Dokumentvorlage, Version vom 18.04.2013

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Dolutegravir/Rilpivirin (Juluca)

ViiV Healthcare GmbH

Modul 4 A

Antiretrovirale Kombinationstherapie von Erwachsenen, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert und virologisch supprimiert sind

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 22.05.2018

Inhaltsverzeichnis

Seite

Tabellenverzeichnis **5**
Abbildungsverzeichnis **14**
Abkürzungsverzeichnis **16**
4 Modul 4 – allgemeine Informationen **18**
4.1 Zusammenfassung der Inhalte von Modul 4 19
4.2 Methodik 36
4.2.1 Fragestellung 36
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung 39
4.2.3 Informationsbeschaffung 42
4.2.3.1 Studien des pharmazeutischen Unternehmers 42
4.2.3.2 Bibliografische Literaturrecherche 42
4.2.3.3 Suche in Studienregistern 43
4.2.3.4 Selektion relevanter Studien 44
4.2.4 Bewertung der Aussagekraft der Nachweise 45
4.2.5 Informationssynthese und -analyse 46
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien 46
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien 47
4.2.5.3 Meta-Analysen 52
4.2.5.4 Sensitivitätsanalysen 54
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren 55
4.2.5.6 Indirekte Vergleiche 57
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen 59
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel 59
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel 59
4.3.1.1.1 Studien des pharmazeutischen Unternehmers 59
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche 61
4.3.1.1.3 Studien aus der Suche in Studienregistern 64
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel 65
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 66
4.3.1.2.1 Studiendesign und Studienpopulationen 66
4.3.1.2.2 Verzerrungspotenzial auf Studienebene 73
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien 74
4.3.1.3.1 Viruslast < 50 Kopien/ml zu Woche 48 – RCT 76
4.3.1.3.2 EQ-5D zu Woche 48 – RCT 86
4.3.1.3.3 Symptom Distress Module zu Woche 48 – RCT 93
4.3.1.3.4 HIVTSQ Total Score zu Woche 48 – RCT 100
4.3.1.3.5 Mortalität bis Woche 48 – RCT 105
4.3.1.3.6 Unerwünschte Ereignisse bis Woche 48 – RCT 108

4.3.1.3.7 Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 – RCT 112
4.3.1.3.8 Schwerwiegende unerwünschte Ereignisse bis Woche 48 – RCT 117
4.3.1.3.9 Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten – RCT 122
4.3.1.3.10 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen – RCT 126
4.3.1.3.11 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts – RCT 131
4.3.1.3.12 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen – RCT 136
4.3.1.3.13 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität – RCT 140
4.3.1.3.14 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen – RCT 143
4.3.1.3.15 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen – RCT 147
4.3.1.3.16 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems – RCT 152
4.3.1.3.17 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen – RCT 157
4.3.1.3.18 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege – RCT 161
4.3.1.3.19 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes – RCT 165
4.3.1.3.20 Knochenmineraldichte zu Woche 48 – RCT 169
4.3.1.3.21 Knochenspezifischer Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 – RCT 174
4.3.1.3.22 Knochenspezifischer Biomarker: Prokollagen 1-N-terminales Propeptid (P1NP) zu Woche 48 – RCT 178
4.3.1.3.23 Knochenspezifischer Biomarker: type 1 collagen cross-linked C-telopeptide (CTx) zu Woche 48 – RCT 182
4.3.1.3.24 Knochenspezifischer Biomarker: Osteocalcin zu Woche 48 – RCT 186
4.3.1.3.25 Subgruppenanalysen – RCT 190
4.3.1.3.25.1 Subgruppenanalysen: Viruslast < 50 Kopien/ml zu Woche 48 - RCT 195
4.3.1.3.25.2 Subgruppenanalysen: EQ-5D zu Woche 48 - RCT 196
4.3.1.3.25.3 Subgruppenanalysen: Symptom Distress Module zu Woche 48 - RCT 201
4.3.1.3.25.4 Subgruppenanalysen: HIVTSQ Total Score zu Woche 48 - RCT 203
4.3.1.3.25.5 Subgruppenanalysen: unerwünschte Ereignisse bis Woche 48 - RCT 204
4.3.1.3.25.6 Subgruppenanalysen: unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 - RCT 205
4.3.1.3.25.7 Subgruppenanalysen: schwerwiegende unerwünschte Ereignisse bis Woche 48 - RCT 207
4.3.1.3.25.8 Subgruppenanalysen: Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten - RCT 208

4.3.1.3.25.9 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Herzerkrankungen bis Woche 48 - RCT 209
4.3.1.3.25.10 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Erkrankungen des Gastrointestinaltrakts bis Woche 48 - RCT 210
4.3.1.3.25.11 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Leber- und Gallenerkrankungen bis Woche 48 - RCT 211
4.3.1.3.25.12 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Hypersensitivität Woche 48 - RCT 212
4.3.1.3.25.13 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Stoffwechsel- und Ernährungsstörungen bis Woche 48 - RCT 213
4.3.1.3.25.14 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen bis Woche 48 - RCT 214
4.3.1.3.25.15 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Erkrankungen des Nervensystems bis Woche 48 - RCT 215
4.3.1.3.25.16 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: psychiatrische Erkrankungen bis Woche 48 - RCT 216
4.3.1.3.25.17 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Erkrankungen der Niere und Harnwege bis Woche 48 - RCT 218
4.3.1.3.25.18 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 - RCT 219
4.3.1.3.25.19 Subgruppenanalysen: alkalische Phosphatase (BSAP) zu Woche 48 - RCT 224
4.3.1.3.25.20 Subgruppenanalysen: Prokollagen 1-N-terminales Propeptid (P1NP) zu Woche 48 - RCT 225
4.3.1.3.25.21 Subgruppenanalysen: type 1 collagen cross-linked C-telopeptide (CTx) zu Woche 48 - RCT 227
4.3.1.3.25.22 Subgruppenanalysen: Osteocalcin zu Woche 48 - RCT 237
4.3.1.3.26 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien 244
4.3.2 Weitere Unterlagen 248
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien 248
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche 248
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche 248
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen 249
4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT 249
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT 251
4.3.2.2 Nicht randomisierte vergleichende Studien 252
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien 252

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien 252
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien 253
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien..... 253
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien 254
4.3.2.3 Weitere Untersuchungen 255
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen 255
4.3.2.3.2 Charakteristika der weiteren Untersuchungen 255
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen 256
4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen 256
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen 256
4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen 257
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens 258
4.4.1 Beurteilung der Aussagekraft der Nachweise 258
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß 259
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht 270
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte 270
4.5.1 Begründung für die Vorlage indirekter Vergleiche 270
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen 271
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen 271
4.5.4 Verwendung von Surrogatendpunkten 271
4.6 Liste der eingeschlossenen Studien 273
4.7 Referenzliste 274
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche 278
Anhang 4-B : Suchstrategien – Suche in Studienregistern 282
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche) 284
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern) 286
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT 291
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 323

Tabellenverzeichnis

Seite

Tabelle 4-1: Studienpool - RCT mit zu bewertenden Arzneimittel 21
Tabelle 4-2: Ein- und Ausschlusskriterien bei der Studienselektion 23
Tabelle 4-3: Übersicht der Ergebnisse pro Endpunkt – Teilpopulation ohne Umstellindikation 25
Tabelle 4-4: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens 35
Tabelle 4-5: Ein- und Ausschlusskriterien bei der Studienselektion 40
Tabelle 4-6: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel 60
Tabelle 4-7: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel 60
Tabelle 4-8: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel 64
Tabelle 4-9: Studienpool – RCT mit dem zu bewertenden Arzneimittel 65
Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation) 67
Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation) 69
Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation) [ITT-E Population] 70
Tabelle 4-13: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation) 73
Tabelle 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel - Teilpopulation ohne Umstellindikation 74
Tabelle 4-15: Operationalisierung von Viruslast < 50 Kopien/ml zu Woche 48 77
Tabelle 4-16: Bewertung des Verzerrungspotenzials für Viruslast < 50 Kopien/ml zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel 79
Tabelle 4-17: Ergebnisse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - ITT-E Population 80
Tabelle 4-18: Meta-Analyse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel 81
Tabelle 4-19: Sensitivitätsanalyse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Sensitivitätsanalyse 1: Patienten ohne Angabe zur Viruslast ausgeschlossen 83
Tabelle 4-20: Sensitivitätsanalyse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Sensitivitätsanalyse 2: Imputation mittels Responderrate ohne fehlender Werte 84

