter ( $p=0,12$ ;  $I^2=58,3\%$ ). Dieser zeigt an, dass sich der statistisch signifikanter Unterschied zwischen den Behandlungen, zugunsten von EVG/COBI/FTC/TDF, auf die  $\geq 40$ -jährigen Patienten beschränkt (OR [95%-KI]: 0,30 [0,10;0,84],  $p=0,02$ ). Zu Woche 60 und Woche 96 besteht dieser Hinweis auf Effektmodifikation allerdings nicht mehr. Es besteht allerdings zu Woche 48 ein Beleg ( $p=0,03$ ;  $I^2=78,4\%$ ) und zu Woche 96 ein Hinweis ( $p=0,11$ ;  $I^2=59,9\%$ ) auf Effektmodifikation durch Rasse. Das OR [95%-KI] für die kaukasischen Patienten beträgt zu Woche 48 0,42 [0,28;0,65] mit  $p<0,0001$  und zu Woche 96 0,51 [0,33;0,77] mit  $p=0,0016$ .

Zusammenfassend besteht hier ein beträchtlicher Zusatznutzen von EVG/COBI/FTC/TDF in der kaukasischen Patientenpopulation. Die Wahrscheinlichkeit des Zusatznutzens wird als Beleg eingestuft. Und da der Gesamtschätzer ebenfalls einen statistisch signifikanten

Unterschied zugunsten EVG/COBI/FTC/TDF anzeigt, besteht für die nicht-kaucasische Population ein Hinweis auf einen beträchtlichen Zusatznutzen.

### ***Renale Ereignisse***

Zu Woche 48 besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 5,12 [0,59;44,02],  $p=0,12$ ; robuste Ergebnisse unter Verwendung eines FEM). Zu Woche 60 ist ein Schätzer nicht berechenbar, da in Studie S104 bis Woche 60 keine renalen Ereignisse auftraten. Zu Woche 96 besteht jedoch ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (POR [95%-KI]: 4,60 [1,14;18,54],  $p=0,037$ ). Da hier die Berechnung eines Konfidenzintervalls für ein OR bzw. RR (Normalapproximation) inkonsistente Ergebnisse zu dem p-Wert mittels Fisher's exaktem Test aus dem CTR lieferten, wurde ein Peto Odds Ratio berechnet. Da dieser statistisch signifikante Unterschied zuungunsten von EVG/COBI/FTC/TDF nur zu Woche 96 besteht, kann hier nur ein Hinweis für einen geringen Schaden abgeleitet werden, der sich auf die Gesamtpopulation bezieht.

### ***Knochenfrakturen***

Es besteht zu keinem der hier betrachteten Zeitpunkte ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: Woche 48: 1,06 [0,36;3,10],  $p=0,92$  [robuste Ergebnisse unter Verwendung eines FEM]; Woche 60: 1,48 [0,06;37,84],  $p=0,8111$ ; Woche 96: 1,27 [0,50;3,26],  $p=0,64$ ). Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### ***Erkrankungen des Gastrointestinaltrakts (SOC)***

Zu Woche 48 besteht ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen (OR [95%-KI]: 1,44 [1,08;1,91],  $p=0,01$ ; robuste Ergebnisse unter Verwendung eines FEM) zuungunsten von EVG/COBI/FTC/TDF. Jedoch zeigt sich zu Woche 60 (OR [95%-KI]: 2,70 [0,94;7,75],  $p=0,0647$ ) und Woche 96 (OR [95%-KI]: 1,34 [1,00;1,81],  $p=0,0538$ ) kein statistisch signifikanter Unterschied mehr zwischen den Behandlungen.

Ein Hinweis auf Effektmodifikation durch Rasse besteht durch CD4-Zellzahl zu Baseline sowohl zu Woche 48 ( $p=0,15$ ;  $I^2=52,6\%$ ) als auch zu Woche 96 ( $p=0,15$ ;  $I^2=51,9\%$ ). Dabei beschränkt sich der statistisch signifikante Unterschied zuungunsten von EVG/COBI/FTC/TDF zu beiden Zeitpunkten lediglich auf die Subgruppe der Patienten mit CD4-Zellzahl zu Baseline von  $\leq 350$  Zellen/ $\mu\text{L}$ . Das OR [95%-KI] zu Woche 48 beträgt 1,87 [1,20;2,93] mit  $p=0,006$  und zu Woche 96 1,74 [1,10;2,75] mit  $p=0,0171$ . Außerdem besteht zu Woche 96 ein Beleg für Effektmodifikation durch Ausgangsviruslast ( $p=0,03$ ;  $I^2=78,6\%$ ). Für Patienten mit Ausgangsviruslast  $>100.000$  Kopien/mL beträgt das OR [95%-KI] 2,16 [1,27;3,65] mit  $p=0,0042$ . Des Weiteren besteht ein Hinweis auf Effektmodifikation durch Therapieadhärenz ( $p=0,09$ ;  $I^2=64,2\%$ ). Da Subgruppenanalysen, stratifiziert nach Therapieadhärenz, allgemein als nicht aussagekräftig bewertet werden, ergibt sich hieraus keine Konsequenz für die Ableitung eines Zusatznutzens.

Zusammenfassend ergibt sich hier ein Hinweis für einen geringen Schaden in der Gesamtpopulation. Für Patienten mit CD4-Zellzahl zu Baseline  $\leq 350$  Zellen/ $\mu\text{L}$  besteht ein Beleg für einen geringen Schaden durch EVG/COBI/FTC/TDF und in der Gruppe der Patienten mit CD4-Zellzahl zu Baseline  $> 350$  Zellen/ $\mu\text{L}$  besteht dieser geringe Schaden mit der Wahrscheinlichkeit eines Hinweises. Allerdings wird postuliert, dass es einen komplexen, kombinierten Effekt aus HIV-Infektion und ART in Bezug auf die Inzidenz der gastrointestinalen Symptome gibt – das bedeutet, dass die Mehrheit der gastrointestinalen Nebenwirkungen in einigen Studien möglicherweise nicht in Zusammenhang mit der ART stehen könnten [22]. Klinische Daten zeigen, dass sich die gastrointestinalen Nebenwirkungen durch die Therapie einer fortgeschrittenen HIV-Infektion als Ergebnis der immunologischen Rekonstitution verbessern. Dies steht im Einklang mit den gefundenen Ergebnissen, in denen sich ein Unterschied zwischen den Armen lediglich in den kränkeren Subgruppen mit Ausgangsviruslast  $> 100.000$  Kopien/mL, resp. CD4-Zellzahl zu Baseline  $\leq 350$  Zellen/ $\mu\text{L}$  zeigt. Der Beleg für Effektmodifikation durch Ausgangsviruslast wird als durch die Effektmodifikation durch CD4-Zellzahl zu Baseline abgedeckt angesehen.

### **Diarrhoe**

Hinsichtlich des Auftretens von Diarrhoe besteht zu keinem der hier betrachteten Zeitpunkte ein statistisch signifikanter Unterschied zwischen den Behandlungen. Zu Woche 48 beträgt das OR [95%-KI] 1,29 [0,91;1,83] mit  $p=0,16$  (robuste Ergebnisse unter Verwendung eines FEM), zu Woche 60 beträgt es 1,76 [0,50,6,17] mit  $p=0,3742$  und zu Woche 96 ergibt es sich zu 1,08 [0,76;1,53] mit  $p=0,6613$ . Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### **Übelkeit**

Zu Woche 48 besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,61 [1,09;2,38],  $p=0,02$ ; robuste Ergebnisse unter Verwendung eines FEM). Und auch zu Woche 96 zeigt sich ein statistisch signifikanter Unterschied zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,58 [1,07;2,32],  $p=0,0213$ ).

Zu Woche 48 ( $p=0,18$ ;  $I^2=45,0\%$ ) und zu Woche 96 ( $p=0,11$ ;  $I^2=59,8\%$ ) besteht jeweils ein Hinweis auf Effektmodifikation durch Ausgangsviruslast. Dabei beschränkt sich der statistisch signifikante Unterschied zuungunsten von EVG/COBI/FTC/TDF lediglich auf die Subgruppe von Patienten mit Ausgangsviruslast  $> 100.000$  Kopien/mL. Das OR [95%-KI] zu Woche 48 ergibt sich zu 2,46 [1,20;5,04],  $p=0,0135$  und zu Woche 96 beträgt es 2,48 [1,24;4,98],  $p=0,0104$ .

Zusammenfassend ergibt sich hier für die Gesamtpopulation ein Beleg für einen geringen Schaden. Außerdem besteht für Patienten mit Ausgangsviruslast  $> 100.000$  Kopien/mL ein Beleg für einen geringen Schaden durch EVG/COBI/FTC/TDF. Da auch der Gesamtschätzer statistisch signifikant ist, ergibt sich für Patienten mit Ausgangsviruslast  $\leq 100.000$  Kopien/mL ein Hinweis auf einen geringen Schaden. Da dieser Endpunkt aber durch die SOC der Erkrankungen des Gastrointestinaltrakts abgedeckt wird, wird dieser



Endpunkt nicht zusätzlich in der Ableitung eines gesamten Zusatznutzens bzw. Schadens berücksichtigt.

### ***Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC)***

Zu Woche 48 besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,64 [0,46;0,89], p=0,0009; robuste Ergebnisse unter Verwendung eines FEM). Dieser Unterschied besteht zu Woche 60 (OR [95%-KI]: 0,76 [0,25;2,29], p=0,6288) und zu Woche 96 (OR [95%-KI]: 0,72 [0,52;1,01], p=0,0537) nicht mehr.

Es besteht zu Woche 48 ein Hinweis (p=0,08; I<sup>2</sup>=67,4%) und zu Woche 96 (p=0,04; I<sup>2</sup>=75,6%) ein Beleg für Effektmodifikation durch CD4-Zellzahl zu Baseline. Der statistisch signifikante Unterschied zugunsten von EVG/COBI/FTC/TDF besteht dabei jeweils in der Subgruppe der Patienten mit CD4-Zellzahl zu Baseline ≤350 Zellen/μL. Zu Woche 48 beträgt das OR [95%-KI] 0,46 [0,27;0,75] mit p=0,002 und zu Woche 96 0,49 [0,29;0,81] mit p=0,0052. Ein weiterer Beleg für Effektmodifikation besteht zu Woche 96 durch Geschlecht (p=0,03; I<sup>2</sup>=78,8%). 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.

Zusammenfassend besteht hier ein Hinweis auf einen geringen Zusatznutzen durch EVG/COBI/FTC/TDF in der Gesamtpopulation. Außerdem kann hier für die Population der Patienten, die eine CD4-Zellzahl zu Baseline ≤350 Zellen/μL aufweisen, ein geringer Zusatznutzen für EVG/COBI/FTC/TDF hinsichtlich der Vermeidung allgemeiner Erkrankungen und Beschwerden am Verabreichungsort abgeleitet werden. Die Wahrscheinlichkeit wird als Beleg eingestuft. Für Patienten mit höherer CD4-Zellzahl zu Baseline, besteht der geringe Zusatznutzen noch mit der Wahrscheinlichkeit eines Hinweises.

### ***Fatigue***

Zu keinem der hier betrachteten Zeitpunkte besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen. Zu Woche 48 beträgt das OR [95%-KI] 0,89 [0,58;1,37] (robuste Ergebnisse unter Verwendung eines FEM) mit p=0,60, zu Woche 60 0,95 [0,25;3,55] mit p=0,9392 und zu Woche 96 ergibt es sich zu 0,86 [0,56;1,32] mit p=0,4810. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### ***Erkrankungen des Nervensystems (SOC)***

Hinsichtlich der Vermeidung von Erkrankungen des Nervensystems besteht zu Woche 48 ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,55 [0,41;0,75], p=0,0001; robuste Ergebnisse unter Verwendung eines FEM). Dieser Unterschied besteht auch zu Woche 96 (OR [95%-KI]: 0,58 [0,42;0,78], p=0,0004). An Woche 60 (nur S104) kann kein statistisch signifikanter Unterschied zwischen den Behandlungen gezeigt werden (OR [95%-KI]: 0,85 [0,30;2,44], p=0,7661).

Es besteht zu Woche 48 ( $p=0,0532$ ;  $I^2=73,2\%$ ) und 96 ( $p=0,18$ ;  $I^2=44,8\%$ ) jeweils ein Hinweis auf Effektmodifikation durch Therapieadhärenz. Da Subgruppenanalysen, stratifiziert nach Therapieadhärenz, allgemein als nicht aussagekräftig bewertet werden, ergibt sich hieraus keine Konsequenz für die Ableitung eines Zusatznutzens.