Tabelle 4-21: Operationalisierung von EQ-5D zu Woche 48 ... 86
Tabelle 4-22: Bewertung des Verzerrungspotenzials für EQ-5D zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel ... 87
Tabelle 4-23: Ergebnisse für EQ-5D zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – ITT-E Population ... 88
Tabelle 4-24: Meta-Analyse für EQ-5D zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel ... 90
Tabelle 4-25: Operationalisierung von SDM zu Woche 48 ... 93
Tabelle 4-26: Bewertung des Verzerrungspotenzials für SDM zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel ... 94
Tabelle 4-27: Ergebnisse für SDM zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – ITT-E Population ... 95
Tabelle 4-28: Meta-Analyse für SDM zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel ... 97
Tabelle 4-29: Operationalisierung von HIVTSQ Total Score zu Woche 48 ... 100
Tabelle 4-30: Bewertung des Verzerrungspotenzials für HIVTSQ Total Score zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel ... 101
Tabelle 4-31: Ergebnisse für HIVTSQ Total Score zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – ITT-E Population ... 102
Tabelle 4-32: Meta-Analyse für HIVTSQ Total Score zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel ... 103
Tabelle 4-33: Operationalisierung von Mortalität bis Woche 48 ... 105
Tabelle 4-34: Bewertung des Verzerrungspotenzials für Mortalität bis Woche 48 in RCT mit dem zu bewertenden Arzneimittel ... 105
Tabelle 4-35: Ergebnisse für Mortalität bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population ... 106
Tabelle 4-36: Operationalisierung von unerwünschte Ereignisse bis Woche 48 ... 108
Tabelle 4-37: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse bis Woche 48 in RCT mit dem zu bewertenden Arzneimittel ... 108
Tabelle 4-38: Ergebnisse für unerwünschte Ereignisse bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population ... 109
Tabelle 4-39: Meta-Analyse für unerwünschte Ereignisse bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel ... 110
Tabelle 4-40: Operationalisierung von unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 ... 112
Tabelle 4-41: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 in RCT mit dem zu bewertenden Arzneimittel ... 112
Tabelle 4-42: Ergebnisse für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population ... 113

Tabelle 4-43: Meta-Analyse für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel 114

Tabelle 4-44: Operationalisierung von schwerwiegende unerwünschte Ereignisse bis Woche 48..... 117

Tabelle 4-45: Bewertung des Verzerrungspotenzials für schwerwiegende unerwünschte Ereignisse bis Woche 48 in RCT mit dem zu bewertenden Arzneimittel..... 118

Tabelle 4-46: Ergebnisse für schwerwiegende unerwünschte Ereignisse bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population..... 119

Tabelle 4-47: Meta-Analyse für schwerwiegende unerwünschte Ereignisse bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel 120

Tabelle 4-48: Operationalisierung von Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten 122

Tabelle 4-49: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten in RCT mit dem zu bewertenden Arzneimittel..... 122

Tabelle 4-50: Ergebnisse für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten aus RCT mit dem zu bewertenden Arzneimittel – Safety Population..... 123

Tabelle 4-51: Meta-Analyse für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten aus RCT mit dem zu bewertenden Arzneimittel..... 124

Tabelle 4-52: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen..... 126

Tabelle 4-53: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen in RCT mit dem zu bewertenden Arzneimittel..... 127

Tabelle 4-54: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population..... 128

Tabelle 4-55: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen aus RCT mit dem zu bewertenden Arzneimittel 129

Tabelle 4-56: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts 131

Tabelle 4-57: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts in RCT mit dem zu bewertenden Arzneimittel..... 131

Tabelle 4-58: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 132

Tabelle 4-59: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts aus RCT mit dem zu bewertenden Arzneimittel..... 134

Tabelle 4-60: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen 136

Tabelle 4-61: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen in RCT mit dem zu bewertenden Arzneimittel 136

Tabelle 4-62: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 137

Tabelle 4-63: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen aus RCT mit dem zu bewertenden Arzneimittel 138

Tabelle 4-64: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität 140

Tabelle 4-65: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität in RCT mit dem zu bewertenden Arzneimittel 141

Tabelle 4-66: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 142

Tabelle 4-67: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen 143

Tabelle 4-68: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen in RCT mit dem zu bewertenden Arzneimittel 143

Tabelle 4-69: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 144

Tabelle 4-70: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen aus RCT mit dem zu bewertenden Arzneimittel 145

Tabelle 4-71: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen 147

Tabelle 4-72: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen in RCT mit dem zu bewertenden Arzneimittel 147

Tabelle 4-73: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 148

Tabelle 4-74: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus RCT mit dem zu bewertenden Arzneimittel 150

Tabelle 4-75: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems 152

Tabelle 4-76: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems in RCT mit dem zu bewertenden Arzneimittel 152

Tabelle 4-77: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 153

Tabelle 4-78: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems aus RCT mit dem zu bewertenden Arzneimittel 155

Tabelle 4-79: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen 157

Tabelle 4-80: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen in RCT mit dem zu bewertenden Arzneimittel 157

Tabelle 4-81: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 158

Tabelle 4-82: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen aus RCT mit dem zu bewertenden Arzneimittel .. 159

Tabelle 4-83: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege 161

Tabelle 4-84: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege in RCT mit dem zu bewertenden Arzneimittel 161

Tabelle 4-85: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 162

Tabelle 4-86: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege aus RCT mit dem zu bewertenden Arzneimittel 163

Tabelle 4-87: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes 165

Tabelle 4-88: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes in RCT mit dem zu bewertenden Arzneimittel 165

Tabelle 4-89: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 166

Tabelle 4-90: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes aus RCT mit dem zu bewertenden Arzneimittel 167

Tabelle 4-91: Operationalisierung von Knochenmineraldichte zu Woche 48 169

Tabelle 4-92: Bewertung des Verzerrungspotenzials für Knochenmineraldichte zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel 170

Tabelle 4-93: Patientencharakteristika der DEXA Substudie 170

Tabelle 4-94: Ergebnisse für Knochenmineraldichte (g/cm^2) zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – DEXA Substudie 172

Tabelle 4-95: Operationalisierung von knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 174

Tabelle 4-96: Bewertung des Verzerrungspotenzials für knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel 174

Tabelle 4-97: Ergebnisse für knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 175

Tabelle 4-98: Meta-Analyse für knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel 176

Tabelle 4-99: Operationalisierung von Knochenspezifischen Biomarker: P1NP zu Woche 48 178

Tabelle 4-100: Bewertung des Verzerrungspotenzials für Knochenspezifischen Biomarker: P1NP zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel 178

Tabelle 4-101: Ergebnisse für Knochenspezifischen Biomarker: P1NP zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 179

Tabelle 4-102: Meta-Analyse für Knochenspezifischen Biomarker: P1NP zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel 180

Tabelle 4-103: Operationalisierung von Knochenspezifischen Biomarker: CTx zu Woche 48 182

Tabelle 4-104: Bewertung des Verzerrungspotenzials für Knochenspezifischen Biomarker: CTx zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel 182

Tabelle 4-105: Ergebnisse für Knochenspezifischen Biomarker: CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 183

Tabelle 4-106: Meta-Analyse für Knochenspezifischen Biomarker: CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel 184

Tabelle 4-107: Operationalisierung von Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 186

Tabelle 4-108: Bewertung des Verzerrungspotenzials für Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel 186

Tabelle 4-109: Ergebnisse für Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population 187

Tabelle 4-110: Meta-Analyse für Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel 188

Tabelle 4-111: p-Werte der Interaktionstests pro Endpunkt auf Einzelstudienebene 192

Tabelle 4-112: p-Werte der meta-analytischen Interaktionstests für Viruslast < 50 Kopien/ml zu Woche 48 195

Tabelle 4-113: p-Werte der meta-analytischen Interaktionstests für EQ-5D zu Woche 48... 196

Tabelle 4-114: Ergebnisse für EQ-5D VAS zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline CD4+ Zellzahl..... 197

Tabelle 4-115: Meta-Analyse für EQ-5D VAS zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Baseline CD4+ Zellzahl 198

Tabelle 4-116: Ergebnisse für SDM Symptom Count Score zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Geschlecht..... 201

Tabelle 4-117: p-Werte der meta-analytischen Interaktionstests für Symptom Distress Module zu Woche 48 202

Tabelle 4-118: p-Werte der meta-analytischen Interaktionstests für HIVTSQ Total Score zu Woche 48..... 203