Zusammenfassend besteht hier für die Gesamtpopulation ein beträchtlicher Zusatznutzen durch EVG/COBI/FTC/TDF. Da er zu Woche 48 und 96 besteht, wird die Wahrscheinlichkeit des Zusatznutzens als Beleg eingestuft.

### ***Schwindel***

Zu Woche 48 besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,21 [0,13;0,34],  $p<0,00001$ ; robuste Ergebnisse unter Verwendung eines FEM). Dieser statistisch signifikante Unterschied besteht auch noch nach 96 Wochen Behandlung (OR [95%-KI]: 0,22 [0,14;0,36],  $p<0,00001$ ). In Studie S104 erreicht der Unterschied trotz des deutlichen numerischen Vorteils von EVG/COBI/FTC/TDF zu Woche 60 keine statistische Signifikanz (OR [95%-KI]: 0,06 [0,00;1,22],  $p=0,0674$ ). Dies liegt vermutlich an der geringen Patientenzahl in dieser Studie. Insgesamt lässt sich hier ein beträchtlicher Zusatznutzen durch EVG/COBI/FTC/TDF für die Gesamtpopulation ableiten, der mit der Wahrscheinlichkeit eines Belegs eingeschätzt wird. Da dieser Endpunkt aber durch die SOC der Erkrankungen des Nervensystems abgedeckt wird, wird dieser Endpunkt nicht zusätzlich in der Ableitung eines gesamten Zusatznutzens bzw. Schadens berücksichtigt.

### ***Kopfschmerzen***

Zu keinem der hier betrachteten Zeitpunkte besteht ein statistisch signifikanter Unterschied zwischen den beiden Behandlungen. Zu Woche 48 beträgt das OR [95%-KI] 1,43 [0,92;2,22] mit  $p=0,11$  (robuste Ergebnisse unter Verwendung eines FEM), zu Woche 60 0,81 [0,21;3,11] mit  $p=0,7599$  und zu Woche 96 1,50 [0,97;2,31] mit  $p=0,0702$ .

Zu Woche 48 gibt es jeweils einen Hinweis auf Effektmodifikation durch Geschlecht ( $p=0,18$ ;  $I^2=43,3\%$ ) und durch Rasse ( $p=0,19$ ;  $I^2=43,0\%$ ). 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. In der Gruppe der kaukasischen Patienten besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zuungunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 1,90 [1,06;3,39],  $p=0,03$ ).

Zusammenfassend wird hier kein Zusatznutzen oder Schaden abgeleitet. Die Effektmodifikation durch Rasse besteht lediglich zu einem der beiden hier relevanten Zeitpunkte und ist außerdem nur als Hinweis eingestuft. Außerdem ist dieser Endpunkt durch die SOC der Erkrankungen des Nervensystems abgedeckt. Somit wird dieser Endpunkt nicht zusätzlich in der Ableitung eines gesamten Zusatznutzens bzw. Schadens berücksichtigt.

### ***Psychiatrische Erkrankungen (SOC)***

Zu Woche 48 besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,59 [0,44;0,79],  $p=0,0003$ ; robuste Ergebnisse unter Verwendung eines FEM). Dieser signifikante Unterschied kann zwar in der kleinen Studie S104 zu Woche 60 nicht gezeigt werden (OR [95%-KI]: 0,55 [0,20;1,50],  $p=0,2417$ ), zu Woche 96 besteht der statistisch signifikante Unterschied zugunsten EVG/COBI/FTC/TDF allerdings (OR [95%-KI]: 0,64 [0,47;0,86],  $p=0,0030$ ). Insgesamt besteht hier somit ein Beleg für einen beträchtlichen Zusatznutzen durch EVG/COBI/FTC/TDF in der Gesamtpopulation.

### ***Abnorme Träume***

Zu Woche 48 besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF (OR [95%-KI]: 0,41 [0,22;0,78],  $p=0,006$ ; robuste Ergebnisse unter Verwendung eines FEM). Dieser signifikante Unterschied zugunsten von EVG/COBI/FTC/TDF besteht auch sowohl zu Woche 60 (OR [95%-KI]: 0,22 [0,06;0,77],  $p=0,0180$ ) als auch zu Woche 96 (OR [95%-KI]: 0,47 [0,32;0,69],  $p=0,0001$ ).

Es bestehen zu allen drei Zeitpunkten Hinweise auf Effektmodifikation durch Rasse (Woche 48:  $p=0,13$ ;  $I^2=56,0\%$ , Woche 60:  $p=0,14$ ;  $I^2=53,6\%$ , Woche 96:  $p=0,14$ ;  $I^2=54,4\%$ ). Demnach besteht in der Gruppe der kaukasischen Patienten ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF. Zu Woche 48 beträgt das OR [95%-KI] 0,26 [0,08;0,87] mit  $p=0,03$ , zu Woche 60 0,10 [0,02;0,56] mit  $p=0,0089$  und zu Woche 96 ergibt es sich zu 0,38 [0,24;0,62] mit  $p=0,0001$ . Außerdem bestehen zu Woche 60 noch jeweils ein Hinweis auf Effektmodifikation durch Alter ( $p=0,08$ ;  $I^2=67,0\%$ ) und durch Geschlecht ( $p=0,19$ ;  $I^2=42,4\%$ ). 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. Da der Hinweis auf Effektmodifikation durch Alter lediglich zum Zeitpunkt Woche 60 besteht, welcher nur auf der kleinen Studie S104 basiert, erscheint es nicht plausibel einen Zusatznutzen für diese Population abzuleiten.

Zusammenfassend ergibt sich ein Beleg für einen beträchtlichen Zusatznutzen von EVG/COBI/FTC/TDF in der Gesamtpopulation. Außerdem besteht ein Beleg für einen beträchtlichen Zusatznutzen in der kaukasischen Population, und in der nicht-kaukasischen Population besteht dieser mit der Wahrscheinlichkeit eines Hinweises. Da dieser Endpunkt aber durch die SOC der psychiatrischen Erkrankungen abgedeckt wird, wird dieser Endpunkt nicht zusätzlich in der Ableitung eines gesamten Zusatznutzens bzw. Schadens berücksichtigt.

### ***Schlaflosigkeit***

Es besteht zu Woche 48 (OR [95%-KI]: 0,62 [0,39;0,98],  $p=0,04$ ; robuste Ergebnisse unter Verwendung eines FEM) und zu Woche 96 (OR [95%-KI]: 0,63 [0,41;0,99],  $p=0,0428$ ) jeweils ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von EVG/COBI/FTC/TDF. An Woche 60 besteht kein statistisch signifikanter Unterschied

zwischen den Behandlungen (OR [95%-KI]: 1,50 [0,28;8,08],  $p=0,6370$ ). Zusammenfassend ergibt sich ein Beleg für einen geringen Zusatznutzen in der Gesamtpopulation. Da dieser Endpunkt aber durch die SOC der psychiatrischen Erkrankungen abgedeckt wird, wird dieser Endpunkt nicht zusätzlich in der Ableitung eines gesamten Zusatznutzens bzw. Schadens berücksichtigt.

### ***Depression***

Weder zu Woche 48 (OR [95%-KI]: 0,85 [0,53;1,36],  $p=0,49$ ; robuste Ergebnisse unter Verwendung eines FEM) noch zu Woche 96 (OR [95%-KI]: 0,85 [0,55;1,32],  $p=0,4668$ ) kann ein statistisch signifikanter Unterschied zwischen den Behandlungen festgestellt werden. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### ***Angst***

Weder zu Woche 48 (OR [95%-KI]: 0,76 [0,22;2,67],  $p=0,67$ ; robuste Ergebnisse unter Verwendung eines FEM) noch zu Woche 60 (OR [95%-KI]: 0,29 [0,04;1,87],  $p=0,1930$ ) kann ein statistisch signifikanter Unterschied zwischen den Behandlungen festgestellt werden. Ein Zusatznutzen oder Schaden ist nicht ableitbar.

### ***Erkrankungen der Haut und des Unterhautzellgewebes (SOC)***

Zu Woche 48 besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,60 [0,44;0,81],  $p=0,001$ ; robuste Ergebnisse unter Verwendung eines FEM). Und auch nach 96 Wochen kann ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF gezeigt werden (OR [95%-KI]: 0,65 [0,48;0,89],  $p=0,0070$ ). Lediglich zu Woche 60 (nur S104) besteht kein statistisch signifikanter Unterschied zwischen den Behandlungen (OR [95%-KI]: 0,85 [0,30;2,44],  $p=0,7661$ ).

Es besteht zu Woche 48 ( $p=0,17$ ;  $I^2=46,0\%$ ) und 96 ( $p=0,06$ ;  $I^2=71,4\%$ ) jeweils ein Hinweis auf Effektmodifikation durch Therapieadhärenz. Da jedoch die Subgruppenanalyse, stratifiziert nach Therapieadhärenz, allgemein als nicht aussagekräftig bewertet wird, ergibt sich hieraus keine Konsequenz für die Ableitung eines Zusatznutzens. Außerdem gibt es zu Woche 48 ( $p=0,04$ ;  $I^2=76,6\%$ ) einen Beleg und zu Woche 96 einen Hinweis auf Effektmodifikation durch Rasse ( $p=0,07$ ;  $I^2=70,2\%$ ). In der Gruppe der kaukasischen Patienten beträgt das OR [95%-KI] zu Woche 48 0,47 [0,32;0,69] mit  $p=0,0001$  und zu Woche 96 0,53 [0,36;0,78] mit  $p=0,0013$ .

Zusammenfassend ergibt sich hier in der kaukasischen Population ein Beleg für einen beträchtlichen Zusatznutzen. Für die nicht-kaukasische Population besteht der beträchtliche Zusatznutzen mit der Wahrscheinlichkeit eines Hinweises.

### ***Hautausschlag***

Zu Woche 48 (OR [95%-KI]: 0,49 [0,30;0,81],  $p=0,006$ ; robuste Ergebnisse unter Verwendung eines FEM) und zu Woche 96 (OR [95%-KI]: 0,50 [0,30;0,82],  $p=0,0066$ ) besteht ein statistisch signifikanter Unterschied zwischen den Behandlungen zugunsten von

EVG/COBI/FTC/TDF. Zu Woche 60 kann, trotz eines numerischen Vorteils von EVG/COBI/FTC/TDF, kein statistisch signifikanter Unterschied zwischen den Behandlungen gezeigt werden (OR [95%-KI]: 0,55 [0,13;2,29],  $p=0,4130$ ).

Es besteht zu Woche 48 ein Hinweis auf Effektmodifikation durch Alter ( $p=0,17$ ;  $I^2=48,1\%$ ). Demnach besteht ein statistisch signifikanter Unterschied zugunsten von EVG/COBI/FTC/TDF in der Gruppe der  $\geq 40$ -jährigen Patienten. Das OR [95%-KI] beträgt 0,31 [0,13;0,72] mit  $p=0,006$ .

Zusammenfassend ergibt sich hier ein Beleg für einen geringen Zusatznutzen von EVG/COBI/FTC/TDF in der Gesamtpopulation. Außerdem besteht in der Gruppe der  $\geq 40$ -jährigen Patienten ein beträchtlicher Zusatznutzen, der, da der Hinweis auf Effektmodifikation nur zu Woche 48 besteht, als Hinweis eingeschätzt wird. Für die  $< 40$ -jährigen Patienten ergibt sich ein Anhaltspunkt für einen beträchtlichen Zusatznutzen. Da dieser Endpunkt aber durch die SOC der Erkrankungen der Haut und der Unterhautzellgewebes 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. 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 [7, 9, 10].

In der vorliegenden Nutzenbewertung wurde anhand von patientenrelevanten und validen Endpunkten gezeigt, dass EVG/COBI/FTC/TDF, bei gleichwertiger Wirksamkeit wie EFV/FTC/TDF (Nichtunterlegenheit zur 12%-Grenze wurde gezeigt), eine vom Nebenwirkungsprofil verbesserte Behandlungsoption darstellt als die zweckmäßige Vergleichstherapie EFV/FTC/TDF. Durch EVG/COBI/FTC/TDF kann der durch die zweckmäßige Vergleichstherapie EFV/FTC/TDF ausgelöste patientenrelevante Schaden vermindert werden.