Tabelle 4-119: p-Werte der meta-analytischen Interaktionstests für unerwünschte Ereignisse bis Woche 48 204

Tabelle 4-120: Ergebnisse für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline CD4+ Zellzahl..... 205

Tabelle 4-121: p-Werte der meta-analytischen Interaktionstests für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 206

Tabelle 4-122: p-Werte der meta-analytischen Interaktionstests für schwerwiegende unerwünschte Ereignisse bis Woche 48 207

Tabelle 4-123: p-Werte der meta-analytischen Interaktionstests für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten 208

Tabelle 4-124: p-Werte der meta-analytischen Interaktionstests für Erkrankungen des Gastrointestinaltrakts bis Woche 48..... 210

Tabelle 4-125: p-Werte der meta-analytischen Interaktionstests für Stoffwechsel- und Ernährungsstörungen bis Woche 48..... 213

Tabelle 4-126: p-Werte der meta-analytischen Interaktionstests für Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen bis Woche 48..... 214

Tabelle 4-127: p-Werte der meta-analytischen Interaktionstests für Erkrankungen des Nervensystems bis Woche 48 215

Tabelle 4-128: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse: psychiatrische Erkrankungen bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Alter 216

Tabelle 4-129: p-Werte der meta-analytischen Interaktionstests für psychiatrische Erkrankungen bis Woche 48 217

Tabelle 4-130: p-Werte der meta-analytischen Interaktionstests für Erkrankungen der Niere und Harnwege bis Woche 48..... 218

Tabelle 4-131: p-Werte der meta-analytischen Interaktionstests für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 219

Tabelle 4-132: Ergebnisse für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline CD4+ Zellzahl ... 220

Tabelle 4-133: Meta-Analyse für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Baseline CD4+ Zellzahl ... 221

Tabelle 4-134: p-Werte der meta-analytischen Interaktionstests für BSAP zu Woche 48 ... 224

Tabelle 4-135: Ergebnisse P1NP zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Alter ... 225

Tabelle 4-136: p-Werte der meta-analytischen Interaktionstests für P1NP zu Woche 48 ... 226

Tabelle 4-137: Ergebnisse CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline Third Agent ... 227

Tabelle 4-138: p-Werte der meta-analytischen Interaktionstests für CTx zu Woche 48 ... 228

Tabelle 4-139: Ergebnisse für CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Geschlecht ... 229

Tabelle 4-140: Meta-Analyse für CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Geschlecht ... 230

Tabelle 4-141: Ergebnisse für CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline CD4+ Zellzahl ... 233

Tabelle 4-142: Meta-Analyse für CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Baseline CD4+ Zellzahl ... 234

Tabelle 4-143: Ergebnisse Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Alter ... 237

Tabelle 4-144: p-Werte der meta-analytischen Interaktionstests für Osteocalcin zu Woche 48 ... 238

Tabelle 4-145: Ergebnisse für Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline Third Agent ... 239

Tabelle 4-146: Meta-Analyse für Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Baseline Third Agent ... 240

Tabelle 4-147: Übersicht der Ergebnisse aus RCT für die Teilpopulation ohne Umstellindikation auf Endpunktebene ... 245

Tabelle 4-148: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche ... 249

Tabelle 4-149: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden ... 249

Tabelle 4-150: Operationalisierung von <Endpunkt xxx> ... 250

Tabelle 4-151: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche ... 250

Tabelle 4-152: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche ... 251

Tabelle 4-153: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien ... 253

Tabelle 4-154: Operationalisierung von <Endpunkt xxx> ... 253
Tabelle 4-155: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien ... 254
Tabelle 4-156: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen ... 256
Tabelle 4-157: Übersicht der Ergebnisse pro Endpunkt – Teilpopulation ohne Umstellindikation ... 260
Tabelle 4-158: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens ... 270
Tabelle 4-159: Studienpool – Studien mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation) ... 273
Tabelle 4-160 (Anhang): Studiendesign und -methodik für Studie 201636 (SWORD-1) ... 291
Tabelle 4-161 (Anhang): Studiendesign und -methodik für Studie 201637 (SWORD-2) ... 303
Tabelle 4-162 (Anhang): Studiendesign und -methodik für Studie 202094 (DEXA-Substudie) ... 315
Tabelle 4-163 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SWORD-1 ... 324
Tabelle 4-164 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SWORD-2 ... 337
Tabelle 4-165 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie DEXA-Substudie ... 350

Abbildungsverzeichnis

Seite

Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel 63

Abbildung 2: Snapshot-Algorithmus 78

Abbildung 3: Meta-Analyse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT; DTG/RPV versus CAR 82

Abbildung 4: Meta-Analyse für EQ-5D Utility Score zu Woche 48 aus RCT; DTG/RPV versus CAR 91

Abbildung 5: Meta-Analyse für EQ-5D VAS zu Woche 48 aus RCT; DTG/RPV versus CAR..... 92

Abbildung 6: Meta-Analyse für SDM Symptom Count Score zu Woche 48 aus RCT; DTG/RPV versus CAR 98

Abbildung 7: Meta-Analyse für SDM Symptom Bother Score zu Woche 48 aus RCT; DTG/RPV versus CAR 99

Abbildung 8: Meta-Analyse für HIVTSQ Total Score zu Woche 48 aus RCT; DTG/RPV versus CAR 104

Abbildung 9: Meta-Analyse für unerwünschte Ereignisse bis Woche 48 aus RCT; DTG/RPV versus CAR 111

Abbildung 10: Meta-Analyse für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 aus RCT; DTG/RPV versus CAR 116

Abbildung 11: Meta-Analyse für schwerwiegende unerwünschte Ereignisse bis Woche 48 aus RCT; DTG/RPV versus CAR 121

Abbildung 12: Meta-Analyse für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten aus RCT; DTG/RPV versus CAR..... 125

Abbildung 13: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen aus RCT; DTG/RPV versus CAR..... 130

Abbildung 14: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts aus RCT; DTG/RPV versus CAR 135

Abbildung 15: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen aus RCT; DTG/RPV versus CAR 139

Abbildung 16: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen aus RCT; DTG/RPV versus CAR 146

Abbildung 17: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus RCT; DTG/RPV versus CAR 151

Abbildung 18: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems aus RCT; DTG/RPV versus CAR 156

Abbildung 19: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen aus RCT; DTG/RPV versus CAR..... 160

Abbildung 20: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege aus RCT; DTG/RPV versus CAR..... 164

Abbildung 21: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes aus RCT; DTG/RPV versus CAR 168

Abbildung 22: Meta-Analyse für knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 aus RCT; DTG/RPV versus CAR 177

Abbildung 23: Meta-Analyse für Knochenspezifischen Biomarker: P1NP zu Woche 48 aus RCT; DTG/RPV versus CAR 181

Abbildung 24: Meta-Analyse für Knochenspezifischen Biomarker: CTx zu Woche 48 aus RCT; DTG/RPV versus CAR 185

Abbildung 25: Meta-Analyse für Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 aus RCT; DTG/RPV versus CAR 189

Abbildung 26: Meta-Analyse für EQ-5D VAS zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl < 500 199

Abbildung 27: Meta-Analyse für EQ-5D VAS zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl ≥ 500 200

Abbildung 28: Meta-Analyse für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl < 500 222

Abbildung 29: Meta-Analyse für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl ≥ 500 223

Abbildung 30: Meta-Analyse für CTx zu Woche 48 aus RCT; DTG/RPV versus CAR – weiblich 231

Abbildung 31: Meta-Analyse für CTx zu Woche 48 aus RCT; DTG/RPV versus CAR – männlich 232

Abbildung 32: Meta-Analyse für CTx zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl < 500 235

Abbildung 33: Meta-Analyse für CTx zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl ≥ 500 236

Abbildung 34: Meta-Analyse für Osteocalcin zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline Third Agent = INI 241

Abbildung 35: Meta-Analyse für Osteocalcin zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline Third Agent = NNRTI 242

Abbildung 36: Meta-Analyse für Osteocalcin zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline Third Agent = PI 243

Abbildung 37: Flowchart für die Studie 201636 (SWORD-1) 302

Abbildung 38: Flowchart für die Studie 201637 (SWORD-2) 314

Abbildung 39: Flowchart für die Studie 202094 (DEXA-Substudie) 322

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AIDS	*Acquired immunodeficiency syndrome*
ALT	Alanin-Aminotransferase
ANCOVA	Kovarianzanalyse (*analysis of covariance*)
ART	Antiretrovirale Therapie
BMD	*bone mineral density*, Knochenmineraldichte
BSAP	*Bone Specific Alkaline Phosphatase*
CAR	*Current Antiretroviral Regimen*
CHMP	*Committee for Medicinal Products for Human Use*
CONSORT	*Consolidated Standards of Reporting Trials*
CTx	*type 1 collagen cross-linked C-telopeptide*
DAIDS	*Division of AIDS*
DEXA	*Dual energy X-ray absorptiometry*
DIMDI	Deutsches Institut für Medizinische Dokumentation
DTG	Dolutegravir
EG	Europäische Gemeinschaft
EQ-5D-5L	*European Quality of Life – 5 Dimensions*
FDA	*Food and Drug Administration*
G-BA	Gemeinsamer Bundesausschuss
HAART	*Highly active antiretroviral therapy*
HIV	Humanes Immundefizienz-Virus
HIV-1	Humanes Immundefizienz-Virus Typ 1
HIVTSQ	*HIV Treatment Satisfaction Questionnaire*
INI	Integrase-Inhibitor
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	*Intention to treat*
ITT-E	*Intention to treat - exposed*
LOCF	*last observation carried forward*
MedDRA	*Medical Dictionary for Regulatory Activities*
MTC	*Mixed Treatment Comparison*
NNRTI	Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren

NRTI	Nukleosidale Reverse-Transkriptase-Inhibitoren
P1NP	*procollagen type 1-N-propeptide*
PI	Protease-Inhibitor
PT	*preferred term*
pU	pharmazeutischer Unternehmer
RCT	*Randomized Controlled Trial*
RNA	Ribonukleinsäure (*ribonucleic acid*)
RPV	Rilpivirin
RR	Relatives Risiko
SDM	*Symptom Distress Module*
SGB	Sozialgesetzbuch
SOC	*System Organ Class*
STE	*Surrogate Threshold Effects*
STROBE	*Strengthening the Reporting of Observational Studies in Epidemiology*
TDF	Tenofovirdisoproxilfumarat
TREND	*Transparent Reporting of Evaluations with Non-Randomized Design*
ULN	obere Grenze des Normalbereichs (*upper limit of normal*)
VAS	visuelle Analogskala
WHO	*World Health Organization*
zVT	zweckmäßige Vergleichstherapie

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.

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

Der Integrase-Inhibitor (INI) Dolutegravir (DTG) ist als Einzelwirkstoff seit Januar 2014 in Europa zur Behandlung von HIV in Kombination mit anderen antiretroviralen Medikamenten zugelassen; Rilpivirin (RPV), ein Nicht-Nukleosidaler Reverse-Transkriptase-Inhibitor (NNRTI), ist seit November 2014 zugelassen. Beide Wirkstoffe besitzen darüber hinaus Zulassungen im Rahmen von weiteren Festkombinationen (1-3).

In der vorliegenden Nutzenbewertung soll nun das Ausmaß des Zusatznutzens der neuen Festkombination Dolutegravir/Rilpivirin (Juluca) bewertet werden. Dolutegravir/Rilpivirin ist angezeigt für die Behandlung von Infektionen mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) bei Erwachsenen, die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen

Als Ergebnis der Beratung zur zweckmäßigen Vergleichstherapie beim G-BA am 24. August 2017 (4, 5) hat der G-BA die Notwendigkeit gesehen, innerhalb des Anwendungsgebietes A zwei Teilpopulationen anhand des Grundes für einen Therapiewechsel separat zu betrachten. Diese beiden Teilpopulationen sind demnach wie folgt definiert:

- Antiretroviral vorbehandelte Patienten mit Umstellindikation (d.h. Patienten, die aufgrund von virologischem Versagen oder Nebenwirkungen oder Unverträglichkeiten die Vortherapie wechseln) bzw.

- antiretroviral vorbehandelte Patienten ohne Umstellindikation.

Patientenpopulation:

Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen.

Die Patientenpopulation wird dabei wie zuvor ausgeführt in zwei Teilpopulationen aufgeteilt und diese werden separat betrachtet:

- Antiretroviral vorbehandelte Patienten mit Umstellindikation (d.h. Patienten die aufgrund von virologischem Versagen oder Nebenwirkungen oder Unverträglichkeiten die Vortherapie wechseln) und

- antiretroviral vorbehandelte Patienten ohne Umstellindikation.

Intervention:

Dolutegravir/Rilpivirin 50 mg/25 mg ist einmal täglich als Tablette einzunehmen.

Vergleichstherapie:

Die Vergleichstherapie für das Anwendungsgebiet A besteht in einer individuellen antiretroviralen Therapie in Abhängigkeit der Vortherapie(n) und ggf. unter Berücksichtigung von Nebenwirkungen (4).

Bezüglich der beiden Teilpopulationen (mit oder ohne Umstellindikation) ergeben sich folgende zweckmäßige Vergleichstherapien:

- **Patienten mit Umstellindikation:** Die zweckmäßige Vergleichstherapie (zVT) entspricht der Festlegung der zVT für das gesamte Anwendungsgebiet A, da alle in Frage kommenden Wirkstoffkombinationen gleichermaßen zweckmäßig sind (4).

- **Patienten ohne Umstellindikation:** Die zVT besteht in der Fortführung der bisherigen Therapie (4).

Endpunkte:

Folgende patientenrelevante Endpunkte, die auch in den eingeschlossenen, randomisierten, kontrollierten Studien (RCT) erhoben wurden, werden in diesem Nutzendossier dargestellt:

Morbidität:

- Viruslast < 50 Kopien/ml zu Woche 48

Lebensqualität:

- EQ-5D
  - Utility Score
  - VAS
- Symptom Distress Module zu Woche 48
  - Symptom Count Score
  - Symptom Bother Score
- HIVTSQ Total Score

Mortalität:

- Anzahl der Todesfälle über 48 Wochen

Sicherheit/Verträglichkeit:

- Unerwünschte Ereignisse bis Woche 48
- Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48
- Schwerwiegende unerwünschte Ereignisse bis Woche 48
- Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten
- Unerwünschte Ereignisse von besonderem Interesse:
    - Herzerkrankungen
    - Erkrankungen des Gastrointestinaltrakts
    - Leber- und Gallenerkrankungen
    - Hypersensitivität
    - Stoffwechsel- und Ernährungsstörungen
    - Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen
    - Erkrankungen des Nervensystems
    - Psychiatrische Erkrankungen
    - Erkrankungen der Nieren und Harnwege
    - Erkrankungen der Haut und des Unterhautzellgewebes
- Knochenmineraldichte zu Woche 48
- Knochenspezifische Biomarker
    - Alkalische Phosphatase (BSAP)
    - Prokollagen 1-N-terminales Propeptid (P1NP)
    - CTx
    - Osteocalcin

Studientypen:

Es wurden zwei randomisierte, kontrollierte Studien (RCT) zur Bewertung des Zusatznutzens eingeschlossen: SWORD-1 und SWORD-2.

Datenquellen

Aus der Literatur- und Registerrecherche und den Studien des pharmazeutischen Unternehmers (pU) ergeben sich folgende Studien als Datenquellen:

Tabelle 4-1: Studienpool - RCT mit zu bewertenden Arzneimittel

Studie	Studientyp
SWORD-1 [201636]	RCT, offen, parallel
SWORD-2 [201637]	RCT, offen, parallel
DEXA Substudie [202094]	Substudie, offen, parallel

Die DEXA Substudie ist keine RCT im eigentlichen Sinne. In dieser Substudie werden bereits in die beiden Zulassungsstudien (SWORD-1 und SWORD-2) randomisierte Patienten eingeschlossen, die zusätzlich bestimmte Kriterien erfüllen (insbesondere das Vorhandensein von Tenofovirdisoproxilfumarat als ART-Komponente). Da sie auch im Zulassungsantrag als eigenständige Studie gelistet ist, wird sie hier extra aufgeführt, obwohl es für die Fragestellung keine eigenständige RCT ist.