Es konnte gezeigt werden, dass EVG/COBI/FTC/TDF eine gleichwertige bzw. nichtunterlegene Wirksamkeit wie EFV/FTC/TDF besitzt. Dies zeigt sich anhand der Endpunkte virologisches Ansprechen, Zeit bis zum Verlust des virologischen Ansprechens und Veränderung der CD4-Zellzahl. Hinsichtlich der Veränderung der CD4-Zellzahl zeigte sich zu Woche 48 sogar ein Vorteil für EVG/COBI/FTC/TDF. 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

ART: Diese sollte gemäß Leitlinien spätestens bei einer CD4-Zellzahl von 350 CD4-Zellen/ $\mu$ L begonnen werden [31]; ab einer CD4-Zellzahl von <200 CD4-Zellen/ $\mu$ L muss mit zunehmender Dauer der Immunsuppression vermehrt mit dem Auftreten von Aids-definierenden Erkrankungen gerechnet werden [43]. 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.

Bei sicherheitsbezogenen Endpunkten zeigten sich unter EVG/COBI/FTC/TDF deutliche patientenrelevante Vorteile hinsichtlich der Vermeidung von Therapieabbrüchen aufgrund unerwünschter Ereignisse, der Vermeidung neurologischer & psychiatrischer Ereignisse, von Hautausschlägen, allgemeiner Erkrankungen und Beschwerden am Verabreichungsort (darunter Fatigue), Erkrankungen des Nervensystems (darunter Schwindel), psychiatrischer Erkrankungen (darunter abnorme Träume und Schlaflosigkeit) und Erkrankungen der Haut und des Unterhautzellgewebes (darunter Hautausschlag). Bei der Ableitung eines Gesamtfazits zum Zusatznutzen werden die Gruppen von PT (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 Gesamtzusatznutzens bei.

Erkrankungen der Haut und des Unterhautzellgewebes werden von den Betroffenen nicht nur als sehr unangenehm empfunden, sondern stellen auch einen 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 Hautausschlag im Kreis der HIV-infizierten Bevölkerung somit 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 Wiederaufstieg der Viruslast mit konsekutivem virologischen Therapieversagen, immunologischer Verschlechterung und erhöhter Morbidität und Mortalität führen kann [1, 2].

Auch weitere direkt durch die Behandlung verursachte Erkrankungen des Nervensystems und psychiatrische Erkrankungen können die Adhärenz mit den beschriebenen klinischen Folgen negativ beeinflussen. Zudem stellen sie für Patienten äußerst unangenehme Belastungen dar, die sie in ihrem Alltag und Berufsleben stark beeinträchtigen und zu einer stark reduzierten Lebensqualität führen. Zu diesen zählen Nebenwirkungen wie Schwindel, Schlaflosigkeit, abnorme Träume und Alpträume. Die hohe und persistierende Inzidenz von Schwindel stellt ein besonderes Risiko bei Personen dar, die beruflich Maschinen bedienen, Kraftfahrzeuge lenken und/oder hochkonzentriert arbeiten müssen. Die hohe Patientenrelevanz des

zentralnervösen Nebenwirkungsspektrums zeigen Daten aus verschiedenen Kohorten (Real-Life-Setting): In einer englischen Kohorte mussten etwa 20% aller Patienten, die eine Therapie mit EFV/FTC/TDF beginnen, innerhalb der ersten 42 Therapiewochen (Median) umgestellt werden – und zwar primär aufgrund der ZNS-Toxizität von Efavirenz [3]. In einer skandinavischen Kohorte zeigte sich, dass etwa 50% der Patienten unter einer EFV-haltigen Initialtherapie innerhalb des ersten Behandlungsjahres umgestellt werden müssen [4]. Zudem besteht das Risiko, dass die Patienten ihre Therapie aufgrund der subjektiv belastenden Nebenwirkungen unkontrolliert abbrechen, was zu Resistenzentwicklung, Verlust an Therapieoptionen und Krankheitsprogression führen kann. Auch der Großteil der Therapieabbrüche in den hier vorliegenden Studien ist psychiatrischer Natur.

Die SOC allgemeine Erkrankungen und Beschwerden am Verabreichungsort beschreibt eine eher heterogene Gruppe von Ereignissen; sie umfasst einerseits allgemeine Erkrankungen wie Fatigue, Fieber und Schmerzen, und berücksichtigt andererseits auch Beschwerden am Verabreichungsort. Während die Beschwerden am Verabreichungsort bei einer oral einzunehmenden Therapie von untergeordneter Bedeutung sind, umfassen die allgemeinen Erkrankungen eine ganze Reihe von Ereignissen mit hoher Patientenrelevanz. Die unter dieser SOC beobachteten Ereignisse in Studie S102 zu Woche 48 umfassen Ereignisse wie unter anderem Fatigue, Pyrexie und Brustschmerzen, die alle direkten Einfluss auf das Wohlbefinden des Patienten haben und somit patientenrelevant sind.

Durch die bessere Verträglichkeit sinkt 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 [5, 6]. 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 ebenfalls zu bedenken ist, dass eine ART durch die Übertragung resistenter Viren von Anfang an komplizierter und im Zweifel auch teurer wird [7]. 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 [73].

Insgesamt zeigen verschiedene Daten, dass beeinträchtigende Nebenwirkungen die Therapieadhärenz und damit Persistenz generell negativ beeinflussen, was wiederum das Risiko der Bildung von Resistenzen und virologischem Versagen mit 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 [7, 9, 10].

Neben den positiven sicherheitsbezogenen Aspekten wurden auch einige Nachteile von EVG/COBI/FTC/TDF beobachtet. Dies zeigte sich bei der Gesamtrate schwerwiegender unerwünschter Ereignisse, renalen Ereignissen und Erkrankungen des Gastrointestinaltrakts (darunter Übelkeit).

Die Detail-Analyse zeigt, dass der gefundene Unterschied bei den beobachteten schwerwiegenden unerwünschten Ereignissen primär getrieben wurde durch einen Unterschied zwischen beiden Behandlungsarmen in der SOC Infektionen und parasitäre Erkrankungen in Studie S102 – dieser Unterschied zeigte sich weder in Studie S104 noch in Studie S103 [11]. Die ist insofern überraschend, als dass ein erhöhter Anteil an Infektionen unter einer bestimmten Behandlung im Gegensatz zur Vergleichstherapie, im HIV-Bereich zumeist ein unter dieser Behandlung nicht ausreichend gestärktes Immunsystem nahelegen würde (siehe Modul 3A) - die Erhöhung der CD4-Zellzahlen unter EVG/COBI/FTC/TDF war jedoch vergleichbar bzw. zu Woche 48 sogar signifikant stärker war als unter EFV/FTC/TDF. Damit kann ein Versagen der Behandlung ebenso wie ein durch die HIV-1-Infektion geschwächtes Immunsystem als Grund ausgeschlossen werden. Darüber hinaus fällt auf, dass der Unterschied zwischen den Armen dabei insbesondere auf die Krankheitsbilder Cellulitis und Appendicitis zurückzuführen ist: Hier zeigte sich in Studie S102 ein deutlich größerer Anteil an EVG/COBI/FTC/TDF-Patienten (Woche 48: Cellulitis 4 [1,1%], Appendicitis 3 [0,9%]) als EFV/FTC/TDF-Patienten (Woche 48: Cellulitis 0 [0,0%], Appendicitis 1 [0,3%]). Bei der Cellulitis handelt es sich um eine verursachte häufige, bakterielle Infektion von Dermis und Unterhautgewebe, primär verursacht durch Streptokokken und Staphylokokkus Aureus. Die Risikofaktoren für eine Cellulitis sind klar umschrieben und umfassen neben Hautläsionen (Risse, Schnitte oder Bisse) und kutanen Pilzinfektionen vor allem Diabetes mellitus, ein geschwächtes Immunsystem (beispielsweise durch HIV oder nach Chemotherapie), Lymphödeme und intravenösen Drogengebrauch. Laut der individuellen Fallberichte („patient narratives“) der EVG/COBI/FTC/TDF-Patienten mit schwerwiegender Cellulitis wies jeder dieser Patienten Risikofaktoren für Cellulitis auf; entsprechend wurden alle aufgetretenen Fälle vom jeweiligen Prüfarzt als nicht in Zusammenhang mit der Studienmedikation stehend beurteilt. Ein Zusammenhang mit EVG/COBI/FTC/TDF ist für die Cellulitis damit ebenso wenig nachvollziehbar wie für die Appendicitis; auch die Appendicitis-Fälle wurden vom jeweiligen Prüfarzt als nicht assoziiert mit der Studienmedikation bewertet. Zudem wurden keine Untersuchungen zu Unbalanziertheiten bezüglich solcher Risikofaktoren durchgeführt und somit kann eine Unbalanziertheit der Risikofaktoren zuungunsten von EVG/COBI/FTC/TDF nicht ausgeschlossen werden. Des Weiteren ist hier anzumerken, dass in keiner anderen EVG/COBI/FTC/TDF-Studie (S103 und S104) ein Unterschied zuungunsten EVG/COBI/FTC/TDF gefunden wurde. Ein Zusammenhang zwischen EVG/COBI/FTC/TDF und schweren Infektionen ist medizinisch nicht plausibel und somit kann das ungewöhnlich häufige Auftreten dieser Infektionen in Studie S102 hier nicht in einen kausalen Zusammenhang mit der Medikation gebracht werden. Korrigiert man um die Cellulitis- und Appendicitis-Fälle, geht der statistisch signifikante Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF verloren.



Bei den renalen Ereignissen zeigt sich ein statistisch signifikanter Unterschied zwischen den beiden Armen zuungunsten von EVG/COBI/FTC/TDF; renale Ereignisse waren dabei selten und traten in Studie S102 bei 5 Patienten unter EVG/COBI/FTC/TDF und 1 Patienten unter EFV/FTC/TDF auf. Unter denselben zu vergleichenden Therapien trat in Studie S104 bei keinem Patienten ein renales Ereignis auf. Potentielle renale Nebenwirkungen im Sinne einer Schädigung des proximalen Tubulus unter Tenofovir sind seit langem bekannt, und werden durch den „Risk Management Plan“ durch die Europäische Zulassungsbehörde adäquat überwacht. Der hier vorliegende statistisch signifikante Unterschied zwischen EVG/COBI/FTC/TDF und EFV/FTC/TDF hinsichtlich renaler Ereignisse ist insofern überraschend, als dass Tenofovir als potentiell verursachendes Agens in gleicher Dosierung in beiden Armen enthalten ist – entsprechend findet sich in Studie S104 ebenso wie im Vergleich zwischen EVG/COBI/FTC/TDF und ATV/r+FTC/TDF in Studie S103 kein Unterschied hinsichtlich des Auftretens renaler Ereignisse [11].

Erkrankungen des Gastrointestinaltrakts gehören zu den häufigsten Nebenwirkungen; sie können insbesondere in der Anfangsphase bei fast allen antiretroviralen Arzneimitteln auftreten, wobei NNRTI ein günstiges gastrointestinales Nebenwirkungsprofil aufweisen, geboosterte Protease-Inhibitoren hingegen häufig mit gastrointestinalen Nebenwirkungen assoziiert sind [21]. Hill et al. zeigen, dass es einen komplexen, kombinierten Effekt aus HIV-Infektion und ART in Bezug auf die Inzidenz gastrointestinaler Symptome gibt – das bedeutet, dass die Mehrheit der gastrointestinalen Nebenwirkungen in einigen Studien nicht in Zusammenhang mit der ART stehen könnten. Entsprechend gibt es klinische Daten, die zeigen, dass sich die gastrointestinalen Nebenwirkungen durch die Therapie einer fortgeschrittenen HIV-Infektion als Ergebnis der immunologischen Rekonstitution verbessern [22]. Dies steht im Einklang mit den gefundenen Ergebnissen: Während der Unterschied in den kränkeren Subgruppen mit Ausgangsviruslast  $>100.000$  Kopien/mL, resp. CD4-Zellzahl  $\leq 350$  CD4-Zellen/ $\mu$ L signifikant zuungunsten von EVG/COBI/FTC/TDF ist, zeigt sich bei den Patienten mit Ausgangsviruslast  $\leq 100.000$  Kopien/mL, resp. CD4-Zellzahl  $>350$  CD4-Zellen/ $\mu$ L kein Unterschied zwischen den Behandlungen. Möglicherweise führt der zusätzliche gastrointestinale Einfluss der HIV-Infektion zu einem additiven Effekt von Infektion und Medikation, der bei den kränkeren Patienten unter EVG/COBI/FTC/TDF ausgeprägter ist als unter dem gastrointestinal gut verträglichen EFV-basierten Regime. Verbessert sich also der Immunstatus durch die Therapie, ist davon auszugehen, dass der Unterschied wieder verschwindet. Die Transienz der gastrointestinalen Nebenwirkungen bestätigt auch eine kürzlich veröffentlichte Analyse der Studien S102 und S103, in der gezeigt wird, dass die Inzidenz gastrointestinaler Nebenwirkungen EVG/COBI/FTC/TDF zu Beginn einer Therapie am höchsten ist und über die Zeit deutlich abnimmt [23]. Diarrhoe lässt sich auch nicht-medikamentös sehr gut behandeln, desgleichen Übelkeit - insbesondere wenn diese, wie im Fall von EVG/COBI/FTC/TDF, nicht in Verbindung mit Erbrechen steht. Die überschaubare klinische Relevanz dieser Nebenwirkung wird zudem deutlich in der Tatsache, dass lediglich ein Patient aufgrund von gastrointestinalen Nebenwirkungen (hier Übelkeit) abgebrochen hat. Des Weiteren wurde der Großteil der Erkrankungen des Gastrointestinaltrakts in Studie S102 zu Woche 48 als Grad 1-2 eingestuft.