Ein-/Ausschlusskriterien für Studien

Tabelle 4-2: Ein- und Ausschlusskriterien bei der Studienselektion

	Einschlusskriterien	**E**	**Ausschlusskriterien**	**A**
Patientenpopulation	Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen.	E1	Von E1 abweichende Patientenpopulation oder andere Altersgruppe der unter E1 genannten Patientenpopulation	A1
Intervention	Dolutegravir/Rilpivirin 50 mg/25 mg einmal täglich	E2	andere Intervention bzw. abweichende Dosierung	A2
Vergleichstherapie	individuelle antiretrovirale Therapie in Abhängigkeit der Vortherapie(n) und ggf. unter Berücksichtigung von Nebenwirkungen (bei Patienten ohne Umstellindikation wird diese Vergleichstherapie durch Fortführung ihrer bisherigen Therapie realisiert)	E3	andere Vergleichstherapie	A3
Endpunkte	Ergebnisse zu mind. einem patientenrelevanten Endpunkt bzgl. • Morbidität • Lebensqualität • Mortalität • Sicherheit/Verträglichkeit sind berichtet	E4	Ergebnisse zu keinem der nebenstehend genannten Endpunkte sind berichtet	A4
Studientyp	randomisierte, kontrollierte Studie (RCT)	E5	nicht randomisierte Studie	A5
Publikationstyp	Primärpublikation oder Sekundärpublikation mit neuen Daten zu den unter E4 genannten Endpunkten oder ausführlicher Studienbericht	E6	Sekundärpublikation ohne neue Daten oder ohne Angaben zu patientenrelevanten Endpunkten	A6
Studiendauer	mindestens 48 Wochen	E7	kürzer als 48 Wochen	A7

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

Studienqualität:

Für die vorliegende Darstellung des Zusatznutzens wurden die Ergebnisse zweier randomisierter klinischer Studien (RCT) und einer Substudie von hoher Qualität herangezogen. Die Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial und ermöglichen den direkten Vergleich des zu bewertenden Arzneimittels Dolutegravir/Rilpivirin mit der zweckmäßigen Vergleichstherapie.

Validität der herangezogenen Endpunkte:

Alle Endpunkte, die für die Darstellung der Wirksamkeit und Sicherheit im Rahmen der Nutzenbewertung von Dolutegravir/Rilpivirin herangezogen wurden, sind patientenrelevant.

Die beiden Instrumente zur Erhebung der Lebensqualität (EQ-5D und HIVTSQ) sind im jeweils vorliegenden Indikationsgebiet der HIV-Infektion validierte und allgemein anerkannte Messinstrumente.

Evidenzstufe:

Der Nachweis für den Zusatznutzen beruht auf den Ergebnissen einer Meta-Analyse zweier randomisierter klinischer Studien und einer Substudie und entspricht somit dem höchstmöglichen Evidenzgrad.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Tabelle 4-3: Übersicht der Ergebnisse pro Endpunkt – Teilpopulation ohne Umstellindikation

	Effektschätzer der Meta-Analyse (Modell mit festen Effekten) **Hedges' g oder mittlere Änderung oder RR [95% KI], p-Wert**	**Wahrscheinlichkeit und Ausmaß des Zusatznutzens für DTG/RPV**
Viruslast < 50 Kopien/ml zu Woche 48	RR: 1,00 [0,98; 1,03], p=0,970	Zusatznutzen nicht belegt
EQ-5D zu Woche 48 Utility Score	Hedges' g: 0,014 [-0,110; 0,137], p=0,830	Zusatznutzen nicht belegt
VAS	Hedges' g: -0,005 [-0,128; 0,118], p=0,958 CD4+ Zahl < 500 Hedges' g: 0,193 [-0,030; 0,416], p=0,090 CD4+ Zahl ≥ 500 Hedges' g: -0,093 [-0,241; 0,054], p=0,216	- Zusatznutzen nicht belegt Zusatznutzen nicht belegt
Symptom Distress Module zu Woche 48 Symptom Count Score	Hedges' g: -0,118 [-0,250; 0,015], p=0,081	Zusatznutzen nicht belegt
Symptom Bother Score	Hedges' g: -0,167 [-0,299; -0,034], p=0,014	Beleg für geringen Zusatznutzen
HIVTSQ Total Score zu Woche 48	Hedges' g: 0,168 [0,045; 0,291], p=0,007	Beleg für geringen Zusatznutzen
Mortalität bis Woche 48	wegen sehr weniger Ereignisse keine Meta-Analyse durchgeführt	Zusatznutzen nicht belegt
Unerwünschte Ereignisse bis Woche 48	RR: 1,08 [1,01; 1,16], p=0,037	Zusatznutzen nicht belegt[b]
Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48[a]	SWORD-1 RR: 0,98 [0,49; 1,95], p=0,955 SWORD-2 RR: 3,28 [1,23; 8,76], p=0,018	Zusatznutzen nicht belegt[b]
Schwerwiegende unerwünschte Ereignisse bis Woche 48	RR: 1,31 [0,73; 2,34], p=0,361	Zusatznutzen nicht belegt
Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten	RR: 6,33 [1,87; 21,38], p=0,003	Zusatznutzen nicht belegt[b]
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen	RR: 1,01 [0,14; 7,27], p=0,992	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts[a]	SWORD-1 RR: 1,28 [0,95; 1,74], p=0,104 SWORD-2 RR: 2,16 [1,41; 3,32], p<0,001	Zusatznutzen nicht belegt[b]

Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen		RR: 2,52 [0,49; 12,97], p=0,268	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität		wegen sehr weniger Ereignisse keine Meta-Analyse durchgeführt	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen		RR: 1,05 [0,41; 2,69], p=0,926	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen[a]	SWORD-1 SWORD-2	RR: 0,71 [0,48; 1,05], p=0,088 RR: 1,33 [0,87; 2,03], p=0,184	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems[a]	SWORD-1 SWORD-2	RR: 1,36 [0,86; 2,14], p=0,193 RR: 2,78 [1,55; 4,98], p<0,001	Zusatznutzen nicht belegt[b]
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen		RR: 1,88 [1,25; 2,83], p=0,003	Zusatznutzen nicht belegt[b]
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege		RR: 1,61 [0,70; 3,72], p=0,263	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes	 CD4+ Zahl < 500 CD4+ Zahl ≥ 500	RR: 1,49 [1,05; 2,13], p=0,027 RR: 0,89 [0,50; 1,60], p=0,696 RR: 2,02 [1,28; 3,20], p=0,003	- Zusatznutzen nicht belegt Zusatznutzen nicht belegt[b]
Knochenmineraldichte [g/cm²][c] Hüfte		adj. mittlere Änderung: 1,29 [0,27; 2,31], p=0,014	Hinweis auf geringen Zusatznutzen
Lendenwirbel		adj. mittlere Änderung: 1,32 [0,07; 2,57], p=0,039	Hinweis auf geringen Zusatznutzen
Knochenspezifischer Biomarker: BSAP zu Woche 48		Hedges' g: -1,020 [-1,154; -0,886], p<0,001	Zusatznutzen nicht belegt[b]
Knochenspezifischer Biomarker: P1NP zu Woche 48		Hedges' g: -0,529 [-0,658; -0,400], p<0,001	Zusatznutzen nicht belegt[b]
Knochenspezifischer Biomarker: CTx zu Woche 48	 weiblich männlich CD4+ Zahl < 500 CD4+ Zahl ≥ 500	Hedges' g: -0,252 [-0,380; -0,124], p<0,001 Hedges' g: -0,523 [-0,801; -0,245], p<0,001 Hedges' g: -0,208 [-0,352; -0,065], p=0,004 Hedges' g: -0,595 [-0,830; -0,360], p<0,001 Hedges' g: -0,173 [-0,326; -0,021], p=0,026	Zusatznutzen nicht belegt[b,d] - - - -

Knochenspezifischer Biomarker: Osteocalcin zu Woche 48		Hedges' g: -0,734 [-0,866; -0,602], p<0,001	Zusatznutzen nicht belegt[b,d]
	Third Agent = INI	Hedges' g: -0,608 [-0,900; -0,316], p<0,001	-
	Third Agent = NNRTI	Hedges' g: -0,545 [-0,720; -0,370], p<0,001	-
	Third Agent = PI	Hedges' g: -1,255 [-1,530; -0,980], p<0,001	-

[a] Darstellung der Ergebnisse der Studien SWORD-1 und SWORD-2, da in der Meta-Analyse eine hohe Heterogenität vorlag
[b] Zur Interpretation siehe hierzu die Begründung im Text direkt im Anschluss an diese Tabelle
[c] Ergebnis der Einzelstudie (DEXA Substudie)
[d] Bewertung im Gesamtkollektiv möglich, da die Ergebnisse in allen relevanten Subgruppenanalysen (Geschlecht, CD4-Zellzahl, Third Agent) gleichgerichtet sind

Das **Ausmaß und die Wahrscheinlichkeit des Zusatznutzens für die gesamte Zielpopulation, d.h. für vorbehandelte erwachsene Patienten,** wird jeweils auf der Grundlage der Ergebnisse aller patientenrelevanter Endpunkte über die längste verfügbare Behandlungsdauer von 48 Wochen bewertet.

Weiterhin folgt aus der Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA, dass die **Zielpopulation** in zwei Teilpopulationen entsprechend des Grundes zur Therapieumstellung einzuteilen, separat zu betrachten und zu bewerten ist. Dies sind zum einen die **Teilpopulation mit Umstellindikation** (Umstellindikationen sind nach der Definition des G-BA virologisches Versagen, Nebenwirkungen oder Toxizitäten), sowie die **Teilpopulation ohne Umstellindikation**.

Diese beiden Teilpopulationen werden im Folgenden separat diskutiert.

Patienten der Zielpopulation ohne Umstellindikation:

Der Zusatznutzen wird bezüglich der zuvor bereits in Tabelle 4-3 pro Endpunkt gegebenen Einschätzung über Wahrscheinlichkeit und Ausmaß des Zusatznutzens zusammenfassend beschrieben und begründet.

Morbidität:

Ergebnisse zur Morbidität werden anhand des Endpunktes „Viruslast < 50 Kopien/ml zu Woche 48“ bewertet. In Konsequenz der Indikation und des Studien-Designs, das für diesen Endpunkt in beiden SWORD-Zulassungstudien eine Nicht-Unterlegenheits-Hypothese geprüft hat, war eine Überlegenheit für das zu bewertende Arzneimittel Dolutegravir/Rilpivirin (DTG/RPV) im Vergleich zur Fortführung der bisherigen ART (CAR) für diesen Endpunkt nicht zu erwarten. Die Patientenpopulation der beiden SWORD-Studien bestand entsprechend der Indikation aus Patienten, zu Studienbeginn die eine Viruslast < 50 Kopien/ml aufweisen. Dieses Kriterium war ein Einschlusskriterium für die SWORD-Studien, es gab einen sehr geringen Anteil von weniger als 2 % aller Patienten (für beide Studien SWORD-1 und SWORD-2), die bei Baseline dieses Kriterium nicht erfüllten, obwohl es auch für diese Patienten, also insgesamt für 100 % der Patienten, zum Screening-Zeitpunkt erfüllt war. Insofern waren Verbesserungen im Vergleich zu Baseline nicht zu erwarten und die Fragestellung bestand darin festzustellen, inwieweit sich ein Verlust des virologischen Erfolgs beider Therapien über 48 Wochen einstellen und ggfs. unterscheiden würde.

Im Ergebnis gab es keinen Unterschied bezüglich dieses Endpunktes zwischen DTG/RPV und CAR. Die Viruslast war auch am Ende beider Studien nach 48 Wochen Behandlungsdauer für mindestens 94 % der Patienten in beiden Behandlungsgruppen mit Werten < 50 Kopien/ml kontrolliert und die Nicht-Unterlegenheit wurde damit statistisch signifikant belegt. Die meta-analytische Auwertung zum Endpunkt „Viruslast < 50 Kopien/ml zu Woche 48“ konnte die nahezu perfekte Übereinstimmung zwischen DTG/RPV und der zweckmäßigen Vergleichstherapie CAR zeigen (RR = 1,0). Ein Zusatznutzen für DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie CAR wurde hinsichtlich dieses Endpunktes nicht belegt.

Ergänzend zu diesem Endpunkt sollte auch das Auftreten von AIDS-definierenden Ereignissen (CDC Klasse C) und/oder Todesfällen überprüft werden, um zu belegen, dass sich die beiden Behandlungsgruppen auch in diesem klinischen Endpunkt nicht unterscheiden. Wie im Patientenkollektiv ebenfalls zu erwarten war, traten unter beiden Behandlungen nur sehr wenige Ereignisse einer Progression zur Klasse CDC C im Studienverlauf auf und die relative Häufigkeit unterschied sich nicht relevant und auch nicht statistisch signifikant zwischen den Behandlungen.

Damit stützen diese Ergebnisse die zum Endpunkt „Viruslast < 50 Kopien/ml zu Woche 48“ getroffenen Aussagen sehr gut.

Mortalität:

Aufgrund des eingeschlossenen Patientenkollektivs war auch das Auftreten von Todesfällen im Studienverlauf beider SWORD-Zulassungsstudien nur äußerst selten zu erwarten. Es traten in beiden SWORD-Studien nur insgesamt zwei Todesfälle auf, einer je Behandlungsgruppe. Ein Zusatznutzen für DTG/RPV wurde hinsichtlich der Mortalität nicht belegt.

Gesundheitsbezogene Lebensqualität:

Zur Bewertung der gesundheitsbezogenen Lebensqualität wurden die Endpunkte „EQ-5D“, „*Symptom Distress Module*“ (SDM) und „*HIV Treatment Satisfaction Questionnaire*“ (HIVTSQ) in beiden SWORD-Zulassungsstudien erhoben und die Ergebnisse jeweils zu Woche 48 dargestellt.

Der EQ-5D wurde hinsichtlich des *Utitlity Scores* und der visuellen Analogskala (VAS) separat dargestellt. Es ergaben sich jeweils bei nur sehr geringen Änderungen keine statistisch signfikanten Unterschiede zwischen den Behandlungsgruppen DTG/RPV und CAR. Ein Zusatznutzen ist damit aus den Ergebnissen des EQ-5D nicht belegt.

Für das SDM wurden zwei Dimensionen ausgewertet und dargestellt: der *Symptom Count Score*, der die Häufigkeit aufgetretener Symptome angibt und der *Symptom Bother Score*, bei dem der Patient die Belastung durch aufgetretene Symptome der Erkrankung einschätzt. Verringerungen in diesen beiden Scores zeigen jeweils Verbesserungen an. In beiden Scores treten numerische Verbesserungen zugunsten von DTG/RPV im Vergleich zu CAR auf, die für den *Symptom Bother Score* auch statistisch signifikant sind. Dies bedeutet in der Gesamtschau für das SDM, dass die Patienten im DTG/RPV-Arm im Vergleich zu CAR weniger Symptome berichten und diese zusätzlich statistisch signifikant als weniger belastend eingeschätzt werden, was eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung der Lebensqualität darstellt. Da das Ergebnis durch die Meta-Analyse von zwei hochwertigen RCT nachgewiesen wurde, ergibt sich die Aussagesicherheit eines Belegs. Insgesamt wird aus den Ergebnissen zum SDM daher ein Beleg für einen geringen Zusatznutzen für DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie (CAR) abgeleitet.

Für den HIVTSQ ergibt sich zu Woche 48 eine statistisch signifikante Verbesserung der Therapiezufriedenheit für DTG/RPV im Vergleich zu CAR. Dies bedeutet eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung der Lebensqualität. Es ist dabei hervorzuheben, dass die Behandlungszufriedenheit bei Baseline mit Werten von 54 bis zu 55 Punkten bei einem Maximalscore von 60 (entspricht höchster Therapiezufriedenheit) bereits sehr hoch lag und von daher Verbesserungen im Patientenkollektiv nur schwer zu erreichen waren, weshalb deren Eintreten besonders hoch zu bewerten ist. Da das Ergebnis durch die Meta-Analyse von zwei hochwertigen RCT nachgewiesen wurde, ergibt sich die Aussagesicherheit eines Belegs. Insgesamt wird aus den Ergebnissen des HIVTSQ daher ein Beleg für einen geringen Zusatznutzen für DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie (CAR) abgeleitet.

Sicherheit/Verträglichkeit:

Aufgrund des in den beiden Studien bestehenden Verzerrungspotenzials, ergibt sich die Notwendigkeit, insbesondere die Ergebnisse zur Sicherheit und Verträglichkeit kritisch zu betrachten. In den beiden offenen RCT SWORD-1 und SWORD-2 besteht ein potenzieller Bias zur kritischeren Beurteilung insbesondere von Sicherheit und Verträglichkeit Die Patienten in der CAR-Gruppe setzten ihre seit mindestens 6 Monaten bestehende, effektive (Viruslast < 50 Kopien/ml) und stabil ohne Nebenwirkungen und Toxizitäten verlaufende ART fort, wobei die mediane Dauer der bestehenden Therapie sogar noch sehr viel größerwar, als die geforderten 6 Monate. Der randomisierte Vergleich in beiden Studien erfolgt jedoch gegenüber einer für 87 % der auf DTG/RPV randomisierten Patienten komplett neuen ART, d.h. 87 % der Patienten hatten bisher noch keinen der beiden Wirkstoffe jemals erhalten und für 100 % der Patienten auf DTG/RPV war die Wirkstoff-Kombination aus DTG/RPV neu. Diese Zuteilung erfolgte offen, so dass Prüfärzte und Patienten diese kannten. Es ist daher naheliegend, dass primär im Sinne der Patientensicherheit und Aufrechterhaltung einer effektiven ART das Potential besteht, die neue Therapie mit DTG/RPV, trotz der Bekanntheit ihrer Einzelkomponenten, im Studienkontext besonders kritisch zu bewerten, um im Zweifelsfall möglichst schnell die Studie für einen Patienten abzubrechen, um den Patienten auf eine andere Therapie umstellen zu können. Auch die Meldung und Bewertung von unerwünschten Ereignissen, für die die subjektive Einschätzung, ob diese als bereits melderelevant anzusehen sind oder nicht, entscheidend sein kann, ist das Verzerrungspotenzial im hier vorliegenden offenen Studiendesigns kritisch zu bewerten.