Die Studiendaten zeigen außerdem deutliche Unterschiede in der Patientenrelevanz von neurologischen % psychiatrischen Nebenwirkungen einerseits und gastrointestinalen Ereignisse andererseits: Während in der gepoolten Analyse der Studien S102 und S104 lediglich ein Patient die Studie aufgrund von gastrointestinalen Nebenwirkungen abbrach, führten in Studie S102 neurologische (bis Woche 96: 4 Patienten) und psychiatrische (bis Woche 96: 13 Patienten) Ereignisse (mittels SOC) zu deutlich mehr Studien-Abbrüchen.

#### ***Ableitung eines Gesamt-Zusatznutzens***

Zur Abwägung des Gesamt-Zusatznutzens lässt sich ableiten, dass der Hinweis für einen geringen Schaden durch EVG/COBI/FTC/TDF bei Erkrankungen des Gastrointestinaltrakts durch den Hinweis auf einen geringen Zusatznutzen hinsichtlich allgemeiner Erkrankungen und Beschwerden am Verabreichungsort ausgeglichen wird. Ebenfalls ausgeglichen ist die Nutzen-Schaden-Abwägung in der Subgruppe der Patienten mit CD4-Zellzahl zu Baseline  $\leq 350$  Zellen/ $\mu\text{L}$ , in der einem Beleg für einen geringen Schaden hinsichtlich Erkrankungen des Gastrointestinaltrakts unter EVG/COBI/FTC/TDF ein Beleg für einen geringen Zusatznutzen hinsichtlich allgemeiner Erkrankungen unter EVG/COBI/FTC/TDF gegenübersteht.

Die tatsächliche Patientenrelevanz des Schadens hinsichtlich schwerwiegender unerwünschter Ereignisse wird als gering eingeschätzt, da er hauptsächlich auf die unter EVG/COBI/FTC/TDF ungewöhnlich häufig aufgetretenen schweren Infektionen in Studie S102 (vor allem Cellulitis und Appendicitis) zurückzuführen ist, die medizinisch nicht in Zusammenhang mit dem zu bewertenden Arzneimittel gebracht werden können und entsprechend auch vom Prüfarzt als nicht in Zusammenhang mit der Studienmedikation stehend beurteilt wurden. Medizinisch plausibel kann für schwerwiegende unerwünschte Ereignisse also kein Schaden abgeleitet werden.

Damit bleiben die Belege für einen beträchtlichen Zusatznutzen von EVG/COBI/FTC/TDF sowohl bei Hautausschlag als auch bei den Erkrankungen des Nervensystems (SOC), darunter Schwindel, und bei neurologischen & psychiatrischen Ereignissen (Gruppe von PT). Des Weiteren zeigt sich ein Beleg für einen beträchtlichen Zusatznutzen von EVG/COBI/FTC/TDF bei den psychiatrischen Ereignissen, darunter ein Beleg für einen beträchtlichen Zusatznutzen hinsichtlich der Vermeidung abnormer Träume, sowie ein Beleg für einen geringen Zusatznutzen von EVG/COBI/FTC/TDF im Hinblick auf Schlaflosigkeit.

Es zeigt sich außerdem ein Beleg für einen geringen Zusatznutzen durch EVG/COBI/FTC/TDF hinsichtlich der Veränderung der CD4-Zellzahl. Da durch eine raschere Dynamik der Entwicklung der CD4-Zellzahl das Immunsystem schneller rekonstituiert und so der für die Entwicklung Aids-definierender Erkrankungen relevante, niedrige CD4-Zellbereich rascher verlassen wird, ist dieser Vorteil von besonderer Patientenrelevanz. Diesem Zusatznutzen steht ein Hinweis auf einen geringen Schaden hinsichtlich renaler Ereignisse gegenüber. Diese sind insgesamt selten und unter Tenofovir seit langem bekannt; ein entsprechender Risk-Management-Plan wurde implementiert.

Insgesamt sind somit unter einer Behandlung mit EVG/COBI/FTC/TDF deutlich weniger Nebenwirkungen zu beobachten als unter einer Behandlung mit EFV/FTC/TDF – dies betrifft insbesondere die Nebenwirkungen, die für die Patienten besonders relevant sind: Vor allem dermatologische, psychiatrische und neurologische Nebenwirkungen, führen zu einer erheblichen Beeinträchtigung des Patienten sowohl hinsichtlich seiner Funktionsfähigkeit als auch hinsichtlich seiner Lebensqualität. Derart beeinträchtigende Nebenwirkungen können die Therapieadhärenz negativ beeinflussen und somit das Risiko eines Therapieversagens mit Resistenzentwicklung und konsekutiver Krankheitsprogression erhöhen. Virologisches Versagen aufgrund von Resistenzentwicklung und nicht tolerierbarer Nebenwirkungen schränken in einer Indikation, die eine lebenslange tägliche Einnahme von Medikamenten erfordert, verfügbare Therapieoptionen ein.

In der Gesamtschau des auf Endpunktebene abgeleiteten Zusatznutzens (siehe Tabelle 4-138) verbleibt nach der Nutzen-Schaden Abwägung ein positiver Zusatznutzen für EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF, dessen Ausmaß als **gering** eingestuft wird. Da hier zwei randomisierte, aktiv-kontrollierte Studien zugrunde liegen, wird dieser Zusatznutzen mit der Wahrscheinlichkeit eines **Belegs** eingestuft.

Tabelle 4-138: Gegenüberstellung von Zusatznutzen und Schaden von EVG/COBI/FTC/TDF im Vergleich zu EFV/FTC/TDF

<b>Zusatznutzen durch EVG/COBI/FTC/TDF</b> Format: <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des ZN	<b>Schaden durch EVG/COBI/FTC/TDF</b> Format: <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des Schaden
<ul style="list-style-type: none"> <li>• <b>Veränderung der CD4-Zellzahl</b> Gesamt: Hinweis auf geringen ZN</li> <li>• <b>Therapieabbrüche aufgrund UEs</b> Nicht-kaukasisch: Beleg für beträchtlichen ZN</li> <li>• <b>Allg. Erkrankungen und Beschwerden am Verabreichungsort SOC</b> Gesamt: Hinweis auf geringen ZN ≤350 Zellen/μL: Beleg für geringen ZN &gt;350 Zellen/μL: Hinweis auf geringen ZN</li> <li>• <b>Erkrankungen des Nervensystems SOC</b> Gesamt: Beleg für beträchtlichen ZN <i>Schwindel</i> Gesamt: Beleg für beträchtlichen ZN <b>Neurologische &amp; psychiatrische Ereignisse (Gruppe von PTs)</b> Gesamt: Beleg für beträchtlichen ZN</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Gesamtrate SUEs</b> Gesamt: Beleg für geringen Schaden Kaukasisch: Hinweis auf geringen Schaden Nicht-kaukasisch: Anhaltspunkt für geringen Schaden</li> <li>• <b>Renale Ereignisse</b> Gesamt: Hinweis auf geringen Schaden</li> <li>• <b>Erkrankungen des Gastrointestinaltrakts SOC</b> Gesamt: Hinweis auf geringen Schaden ≤350 Zellen/μL: Beleg für geringen Schaden &gt;350 Zellen/μL: Hinweis auf geringen Schaden <i>Übelkeit</i> Gesamt: Beleg für geringen Schaden &gt;100.000 Kopien/mL: Beleg für geringen Schaden ≤100.000 Kopien/mL: Hinweis auf</li> </ul>

<b>Zusatznutzen durch EVG/COBI/FTC/TDF</b> Format: <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des ZN	<b>Schaden durch EVG/COBI/FTC/TDF</b> Format: <b>Endpunkt</b> Population: Wahrscheinlichkeit und Ausmaß des Schaden
<ul style="list-style-type: none"> <li>• <b>Psychiatrische Erkrankungen SOC</b>                Gesamt: Beleg für beträchtlichen ZN  <i>Abnorme Träume</i>                Gesamt: Beleg für beträchtlichen ZN                Kaukasisch: Beleg für beträchtlichen ZN                Nicht-kaukasisch: Hinweis auf beträchtlichen ZN  <i>Schlaflosigkeit</i>                Gesamt: Beleg für geringen ZN</li>   <li>• <b>Erkrankungen der Haut und des Unterhautzellgewebes SOC</b>                Kaukasisch: Beleg für beträchtlichen ZN                Nicht-kaukasisch: Hinweis auf beträchtlichen ZN  <i>Hautausschlag</i>                Gesamt: Beleg für beträchtlichen ZN                &lt;40 Jahre: Hinweis auf beträchtlichen ZN                ≥40 Jahre: Anhaltspunkt für beträchtlichen ZN  <b>Hautausschläge (Gruppe von PTs)</b>                Kaukasisch: Beleg für beträchtlichen ZN                Nicht-kaukasisch: Hinweis auf beträchtlichen ZN</li> </ul>	<p style="text-align: center;">geringen Schaden</p>
ZN = Zusatznutzen	

Die von Gilead Sciences vorgeschlagene zweckmäßige Vergleichstherapie umfasste zusätzlich zu der vom G-BA festgelegten und hier untersuchten zweckmäßigen Vergleichstherapie EFV/FTC/TDF zwei weitere Regime, i.e. ATV/r+FTC/TDF und RAL+FTC/TDF (siehe Modul 3), die ebenfalls vollumfänglich untersucht wurden:

In der Gesamtschau des auf Endpunktebene abgeleiteten Zusatznutzens verbleibt ein positiver Zusatznutzen für EVG/COBI/FTC/TDF im Vergleich zu ATV/r+FTC/TDF, der auf Basis einer randomisierten, aktiv-kontrollierten Studie mit der Wahrscheinlichkeit eines Hinweises eingestuft wird [11]. Im Vergleich zur zweckmäßigen Vergleichstherapie RAL+FTC/TDF wird für EVG/COBI/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 [24].

Damit ergibt sich insgesamt ein **Beleg** für einen **geringen** Zusatznutzen für EVG/COBI/FTC/TDF.

#### 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 4-139: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Antiretroviral nicht-vorbehandelte erwachsene Patienten mit einer HIV-1 Infektion	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 4-140: 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.*

Nicht zutreffend.

##### **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.*

Nicht zutreffend.

#### 4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005<sup>13</sup>, Molenberghs 2010<sup>14</sup>). 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 2006<sup>15</sup>) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006<sup>16</sup>) 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

<sup>13</sup> Burzykowski T (Ed.: *The evaluation of surrogate endpoints*. New York: Springer; 2005.

<sup>14</sup> 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.

<sup>15</sup> Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

<sup>16</sup> Weir CJ, Walley RJ. *Statistical evaluation of biomarkers as surrogate endpoints: a literature review*. *Stat Med* 2006; 25(2): 183-203.

- 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 [29]: 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 [30].
- 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.“ [31]. Die EMA empfiehlt entsprechend die Erfassung des Erreichens und Aufrechterhaltens des virologischen Ansprechens im Rahmen klinischer Studien [28]:

„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 [32]. 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 [33]. 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 [34]. Darüber hinaus steigt das Risiko eines virologischen Rebounds mit der Komplexität des Regimes [8].

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 [35].

Ein langfristiges, dauerhaftes virologisches Ansprechen führt zu einer fortschreitenden immunologischen Rekonstitution, wie die über Jahre ansteigenden CD4-Zellzahlen belegen [36].

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 [8, 37, 38]. Zudem ist das Risiko eines virologischen Rebounds bei nicht-vorbehandelten Patienten niedriger als bei behandelten Patienten, die ihre Therapie wechseln [8].