Vor diesem Hintergrund wurden die Endpunkte „unerwünschte Ereignisse bis Woche 48“, „unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten“ sowie alle dargestellten „unerwünschten Ereignisse von besonderem Interesse bis Woche 48“ als mit hohem Verzerrungspotenzial behaftet eingestuft. Die Kategorien „unerwünschte Ereignisse vom Grad 3 oder 4 bis Woche 48“ und „schwerwiegende unerwünschte Ereignisse bis Woche 48“ wurden wegen ihrer Schwere bzw. der aus den Ereignissen ggfs. entstandenen Konsequenzen oder der erforderlichen Gegenmaßnahmen als nur mit geringem Verzerrungspotenzial behaftet eingestuft. Der spezifische Charakter schwerer Ereignisse lässt

das oben beschriebene Verzerrungspotenzial somit eher in den Hintergrund treten. Diese Einschätzung zum Verzerrungspotenzial der Endpunkte zur Sicherheit/Verträglichkeit steht auch in Übereinstimmung mit der Bewertung, die das IQWiG kürzlich im Rahmen der Nutzenbewertung von Darunavir/Cobicistat/Emtricitabin/Tenofoviralafenamid in der gleichen Indikation und einem vom Studiendesign her weitgehend vergleichbaren Setting getroffen hat (6).

In ihrer Bewertung der SWORD-Studien hat das CHMP (Committee for Medicinal Products for Human Use) im Zusammenhang mit der Bewertung der Sicherheit/Verträglichkeit ebenfalls deutlich gemacht, dass unterschiedliche Häufigkeiten gemeldeter unerwünschter Ereignisse (allgemein), von Therapieabbrüchen wegen unerwünschter Ereignisse und zu psychiatrischen unerwünschten Ereignissen in diesem unverblindeten Switch-Studiendesign – was zu teilweise häufigeren Meldungen im DTG/RPV-Arm im Vergleich zum CAR-Arm führte – sehr plausibel durch Verzerrungen (Reporting-Bias) bedingt sein können. Insgesamt werden die Arten der berichteten Ereignisse von den Einzelsubstanzen her als generell bekannt und auch (in der hier geprüften Kombination von DTG und RPV) nicht häufiger auftretend, als bei den Einzelsubstanzen, eingeschätzt. Dies führt das CHMP zu der summarischen Bewertung, dass keine zusätzlichen Risiken oder Sicherheitsprobleme identifiziert wurden. (7)

Vor diesem Hintergrund werden im Folgenden die Aspekte des Zusatznutzens hinsichtlich der Sicherheit/Verträglichkeit diskutiert. Zu den Endpunkten „unerwünschte Ereignisse bis Woche 48“ und „Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten“ liegen Ergebnisse vor, die eine höhere relative Häufigkeit unter DTG/RPV im Vergleich zu CAR beschreiben. Die Ergebnisse sind für beide Endpunkte statistisch signifikant.

Der meta-analytische Effektschätzer zu allen „unerwünschten Ereignissen bis Woche 48“ beschreibt trotz der statistischen Signifikanz einen sehr kleinen Effekt. Bei Invertieren der Effektrichtung im Vergleich zur Darstellung in Tabelle 4-3 ergibt sich eine obere Konfidenzintervallgrenze für das relative Risiko von 0,99. Damit liegt nach Einschätzung des IQWiG (wie kürzlich in der Nutzenbewertung für Darunavir/Cobicistat/Emtricitabin/Tenofoviralafenamid getroffen (6)), ein für diese Kategorie der nicht schwerwiegenden bzw. nicht schweren unerwünschten Ereignisse, ein nicht mehr als geringfügiges Schadenspotenzial für DTG/RPV vor. Unter zusätzlicher Berücksichtigung des zuvor beschriebenden hohen Verzerrungspotenzials dieses Endpunktes kann aus dem Ergebnis zu diesem Endpunkt kein Aspekt eines geringeren Nutzens des zu bewertenden Kombinationsarzneimittels DTG/RPV abgeleitet werden. Zur Häufigkeit aller unerwünschter Ereignisse konstatiert das CHMP ebenfalls, dass der zuungunsten von DTG/RPV verzerrende Einfluss des Switch-Designs eine plausible Erklärung für die berichteten unterschiedlichen Häufigkeiten ist (7).

Bei dem Endpunkt „Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten“ ist das Ausmaß des meta-analytischen Effektschätzers größer, jedoch liegt für diesen Endpunkt ebenfalls ein hohes Verzerrungspotenzial vor. Es liegt die Vermutung nahe, dass es sich um einen Trend handelt, der zu erheblichen Anteilen durch einen Reporting-Bias verursacht wird. Gerade offene RCT sowie die hier vorliegenden Studien (SWORD-1 und

SWORD-2) jeweils zusätzlich durch Vorhandensein des Switch-Designs, sind anfällig für diese Art von Bias. Da auch die Ergebnisse zur Lebensqualität, zum SDM sowie zu schwerwiegenden unerwünschten Ereignissen und zu unerwünschten Ereignissen vom Grad 3 oder 4 das Ergebnis dieses Endpunktes „Unerwünschten Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten“ nicht widerspiegeln, wird auch aus diesem Ergebnis kein Aspekt eines geringeren Nutzens des zu bewertenden Kombinationsarzneimittels DTG/RPV abgeleitet. Auch das CHMP bewertet die häufigeren Meldungen von unerwünschten Ereignissen, die zum Studienabbruch führten, als primär durch die Gegebenheiten des Switch-Designs verursacht und kommt zu dem Schluss, dass sich daraus kein Sicherheitsproblem ableiten lässt (7).

Die Analyse der Endpunkte „Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48“ sowie der beiden Kategorien unerwünschter Ereignisse von besonderem Interesse bis Woche 48 „Erkrankungen des Gastrointestinaltrakts“ und „Erkrankungen des Nervensystems“ ergaben, dass die Meta-Analysen über beide SWORD-Studien eine hohe Heterogenität aufwiesen. Wegen der begründeten Auswertung im Modell mit festen Effekten ist die Meta-Analyse zur Interpretation der Gesamtaussage daher nicht geeignet und es werden daher bei der Diskussion dieser Endpunkte sowie in der Gesamtschau die Ergebnisse der Einzelstudien herangezogen. Es zeigt sich für alle drei Endpunkte, dass ein statistisch signifikanter Effekt im Sinne einer höheren relativen Auftretenshäufigkeit unter DTG/RPV im Vergleich zu CAR in jeweils nur einer Studie (SWORD-2) auftritt. Die jeweils andere Studie (SWORD-1) zeigte ein statistisch nicht signifikantes Ergebnis, wobei ein Trend zuungunsten von DTG/RPV entweder nur sehr schwach (unerwünschte Ereignisse von besonderem Interesse „Erkrankungen des Gastrointestinaltrakts“ und „Erkrankungen des Nervensystems“) oder gar nicht („unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48“) vorlag. Es ist auch auffällig, dass für alle drei Endpunkte in der Studie SWORD-2 ein hauptsächlicher Unterschied in den Ergebnissen in einer im Vergleich zu SWORD-1 deutlich geringeren relativen Häufigkeit der jeweiligen Ereignisse im CAR-Arm lag und nicht etwa in einer Erhöhung im DTG/RPV-Arm. Die Ergebnisse im DTG/RPV-Arm waren hingegen für diese drei hier diskutierten Endpunkte über beide Studien sehr gut vergleichbar. Ohne, dass dafür eine plausible Erklärung identifiziert werden konnte, ist das Ergebnis der hohen Heterogenität der Meta-Analysen im Wesentlichen auf diesen Unterschied im CAR-Arm zwischen den beiden SWORD-Studien zurückzuführen. Folglich kann für diese drei Endpunkte jeweils nicht von einer klaren Evidenzsituation über die vorhandene Gesamtevidenz ausgegangen werden. Daher wird festgestellt, dass für diese Endpunkte weder ein Zusatznutzen noch ein geringerer Nutzen für DTG/RPV als belegt angesehen werden kann. Bei der Interpretation ist ebenfalls das hohe Verzerrungspotenzial bezüglich der beiden Kategorien der unerwünschten Ereignisse von besonderem Interesse zu berücksichtigen.