Dauer des virologischen Ansprechens und Persistenz sind somit kritische Parameter für die Reduktion von Morbidität und Mortalität, dies bedingt durch das assoziierte immunologische Ansprechen [39] 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 [40]:

„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 [28]:

„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 (2008) [41] 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/ $\mu$ 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 (1998) [42] 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 52: Behandlungen, die einen größeren Anstieg der CD4-Zellzahl bedingen, zeigen eine höhere Reduktion der Rate an klinischer Progression. Mit zwei Ausnahmen zeigen alle einen Anstieg der CD4-Zellzahl zusammen mit einem klinischen Vorteil; das Ausmaß des klinischen Vorteils war assoziiert mit dem Anstieg der CD4-Zellzahl.

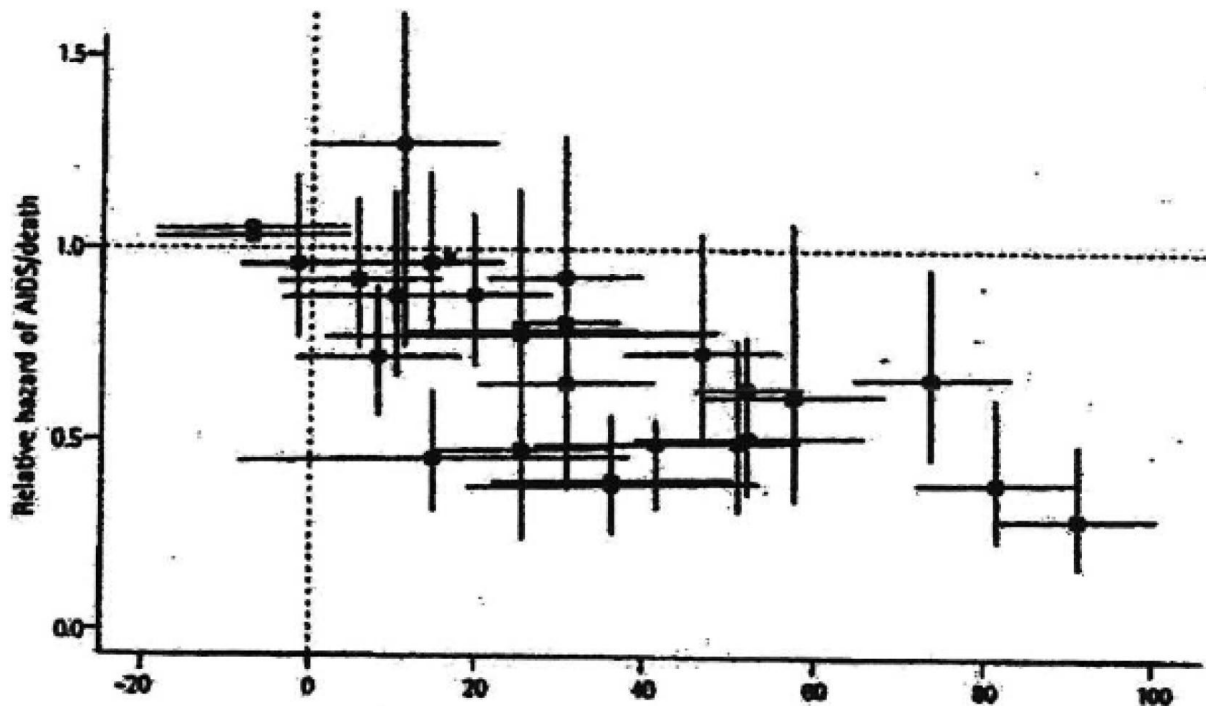


Abbildung 52: Relation zwischen Veränderung der CD4-Zellzahl und der Hazard für klinische Ereignisse (Aids/Tod) [42]

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 gemäß Leitlinien spätestens bei einer CD4-Zellzahl von 350 CD4-Zellen/ $\mu\text{L}$  begonnen werden [31]; ab einer CD4-Zellzahl von  $<200$  CD4-Zellen/ $\mu\text{L}$  muss mit zunehmender Dauer der Immunsuppression vermehrt mit dem Auftreten von Aids-definierenden Erkrankungen gerechnet werden [43]. 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 [58]
- Publikation [60]
- Registereintrag [54, 55]
- Post-hoc Analysen [65, 68]

**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 [59]
- Publikation [74]
- Registereintrag [54, 55]
- Post-hoc Analysen [66, 69]

**Studie GS-US-236-0104****96 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 [61]
- Publikation [62]

- Registereintrag [56, 57]
- Post-hoc Analysen [63, 64, 67]

#### 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] 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.
- [2] 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.
- [3] Scourfield A, Zheng J, Chinthapalli S, Waters L, Martin T, Mandalia S, et al. Discontinuation of Atripla as first-line therapy in HIV-1 infected individuals. *Aids*. 2012 Jul 17;26(11):1399-401.
- [4] Leutscher PD, Stecher C, Storgaard M, Larsen CS. Discontinuation of efavirenz therapy in HIV patients due to neuropsychiatric adverse effects. *Scandinavian journal of infectious diseases*. 2013 Feb 21.
- [5] 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 North America 2002-2009. *Aids*. 2013 Mar 13;27(5):803-13.
- [6] 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.
- [7] 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).
- [8] 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.
- [9] 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.
- [10] 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.
- [11] Gilead Sciences GmbH. Modul 4: Direkter Vergleich von EVG/COBI/FTC/TDF und ATV/r+FTC/TDF. Zur Behandlung der Infektion mit dem Humanen

- Immundefizienzvirus 1 (HIV-1) bei Erwachsenen im Alter von 18 Jahren und darüber, die nicht antiretroviral vorbehandelt sind. 2013.
- [12] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 2728-6184. 2012.
- [13] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 1602-6155. 2012.
- [14] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 1598-6330. 2012.
- [15] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 0744-6032. 2012.
- [16] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 0744-6021. 2012.
- [17] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 0033-6339. 2012.
- [18] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 2475-6073. 2012.
- [19] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 1407-6018. 2012.
- [20] Gilead Sciences Inc. GS-US-236-0102: Patient Narrative 0947-6565. 2012.
- [21] Hoffmann C, Rockstroh JK, Hrsg. HIV 2012. [www.hivbuch.de](http://www.hivbuch.de); Medizin Fokus Verlag; 2012.
- [22] 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.
- [23] 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.
- [24] Gilead Sciences GmbH. Modul 4: Indirekter Vergleich von EVG/COBI/FTC/TDF und RAL+FTC/TDF. Zur Behandlung der Infektion mit dem Humanen Immundefizienzvirus 1 (HIV-1) bei Erwachsenen im Alter von 18 Jahren und darüber, die nicht antiretroviral vorbehandelt sind. 2013.
- [25] 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 nicht-vorbehandelten erwachsenen Patienten. 2013 21.02.2013.
- [26] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Pixantron – Nutzenbewertung gemäß § 35a SGB V. IQWiG-Berichte – Nr. 154. . 2013.
- [27] Gilead Sciences GmbH. Fachinformation Stribild 150 mg/150 mg/200 mg/245 mg Filmtabletten (Elvitegravir/Cobicistat/Emtricitabin/Tenofoviridisoproxil). Stand der Information: Mai 2013. 2013.
- [28] European Medicines Agency. Guideline on the Clinical Development of Medicinal Products for the Treatment of HIV Infection. EMEA/CPMP/EWP/633/02. 2008.
- [29] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Rilpivirin – Nutzenbewertung gemäß § 35a SGB V. IQWiG-Berichte – Nr. 127. . 2012.
- [30] 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.
- [31] 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).
- [32] 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.

- [33] 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.
- [34] 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.
- [35] 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.
- [36] 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.
- [37] 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.
- [38] 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.
- [39] 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.
- [40] Gemeinsamer Bundesausschuss. HIV-Resistenztestung. Unterausschuss „Ärztliche Behandlung“ des Gemeinsamen Bundesausschusses, 2004.
- [41] 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.
- [42] 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.
- [43] Hoffmann C. 6.11. Monitoring der ART. In: Hoffmann C, Rockstroh JK, Hrsg. HIV 2012: Medizin Fokus Verlag; 2012. S. 171-95.
- [44] Gilead Sciences GmbH. Fachinformation Atripla 600 mg/200 mg/245 mg Filmtabletten (Efavirenz/Emtricitabin/Tenofoviridisoproxil). Stand der Information: Januar 2013. 2013.
- [45] Bristol-Myers Squibb GmbH & Co. KGaA. Fachinformation Sustiva 600 mg Filmtabletten (Efavirenz). Stand der Information: Juni 2012. 2012.
- [46] 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).
- [47] 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).
- [48] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj*. 2003 Sep 6;327(7414):557-60.
- [49] Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. 2008.
- [50] 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.
- [51] Hoffmann C. 6. ART. In: Hoffmann C, Rockstroh JK, Hrsg. *HIV 2012: Medizin Fokus Verlag*; 2012. S. 60-263.
- [52] De Wit S, Sabin CA, Weber R, Worm SW, Reiss P, Cazanave 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.
- [53] 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.
- [54] 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>.
- [55] 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>.
- [56] 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>.
- [57] 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>.
- [58] 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.
- [59] 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.
- [60] 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.
- [61] 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.
- [62] 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.
- [63] Gilead Sciences Inc. Posthoc Analysis (5697) for Study GS-US-236-0104 Week 48. 2013.
- [64] Gilead Sciences Inc. Posthoc Analysis (5698) for Study GS-US-236-0104 Week 60. 2013.
- [65] Gilead Sciences Inc. Hazard Ratio for TLOVR Endpoint (5682) for Study GS-US-236-0102 Week 48. 2013.
- [66] Gilead Sciences Inc. Hazard Ratio for TLOVR Endpoint (5695) for Study GS-US-236-0102 Week 96. 2013.
- [67] Gilead Sciences Inc. Hazard Ratio for TLOVR Endpoint (5698) for Study GS-US-236-0104 Week 48. 2013.
- [68] Gilead Sciences Inc. Posthoc Analysis (5682) for Study GS-US-236-0102 Week 48. 2013.
- [69] Gilead Sciences Inc. Posthoc Analysis (5695) for Study GS-US-236-0102 Week 96. 2013.
- [70] Gilead Sciences GmbH. Subgruppenanalysen zu EVG/COBI/FTC/TDF im Vergleich zur zweckmäßigen Vergleichstherapie. 2013.
- [71] 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.
- [72] 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.
- [73] 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.
- [74] 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.

- [75] 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:

<b>Datenbankname</b>	EMBASE	
<b>Suchoberfläche</b>	Ovid	
<b>Datum der Suche</b>	08,12,2010	
<b>Zeitsegment</b>	1980 to 2010 week 50	
<b>Suchfilter</b>	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle <sup>17</sup> ] – Strategy minimizing difference between sensitivity and specificity	
<b>#</b>	<b>Suchbegriffe</b>	<b>Ergebnis</b>
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

<sup>17</sup> 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 nicht erforderlich.

**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:

<b>Studienregister</b>	clinicaltrials.gov
<b>Internetadresse</b>	<a href="http://www.clinicaltrials.gov">http://www.clinicaltrials.gov</a>
<b>Datum der Suche</b>	08,12,2010
<b>Suchstrategie</b>	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
<b>Treffer</b>	23

Die Suche nach registrierten Studien mit Elvitegravir/Cobicistat/Emtricitabin/Tenofovirdisoproxil (EVG/COBI/FTC/TDF) zur Behandlung der HIV-1-Infektion in antiretroviral nicht-vorbehandelten Patienten erfolgte anhand spezifischer Stichwörter. Die Suchen wurden über die jeweilige „Advanced Search“-Funktion der unten aufgeführten Studienregister durchgeführt.

**Elvitegravir/Cobicistat/Emtricitabin/Tenofovirdisoproxil (EVG/COBI/FTC/TDF)**

Die Suche wurde für Abschnitt 4.3.1.1.3 durchgeführt.

<b>Studienregister</b>	clinicaltrials.gov
<b>Internetadresse</b>	<a href="http://www.clinicaltrials.gov/ct2/search/advanced">http://www.clinicaltrials.gov/ct2/search/advanced</a>
<b>Datum der Suche</b>	09.04.2013
<b>Suchstrategie</b>	naive [Search Terms] AND 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** 9

Da das Studienregister [clinicalstudyresults.org](http://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** 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

**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.*

Da die erste Zulassung des zu bewertenden Arzneimittels nicht älter als 12 Monate ist (Zulassung USA: 27. August 2012) und keine bibliografische Literaturrecherche für die zweckmäßige Vergleichstherapie durchgeführt werden musste, wurden keine Volltexte gesichtet.



**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 4-141 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

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

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 4-142 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
<b>ggf. Zwischenüberschrift zur Strukturierung des Studienpools</b>						
<b>placebokontrolliert</b>						
<b>aktivkontrolliert, zweckmäßige Vergleichstherapie(n)</b>						
GS-US-236-0102	ja	ja	nein	ja [58, 59]	ja [54, 55]	ja [60, 74]
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.  b: Studie, für die der Unternehmer Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist.  c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

**Anhang 4-F: Methodik der eingeschlossenen Studien – RCT**

*Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-143 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 4-143 inklusive eines Flow-Charts für den Patientenfluss.*

Tabelle 4-143 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0102

Item <sup>a</sup>	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 Prüfärzten besteht
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<b>Einschlusskriterien:</b> <ul style="list-style-type: none"> <li>• Fähig, eine schriftliche Einwilligungserklärung zu verstehen und zu unterschreiben, welche vor dem Start der Studienprozeduren eingeholt werden musste</li> <li>• Plasma HIV-1-RNA Level <math>\geq 5000</math> Kopien/mL</li> <li>• Kein vorheriger Gebrauch jeglicher zugelassener oder investigativer antiretroviraler Medikamente zu irgendeiner Zeit</li> <li>• Screening HIV-1 Genotyp Report, durch Gilead bereitgestellt, musste Sensitivität auf Emtricitabin, Tenofoviridisoproxil und Efavirenz gezeigt haben</li> <li>• Normales Elektrokardiogramm (EKG); oder wenn unnormal, musste der Prüfärzte bestätigen dass es nicht klinisch relevant</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>ist</p> <ul style="list-style-type: none"> <li>• Adäquate Nierenfunktion: eGFR <math>\geq 70</math> mL/min gemäß Cockcroft-Gault (CG) Formel: Männer: <math>eGFR \text{ (mL/min)} = \frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})}</math> Frauen: <math>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</math></li> <li>• Lebertransaminasen (Aspartat-Aminotransferase (AST) und Alanin-Aminotransferase (ALT)) <math>\leq 5</math> x Obergrenze des Normbereichs (upper limit of normal, ULN)</li> <li>• Gesamt-Bilirubin <math>\leq 1,5</math> mg/dL oder normales direktes Bilirubin</li> <li>• Adäquate hämatologische Funktion (absolute Zahl neutrophiler Granulozyten <math>\geq 1000/\text{mm}^3</math>; Thrombozyten <math>\geq 50,000/\text{mm}^3</math>; Hämoglobin <math>\geq 8,5</math> g/dL)</li> <li>• Serum-Amylase <math>\leq 5</math> x ULN (Patienten mit Serum-Amylase <math>&gt; 5</math> x ULN blieben geeignet falls Serumlipase <math>\leq 5</math> x ULN war)</li> <li>• 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"> <li>○ Weibliche Patienten, die hormonelle Verhütungsmethoden nutzten müssen diese Methode für mindestens 3 Monate vor Studienbehandlung verwendet haben</li> <li>○ Weibliche Patienten, die vor <math>\geq 12</math> 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</li> </ul> </li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Referenzwerten des zentralen Labors, aufweisen.</p> <ul style="list-style-type: none"> <li>• 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</li> <li>• Alter <math>\geq 18</math> Jahre</li> <li>• Lebenserwartung <math>\geq 1</math> Jahr</li> </ul> <p><b>Ausschlusskriterien:</b></p> <ul style="list-style-type: none"> <li>• Diagnose eines neuen Aids-definierenden Zustands innerhalb von 30 Tagen vor dem Screening</li> <li>• Erhalt einer medikamentösen Behandlung des Hepatitis C-Virus (HCV), oder es wurde erwartet, dass sie im Laufe dieser Studie eine HCV-Behandlung erhalten würden</li> <li>• Erleiden einer dekompensierten Zirrhose (d.h. Aszites, Enzephalopathie, etc.)</li> <li>• Frauen die stillen</li> <li>• Positiver Serum-Schwangerschaftstest (gebärfähige Frauen)</li> <li>• Implantierter Defibrillator oder Schrittmacher</li> <li>• Aktueller Alkohol- oder Substanzgebrauch, welcher nach Meinung des Prüfarztes die Studien-Compliance des Patienten potenziell beeinträchtigt</li> <li>• 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</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation																														
		<ul style="list-style-type: none"> <li>• Aktive, ernste Infektionen (anders als HIV-1-Infektion), welche parenterale antibiotische oder antimykotische Therapie innerhalb 30 Tage vor Baseline erfordern.</li> <li>• 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</li> </ul> <table border="1" data-bbox="906 786 1401 2000"> <thead> <tr> <th data-bbox="906 786 1169 857">Medikamentenklasse</th> <th data-bbox="1169 786 1401 857">Verbotene Wirkstoffe*</th> </tr> </thead> <tbody> <tr> <td data-bbox="906 857 1169 929">Alpha-adrenerge Rezeptor-Antagonisten</td> <td data-bbox="1169 857 1401 929">Alfuzosin</td> </tr> <tr> <td data-bbox="906 929 1169 974">Analeptika</td> <td data-bbox="1169 929 1401 974">Modafinil</td> </tr> <tr> <td data-bbox="906 974 1169 1019">Antibiotika</td> <td data-bbox="1169 974 1401 1019">Telithromycin</td> </tr> <tr> <td data-bbox="906 1019 1169 1149">Antikonvulsiva</td> <td data-bbox="1169 1019 1401 1149">Phenobarbital, Phenytoin, Carbamazepin, Oxcarbazepin</td> </tr> <tr> <td data-bbox="906 1149 1169 1193">Antipilzmittel</td> <td data-bbox="1169 1149 1401 1193">Voriconazol</td> </tr> <tr> <td data-bbox="906 1193 1169 1265">Antihistaminika</td> <td data-bbox="1169 1193 1401 1265">Astemizol, Terfenadin</td> </tr> <tr> <td data-bbox="906 1265 1169 1364">Mittel gegen Mykobakterien</td> <td data-bbox="1169 1265 1401 1364">Rifampin, Rifapentin, Rifabutin</td> </tr> <tr> <td data-bbox="906 1364 1169 1408">Kalziumkanalblocker</td> <td data-bbox="1169 1364 1401 1408">Bepridil</td> </tr> <tr> <td data-bbox="906 1408 1169 1588">Ergotaminderivate</td> <td data-bbox="1169 1408 1401 1588">Ergotamin, Ergonovin Dihydroergotamin Methylergonovin Ergometrin</td> </tr> <tr> <td data-bbox="906 1588 1169 1718">Magen- und Darmmotilität beeinflussende Substanzen</td> <td data-bbox="1169 1588 1401 1718">Cisaprid</td> </tr> <tr> <td data-bbox="906 1718 1169 1789">Herbale / Natürliche Nahrungsergänzung</td> <td data-bbox="1169 1718 1401 1789">Johanniskraut, Echinacea</td> </tr> <tr> <td data-bbox="906 1789 1169 1888">HMG-CoA-Reduktase-Inhibitoren</td> <td data-bbox="1169 1789 1401 1888">Simvastatin, Lovastatin, Cerivastatin</td> </tr> <tr> <td data-bbox="906 1888 1169 1933">Neuroleptika</td> <td data-bbox="1169 1888 1401 1933">Pimozid</td> </tr> <tr> <td data-bbox="906 1933 1169 2000">Sedativa/ Hypnotika</td> <td data-bbox="1169 1933 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 <sup>a</sup>	Charakteristikum	Studieninformation	
		Systemische Kortikosteroide mit Ausnahme vom kurzfristigen Gebrauch ( $\leq 1$ 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"> <li>• Teilnahme in anderer klinischer Studie ohne vorheriges Einverständnis des Sponsors war während der Teilnahme an dieser Studie verboten</li> <li>• 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</li> </ul>	
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><b>EVG/COBI/FTC/TDF-Gruppe:</b> 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><b>EFV/FTC/TDF-Gruppe:</b> 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><b>Allgemeine Objektiv:</b> Primär: Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF versus EFV/FTC/TDF, anhand des Erreichens von HIV-1-RNA &lt;50 Kopien/mL zu Woche 48 Sekundär: Bewertung der Wirksamkeit, Sicherheit und Verträglichkeit der beiden Substanzkombinationen nach 96 Wochen Behandlung</p> <p><b>Primäres Zielkriterium:</b> 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 <sup>a</sup>	Charakteristikum	Studieninformation
		<p>virologischen Respondern (Plasma-Viruslast &lt; 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><b>Sekundäre Zielkriterien:</b></p> <ul style="list-style-type: none"> <li>• Anteil an virologischen Respondern (Erreichen und Erhalt einer Plasma-Viruslast &lt;50 Kopien/mL, gemäß Time-to-loss-of-virologic-response (TLOVR)) nach 48 Wochen</li> <li>• Zeit bis Verlust des virologischen Ansprechens nach 48 Wochen</li> </ul> <p><b>Tertiäre Zielkriterien:</b></p> <ul style="list-style-type: none"> <li>• Pures virologisches Versagen nach 48 Wochen</li> <li>• Veränderung des Plasma HIV-1-RNA von Baseline bis Woche 48</li> <li>• Veränderung der CD4-Zellzahl von Baseline bis Woche 48</li> <li>• Veränderung der CD4-Prozente von Baseline bis Woche 48</li> </ul> <p><b>Resistenzanalyse:</b> Genotypische Resistenzen zu Woche 48</p> <p><b>Pharmakokinetik:</b> Plasmakonzentrationen und pharmakokinetische Parameter von Elvitegravir, Cobicistat, Emtricitabin und Tenofoviridisoproxil</p> <p><b>Sicherheits-Zielkriterien:</b></p> <ul style="list-style-type: none"> <li>• Unerwünschte Ereignisse</li> <li>• Schwerwiegende unerwünschte Ereignisse</li> <li>• Therapieabbrüche aufgrund unerwünschter Ereignisse</li> <li>• Unerwünschte Ereignisse von Interesse</li> <li>• Wichtige unerwünschte Ereignisse</li> <li>• Tode</li> </ul>
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 <sup>a</sup>	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 &lt; 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 &lt;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"> <li>• Erste Interim-Analyse: nachdem 350 Patienten die Woche 12-Visiten beendet haben oder die Studienbehandlung vorzeitig beendet haben</li> <li>• Woche 24 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 24 Visite oder die Studienbehandlung vorzeitig beendet hat</li> <li>• Woche 48 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 48 Visite oder die Studienbehandlung vorzeitig beendet hat</li> <li>• Woche 96 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 96 Visite oder die Studienbehandlung vorzeitig beendet hat</li> <li>• Finale Analyse, wenn der letzte Patient die Studie beendet hat</li> </ul> <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 [75]. Für jede Interim-</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		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 ( $\leq 100.000$ 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	<b>Populationen</b> <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 <sup>a</sup>	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><b>Demografie und Baseline Charakteristika:</b> Darstellung von</p> <ul style="list-style-type: none"> <li>• Stichprobengröße</li> <li>• Mittelwert</li> <li>• Standardabweichung</li> <li>• Median</li> <li>• 1. Quartil</li> <li>• 3. Quartil</li> <li>• Minimum</li> <li>• Maximum</li> </ul> <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><b>Dauer der Medikament Einnahme</b> Allgemein: Deskriptive Beschreibung.</p> <ul style="list-style-type: none"> <li>• Zeit bis zum vorzeitigen Abbruch der Studienbehandlung: Kaplan-Meier Analyse</li> <li>• Unterschiede in Dauer der Medikament Einnahme (stratifiziert nach Baseline HIV-1-RNA Level): Log-Rank Test</li> </ul> <p><b>Primärer Endpunkt</b> Analysefenster für Woche 48: Tag 309 bis inklusive Tag 378</p> <p><u>Snapshot-Algorithmus:</u></p> <ul style="list-style-type: none"> <li>• <i>Virologisches Ansprechen:</i> Patienten deren letzter verfügbarer HIV-1-RNA Wert &lt;50 Kopien/mL im Woche 48-</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Analysefenster während der randomisierten Behandlung lag</p> <ul style="list-style-type: none"> <li>• <i>Virologisches Versagen:</i> Patienten deren letzter verfügbarer HIV-1-RNA Wert <math>\geq 50</math> 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"> <li>○ Behandlungsabbruch aufgrund fehlender Wirksamkeit</li> <li>○ Behandlungsabbruch aus anderen Gründen als unerwünschte Ereignisse, Tod oder fehlende Wirksamkeit und letzter verfügbarer HIV-1-RNA-Wert während Behandlung war <math>\geq 50</math> Kopien/mL</li> </ul> </li> <li>• <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"> <li>○ Studienbehandlung wurde aufgrund unerwünschtes Ereignis oder Tod abgebrochen (unabhängig vom letzten HIV-1-RNA-Wert)</li> <li>○ 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 <math>&lt; 50</math> Kopien/mL</li> <li>○ Fehlende Werte während dem Analysefenster, aber in Studie verbleibend</li> </ul> </li> </ul> <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 95,2%-KI, welches als 95%-KI bezeichnet</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>wird</p> <p>Die Baseline HIV-1-RNA Schicht (<math>\leq 100.000</math> oder <math>&gt; 100.000</math> 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"> <li>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-Wert während Behandlung <math>&lt; 50</math> Kopien/mL war</li> <li>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-Wert während Behandlung <math>&lt; 50</math> Kopien/mL war, wurden als virologische Ansprecher gewertet</li> <li>3. Sensitivitätsanalyse: Suche nach Confoundern: <ol style="list-style-type: none"> <li>a. Stratifizierung nach Region</li> <li>b. Ohne jegliche Stratifizierung</li> </ol> </li> </ol> <p><b>Sekundäre und tertiäre Endpunkte</b></p> <p>Allgemein: gleiche Analysemethoden wie primärer Endpunkt</p> <ul style="list-style-type: none"> <li>• Time to loss of virologic response: Kaplan-Meier und Log Rank Test stratifiziert nach Baseline HIV-1-RNA Level (<math>\leq / &gt; 100.000</math> Kopien/mL)</li> <li>• Veränderung <math>\log_{10}</math> HIV-1-RNA, CD4-Zellzahl und CD4%: deskriptiv bzw. mittels ANOVA mit Baseline HIV-1-RNA Level (<math>\leq / &gt; 100.000</math> Kopien/mL) als fester Effekt</li> <li>• Anteil Patienten mit HIV-1-RNA <math>&lt; 50</math> Kopien/mL mittels M=F und M=E Methoden.</li> </ul>
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"> <li>• Ausgangs-HIV-1-RNA Level</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>(<math>\leq</math>/<math>&gt;</math>100.000 Kopien/mL)</p> <ul style="list-style-type: none"> <li>• Alter (<math>&lt;</math>/<math>\geq</math>40 Jahre)</li> <li>• Geschlecht (männlich/weiblich)</li> <li>• Rasse (weiß/nicht-weiß)</li> <li>• CD4-Zellzahl zu Baseline (<math>\leq</math>/<math>&gt;</math>350 Zellen/<math>\mu</math>L)</li> <li>• Adhärenz zur Studienmedikation (<math>&lt;</math>/<math>\geq</math>95%)</li> </ul> <p>Alle Subgruppenanalysen wurden mittels Snapshot-Algorithmus berechnet und bezüglich HIV-1-RNA Level (<math>\leq</math>/<math>&gt;</math>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=353 EFV/FTC/TDF-Gruppe: n=354 b) EVG/COBI/FTC/TDF-Gruppe: n=348 EFV/FTC/TDF-Gruppe: n=352 c) EVG/COBI/FTC/TDF-Gruppe: n=348 EFV/FTC/TDF-Gruppe: n=352
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	Screening des ersten Patienten: 16. März 2010 Letzte Beobachtung des letzten Patienten: 10. August 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.

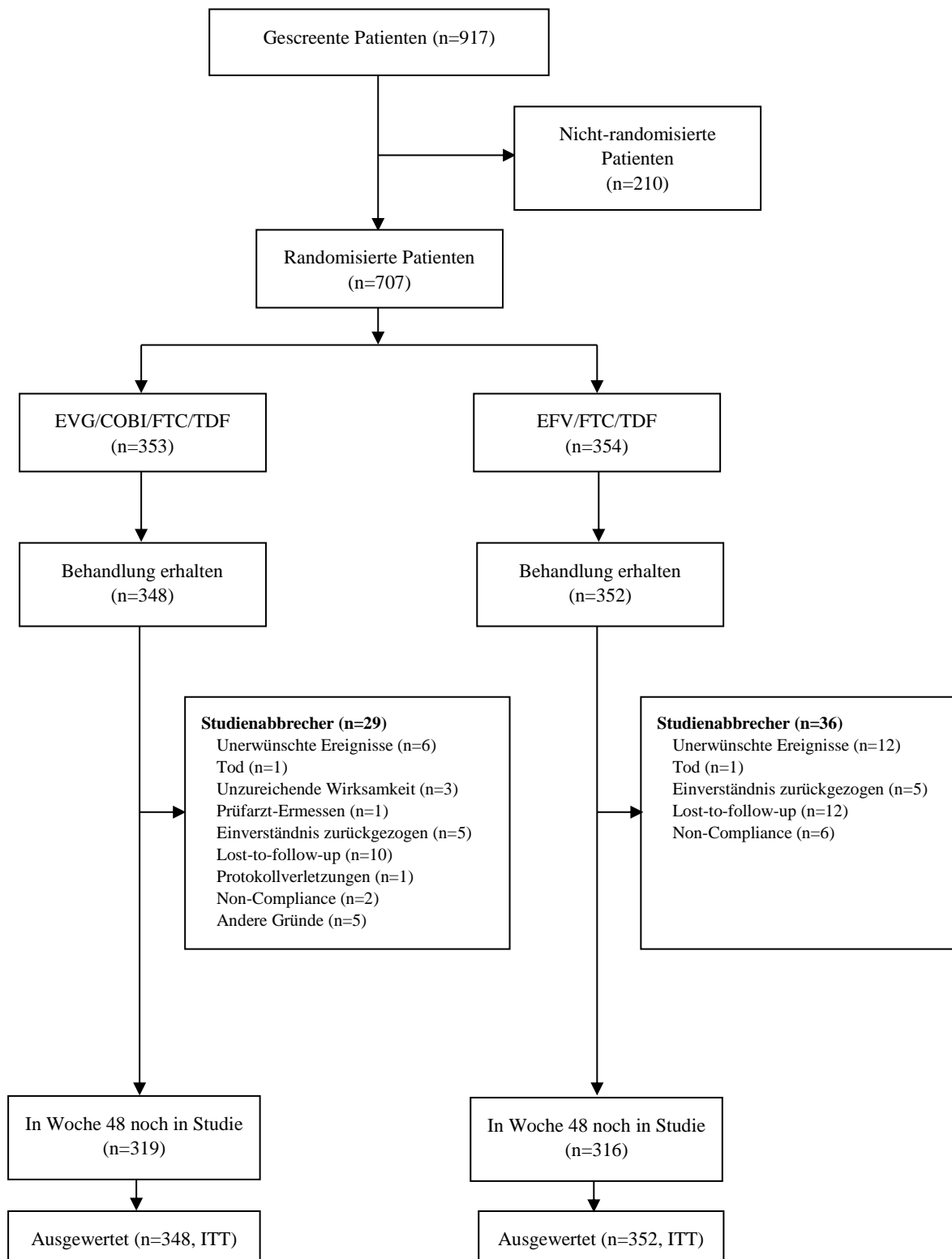


Abbildung 53: Flow-Chart für Studie GS-US-236-0102



Tabelle 4-144 (Anhang): Studiendesign und -methodik für Studie GS-US-236-0104

Item <sup>a</sup>	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Untersuchung der Wirksamkeit und Sicherheit des 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	<b>Einschlusskriterien:</b> <ul style="list-style-type: none"> <li>• Fähig eine schriftliche Einwilligungserklärung zu verstehen und zu unterschreiben, welche vor dem Start der Studienprozeduren eingeholt werden musste</li> <li>• Plasma HIV-1-RNA Level <math>\geq 5000</math> Kopien/mL</li> <li>• Keine vorherige Einnahme eines zugelassenen oder experimentellen HIV-Medikaments</li> <li>• Normales Elektrokardiogramm (EKG); oder wenn unnormal, musste der Prüfarzt bestätigen dass es nicht klinisch relevant ist</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• Adäquate Nierenfunktion: eGFR <math>\geq 80</math> mL/min gemäß Cockcroft-Gault (CG) Formel: Männer: <math>eGFR \text{ (mL/min)} = \frac{(140 - \text{Alter in Jahren}) \cdot (\text{Gewicht in Kilo})}{72 \cdot (\text{Serum Kreatinin in mg/dL})}</math> Frauen: <math>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</math></li> <li>• Lebertransaminasen (Aspartat-Aminotransferase (AST) und Alanin-Aminotransferase (ALT) <math>\leq 2,5</math> x Obergrenze des Normbereichs (upper limit of normal, ULN)</li> <li>• Gesamt-Bilirubin <math>\leq 1,5</math> mg/dL oder normales direktes Bilirubin</li> <li>• Adäquate hämatologische Funktion (absolute Zahl neutrophiler Granulozyten <math>\geq 1000/\text{mm}^3</math>; Thrombozyten <math>\geq 50.000/\text{mm}^3</math>; Hämoglobin <math>\geq 8,5</math> g/dL)</li> <li>• Zahl CD4-Zellen <math>&gt; 50</math> Zellen/<math>\mu\text{L}</math></li> <li>• Serum-Amylase <math>\leq 1,5</math> x ULN (Patienten mit Serum-Amylase <math>&gt; 1,5</math> x ULN blieben geeignet falls Serumlipase <math>\leq 1,5</math> x ULN war)</li> <li>• Normales Thyroid-stimulierende Hormon</li> <li>• Für gebärfähige Frauen: negativer Serum-Schwangerschaftstest</li> <li>• Männer und gebärfähige Frauen (nicht-menopausale Frauen oder Frauen, die <math>&lt; 2</math> Jahre in der Menopause sind und follikelstimulierendes Hormon <math>&lt; 40\text{mIU/mL}</math> aufweisen, und die keine Hysterektomie, bilaterale Ovariectomie 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 oder vasktomierter Partner) ab dem Screening für die Dauer der Studienbehandlung und für 12 Wochen nach der letzten Dosis der Studienbehandlung anzuwenden.</li> <li>• Weibliche Patienten, die hormonelle Verhütungsmethoden nutzten, müssen diese Methode für mindestens 3 Monate vor Studienbehandlung verwendet haben</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• Bei postmenopausalen Frauen musste ein Serum-Follikel-stimulierendes Hormonlevel (FSH) &gt; 40mIU/mL vorliegen. Falls das FSH ≤ 40mIU/mL lag, musste die Patientin einer Verwendung einer hochwirksamen Verhütungsmethode zustimmen.</li> <li>• Alter ≥18 Jahre</li> <li>• Lebenserwartung ≥1 Jahr</li> </ul> <p><b>Ausschlusskriterien:</b></p> <ul style="list-style-type: none"> <li>• Diagnose einer neuen Aids-definierenden Erkrankung innerhalb von 30 Tagen vor dem Screening</li> <li>• 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</li> <li>• Hepatitis B Oberflächenantigen (HBsAg)- positiv</li> <li>• Hepatitis C Virus-Antikörper (HCVAb) – positiv</li> <li>• Vorhandensein einer Zirrhose, Aszites oder Enzephalopathie</li> <li>• Stillende Frauen</li> <li>• Positiver Serum-Schwangerschaftstest (gebärfähige Frauen)</li> <li>• Impfung innerhalb 90 Tage der Studiendosierung</li> <li>• 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 &lt; 30 Jahre</li> <li>• Vorliegen von oder Vorgeschichte bezüglich kardiovaskulärer Erkrankungen, Kardiomyopathie und/oder kardialer Leitungsstörungen</li> <li>• Ausgedehntes, für die Herzfrequenz nach Fridericia's Formel korrigiertes QT-Intervall (QTcF) beim Screening (QTcF &gt; 450ms bei Männern, QtcF &gt;470ms bei Frauen)</li> <li>• PR-Intervall ≥ 200ms oder ≤ 120ms im EKG bei der Screening-Visite</li> <li>• Implantierter Defibrillator oder</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation																										
		<p>Schrittmacher</p> <ul style="list-style-type: none"> <li>• 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.</li> </ul>																										
		<table border="1"> <thead> <tr> <th data-bbox="914 577 1169 645">Medikamentenklasse</th> <th data-bbox="1169 577 1393 645">Verbotene Wirkstoffe*</th> </tr> </thead> <tbody> <tr> <td data-bbox="914 645 1169 712">Alpha-adrenerge Rezeptor-Antagonisten</td> <td data-bbox="1169 645 1393 712">Alfuzosin</td> </tr> <tr> <td data-bbox="914 712 1169 902">Antiarrhythmika</td> <td data-bbox="1169 712 1393 902">Amiodaron, Flecainid, Quinidin, Propafenon, systemisches Lidocain, Mexilitin, Disopyramid</td> </tr> <tr> <td data-bbox="914 902 1169 947">Antibiotika</td> <td data-bbox="1169 902 1393 947">Telithromycin</td> </tr> <tr> <td data-bbox="914 947 1169 1037">Antikonvulsiva</td> <td data-bbox="1169 947 1393 1037">Phenobarbital, Phenytoin, Carbamazepin</td> </tr> <tr> <td data-bbox="914 1037 1169 1126">Antimykotika</td> <td data-bbox="1169 1037 1393 1126">Voriconazol, Ketoconazol, Itraconazol</td> </tr> <tr> <td data-bbox="914 1126 1169 1216">Antihistaminika</td> <td data-bbox="1169 1126 1393 1216">Astemizol, Terfenadin</td> </tr> <tr> <td data-bbox="914 1216 1169 1305">Mittel gegen Mykobakterien</td> <td data-bbox="1169 1216 1393 1305">Rifampin, Rifapentin, Rifabutin</td> </tr> <tr> <td data-bbox="914 1305 1169 1496">Kalziumkanalblocker</td> <td data-bbox="1169 1305 1393 1496">Felodipin, Bepridil, Nifedipin, Nicardipin, Verapamil, Diltiazem, Amlodipin</td> </tr> <tr> <td data-bbox="914 1496 1169 1675">Ergotaminderivaten</td> <td data-bbox="1169 1496 1393 1675">Ergotamin, Ergonovin, Dihydroergotamin, Methylegonovin, Ergometrin</td> </tr> <tr> <td data-bbox="914 1675 1169 1798">Magen- und Darmmotilität beeinflussende Substanzen</td> <td data-bbox="1169 1675 1393 1798">Cisaprid</td> </tr> <tr> <td data-bbox="914 1798 1169 1877">Pflanzliche Nahrungsergänzung</td> <td data-bbox="1169 1798 1393 1877">Johanniskraut, Echinacea</td> </tr> <tr> <td data-bbox="914 1877 1169 1966">HMG-CoA-Reduktase-Inhibitoren</td> <td data-bbox="1169 1877 1393 1966">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 <sup>a</sup>	Charakteristikum	Studieninformation	
		Neuroleptika	Pimozid
		Sedativa/ Hypnotika	Midazolam, Triazolam
		Systemische Kortikosteroide	Alle Mittel, einschließlich Dexamethason
		Mittel zur systemischen antineoplastischen Chemotherapie	Alle Mittel
		<p>* Einnahme jeglicher der genannten Medikamente wurde mindestens 30 Tage vor der Baseline Visite und für die Dauer der Studie abgebrochen</p> <ul style="list-style-type: none"> <li>• Aktueller Alkohol- oder Substanzgebrauch, welcher nach Meinung des Prüfarztes die Studien-Compliance des Patienten potenziell beeinträchtigt</li> <li>• 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-bestätigtem 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</li> <li>• Aktive, ernste Infektionen (anders als HIV-1-Infektion), welche parenterale antibiotische oder antimykotische Therapie innerhalb 30 Tage vor Baseline erfordern</li> <li>• Teilnahme in einer anderen klinischen Studie ohne vorheriges Einverständnis des Sponsors war während der Teilnahme an dieser Studie verboten</li> <li>• Kontraindizierte Medikamente hinsichtlich FTC, EFV oder TDF</li> <li>• Bekannte Allergien gegen sonstige Bestandteile der EVG/COBI/FTC/TDF- oder EFV/FTC/TDF-Tabletten</li> <li>• 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</li> </ul>	
4b	Studienorganisation und Ort der Studiendurchführung	30 Zentren in den USA	

Item <sup>a</sup>	Charakteristikum	Studieninformation
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p><b>EVG/COBI/FTC/TDF-Gruppe:</b></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><b>EFV/FTC/TDF-Gruppe:</b></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><b>Allgemeine Objektivien:</b></p> <p>Primär: Bewertung der Wirksamkeit von EVG/COBI/FTC/TDF versus EFV/FTC/TDF, anhand des Erreichens von HIV-1-RNA &lt;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 &lt;50 Kopien/mL zu Woche 48</p> <p><b>Primäres Zielkriterium:</b></p> <p>Vergleich des Anteils der virologischen Ansprecher (Plasma-Viruslast &lt;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><b>Sekundäre Zielkriterien:</b></p> <ul style="list-style-type: none"> <li>• % Patienten mit virologischem Ansprechen (Plasma-Viruslast &lt;50 Kopien/mL gemäß COBAS Amplicor HIV-1 Monitor Test) zu Woche 48 mit Missing=Failure, Missing=Excluded und Missing/Switch=Failure Methoden</li> <li>• Virologisches Ansprechen (Plasma-Viruslast &lt;50 Kopien/mL) zu Woche 24 und 48 gemäß FDA-definierter Snapshot Analyse</li> <li>• Veränderung der HIV-1-RNA (log<sub>10</sub> Kopien/mL) zwischen Baseline und Woche 24 bzw. 48</li> <li>• Veränderung von Anzahl und Prozent der</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>CD4-Zellzahl zwischen Baseline und Woche 24 bzw. 48</p> <p><b>Pharmakokinetik:</b> Pharmakokinetische Blutprobe, die 20-24 Stunden nach einer beobachteten Dosis EVG/COBI/FTC/TDF (bzw. Placebo) genommen wurde</p> <p><b>Sicherheits-Zielkriterien:</b></p> <ul style="list-style-type: none"> <li>• Unerwünschte Ereignisse</li> <li>• Begleitmedikation</li> <li>• Klinische Labortests</li> <li>• Rückenlägriges 12-Kanal-EKG</li> <li>• Größe, Gewicht</li> </ul>
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 &lt;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 &lt;50 Kopien/mL zu Woche 24 zu zeigen, falls eine Responserate 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"> <li>• Woche 16 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die 16. Studienwoche beendet hat</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• Woche 16 Zwischen-Pharmakokinetik-Analyse</li> <li>• Woche 24 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 24 Visite oder die Studienbehandlung vorzeitig beendet hat</li> <li>• Woche 48 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 48 Visite oder die Studienbehandlung vorzeitig beendet hat</li> <li>• Woche 48+ Interim-Analyse: nachdem der letzte Patient die Studie beendet oder die Studienbehandlung vorzeitig beendet hat, inklusive der 30 Tage Follow-up-Visite</li> <li>• Woche 96 Independent Data Monitoring Committee (IDMC) Interim-Analyse: nachdem der letzte Patient die Woche 96 Visite oder die Studienbehandlung vorzeitig beendet hat</li> <li>• Finale Analyse, wenn der letzte Patient die Studie beendet hat</li> </ul> <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 [75]. Für jede Zwischenanalyse 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 ( $\leq 100.000$ 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 <sup>a</sup>	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><b>Populationen</b></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 <sup>a</sup>	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><b>Demografie und Baseline Charakteristika:</b>  Darstellung von</p> <ul style="list-style-type: none"> <li>• Stichprobengröße</li> <li>• Mittelwert</li> <li>• Standardabweichung</li> <li>• Median</li> <li>• 1. Quartil</li> <li>• 3. Quartil</li> <li>• Minimum</li> <li>• Maximum</li> </ul> <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><b>Dauer der Medikamenten-Einnahme</b>  Allgemein: Deskriptive Beschreibung.</p> <ul style="list-style-type: none"> <li>• Zeit bis zum vorzeitigen Abbruch der Studienbehandlung: Kaplan-Meier Analyse</li> <li>• Unterschiede in Dauer der Medikamenten-Einnahme (stratifiziert)</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>nach Baseline HIV-1-RNA Level): Log-Rank Test</p> <p><b>Primärer Endpunkt</b>  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 (<math>\leq 100.000</math> oder <math>&gt; 100.000</math> Kopien/mL)-gewichtete MH-Anteile und Normalverteilungsapproximation verwendet.</p> <p><b>Sekundäre und tertiäre Endpunkte</b>  Allgemein: gleiche Analysemethoden wie primärer Endpunkt</p> <ul style="list-style-type: none"> <li>• HIV-1-RNA <math>&lt; 50</math> Kopien/mL wurde zu Woche 24 und 48 gemäß FDA-Snapshot-Methode evaluiert</li> <li>• Veränderung <math>\log_{10}</math> HIV-1-RNA, CD4-Zellzahl und CD4%: deskriptiv bzw. mittels ANOVA mit Baseline HIV-1-RNA Level (<math>\leq / &gt; 100.000</math> Kopien/mL) als fester Effekt</li> <li>• Jeweils für Missing=Failure, Missing=Excluded und Missing/Switch=Failure-Methoden</li> </ul>
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 <sup>a</sup>	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.*

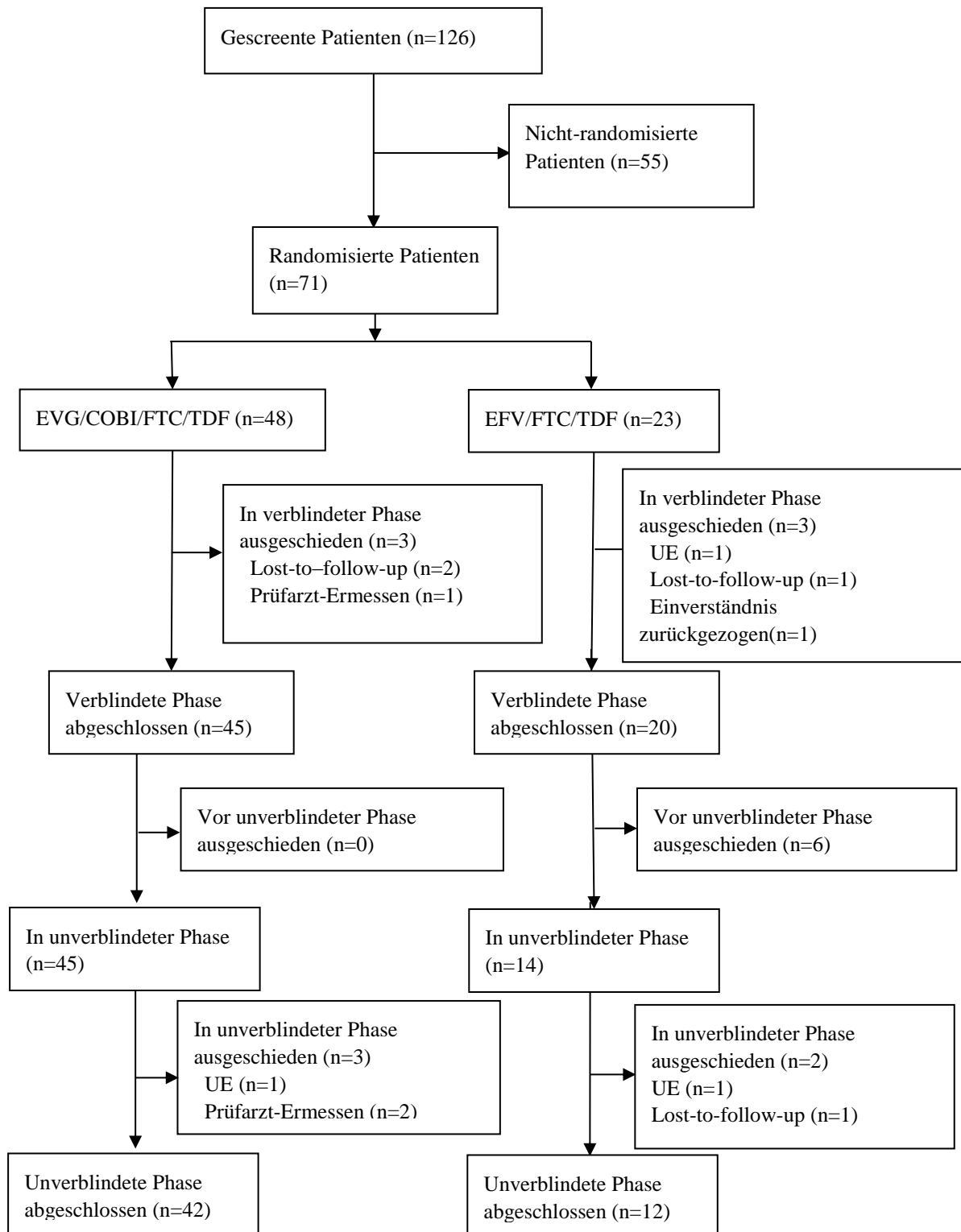


Abbildung 54: Flow-Chart für Studie GS-US-236-0104

**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 4-145 (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: 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 endpunkübergreifende 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 endpunkübergreifenden 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 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 4-146 (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: 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:

---

Der Endpunkt Zeit bis zum Verlust des virologischen Ansprechens 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 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

---

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 zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full 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:

---

---