Für die beiden Kategorien unerwünschter Ereignisse von besonderem Interesse bis Woche 48 „Psychiatrische Erkrankungen“ und „Erkrankungen der Haut und des Unterhautzellgewebes“ ergaben sich bei hohem Verzerrungspotenzial in der Meta-Analyse jeweils statistisch signifikante Ergebnisse einer höheren relativen Häufigkeit dieser Ereignisse unter DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie CAR.

Bei dem Endpunkt „Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes“ lag zusätzlich ein Beleg auf Effektmodifikation hinsichtlich der CD4-Zellzahl bei Baseline vor. Die Effekte in den beiden Subgruppen waren nicht gleichgerichtet, so dass eine Interpretation je Subgruppe erfolgt. Nur in der Subgruppe der Patienten, die bei Baseline eine CD4-Zellzahl von mindestens 500 Zellen/mm^3 hatten, lag ein statistisch signifikantes Ergebnis für eine höhere relative Häufigkeit unter DTG/RPV vor. Im alternativen Stratum gab es einen leichten numerischen Trend in die entgegengesetzte Richtung (weniger unerwünschte Ereignisse der Kategorie unter DTG/RPV im Vergleich zu CAR), der jedoch statistisch nicht signifikant war. Für diesen Endpunkt waren die aufgetretenen Ereignisse für alle Patienten mit Ereignissen maximal vom Grad 1 oder 2 (bis auf einen Patienten im CAR-Arm).

Für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Ereignisse“ waren nur für insgesamt 5 Patienten in der Gruppe DTG/RPV vom Grad 3 oder 4 (das waren 8,2 % aller unerwünschten psychiatrischen Ereignisse unter DTG/RPV) im Vergleich zu 2 Patienten in der CAR-Gruppe (6,3 % aller unerwünschten psychiatrischen Ereignisse unter CAR). Damit waren in beiden Behandlungsgruppen deutlich über 90 % aller unerwünschten psychiatrischen Ereignisse vom Grad 1 oder 2 und der jeweilige Anteil der Ereignisse vom Grad 3 oder 4 unterscheidet sich auch nur geringfügig zwischen DTG/RPV und CAR.

Da es sich somit in der großen Mehrheit der Ereignisse in beiden hier diskutierten Endpunkten um leichte oder maximal mäßig schwere Ereignisse handelte wird – auch vor dem Hintergrund des hohen Verzerrungspotenzials der beiden Endpunkte – insgesamt daraus kein geringerer Nutzen für DTG/RPV als belegt angesehen.

Da generell keine neuen Arten von unerwünschten psychiatrischen Ereignissen aufgetreten sind als solche, die bereits von den Einzelsubstanzen DTG und RPV bekannt sind, und da auch die Häufigkeit dem Erwartungswert dafür entspricht, kommt das CHMP zum Schluss, dass es keinen Hinweis darauf gibt, dass DTG/RPV zu einem additiven oder supraadditiven Risiko für psychiatrische unerwünschte Ereignisse beiträgt (7).

Zu allen weiteren dargestellten Endpunkten der Sicherheit/Verträglichkeit (wie insbesondere schwerwiegenden unerwünschten Ereignissen und den weiteren dargestellten unerwünschten Ereigniskategorien von besonderem Interesse) gab es keine weiteren Aspekte, die auf eine zwischen DTG/RPV und der zweckmäßigen Vergleichstherapie CAR unterschiedliche Sicherheit/Verträglichkeit hinweisen würden.

Bezüglich des Endpunktes „Knochenmineraldichte“ zeigte sich an beiden Messlokalisationen (Hüfte und Lendenwirbel) in der DEXA-Substudie jeweils ein statistisch signifikanter Anstieg über 48 Wochen in der Gruppe DTG/RPV im Vergleich zu CAR. Die ebenfalls dargestellten knochenspezifischen Biomarker (BSAP, P1NP, CTx und Osteocalcin jeweils zu Woche 48) unterstützen dieses Ergebnis zur Knochenmineraldichte durch für jeden dieser Endpunkte in den Meta-Analysen gezeigte statistisch signifikante Verbesserungen unter DTG/RPV im

Vergleich zu CAR. Diese Endpunkte der knochenspezifischen Endpunkte werden jedoch nicht zur separaten Ableitung eines Zusatznutzens herangezogen.

Der Endpunkt Knochenmineraldichte besitzt prädiktive Wertigkeit für Knochenfrakturen. Die Auswertung erfolgte durch verblindete Bewerter, so dass für diesen Endpunkt kein hohes Verzerrungspotenzial vorliegt. Die Ergebnisse zu im Studienverlauf über 48 Wochen aufgetretenen Frakturen zeigen einen leichten positiven Trend zugunsten von DTG/RPV, der aber keine statistische Signifikanz aufweist. Die Aussagesicherheit des Endpunktes „Knochenmineraldichte“ wird (auf der Subpopulation zweier RCT beruhend) als Hinweis eingeschätzt. Insgesamt ergibt sich aus den Ergebnissen zum Endpunkt „Knochenmineraldichte“ damit ein Hinweis für einen geringen Zusatznutzen für DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie.

Zusammenfassung für die Teilpopulation ohne Umstellindikation:

In der **Teilpopulation ohne Umstellindikation** ergibt sich zusammenfassend auf der Grundlage einer vollkommen vergleichbaren Wirksamkeit, einem in der Gesamtschau nicht belegten höheren oder geringeren Nutzen hinsichtlich der dargestellten Endpunkte zu Nebenwirkungen, einem Hinweis auf einen geringen Zusatznutzen hinsichtlich der Knochenmineraldichte und einem Beleg für einen geringen Zusatznutzen hinsichtlich der Lebensqualität in der Gesamtschau ein **Hinweis auf einen geringen Zusatznutzen** für das zu bewertende Kombinationsarzneimittel Dolutegravir/Rilpivirin im Vergleich zur zweckmäßigen Vergleichstherapie „Fortführung der bestehenden ART“.

Die Bewertung zu den diskutierten, teilweise häufigeren, Meldungen unerwünschter Ereignisse berücksichtigt auch die Einschätzung des CHMP, das zu dem Schluss kommt, dass diese plausibel als Folge des diesbezüglich zuungunsten von DTG/RPV verzerrenden Einflusses des (unverblindeten) Switch-Designs sind und daraus keine zusätzlichen Risiken oder Verträglichkeitsprobleme abzuleiten sind (7).

Trotzdem erfolgt die vorgenommene Herabstufung der Aussagesicherheit von einem möglichen Beleg auf einen Hinweis unter Berücksichtigung der Aspekte hinsichtlich der Heterogenität in drei Endpunkten der Sicherheit/Verträglichkeit zwischen den beiden Zulassungsstudien SWORD-1 und SWORD-2 und dem mit hohem Verzerrungspotenzial behafteten Ergebnissen zu dem Endpunkt „unerwünschte Ereignisse bis Woche 48, die zum Studienabbruch geführt hatten“.

<u>Patienten der Zielpopulation mit Umstellindikation:</u>

Für die **Teilpopulation mit Umstellindikation** lag keine RCT-Evidenz vor, somit kann für diese Teilpopulation kein Zusatznutzen begründet werden. Für die Teilpopulation mit Umstellindikation ist damit kein Zusatznutzen belegt.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

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

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen - **antiretroviral vorbehandelte Patienten ohne Umstellindikation.**	Hinweis auf geringen Zusatznutzen
Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen - **antiretroviral vorbehandelte Patienten mit Umstellindikation.**	Kein Zusatznutzen belegt

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

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

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.

Der Integrase-Inhibitor (INI) Dolutegravir ist als Einzelwirkstoff seit Januar 2014 in Europa zur Behandlung von HIV in Kombination mit anderen antiretroviralen Medikamenten zugelassen. Rilpivirin, ein Nicht-Nukleosidaler Reverse-Transkriptase-Inhibitor (NNRTI), ist seit November 2014 zugelassen. Beide Wirkstoffe besitzen darüber hinaus Zulassungen im Rahmen von weiteren Festkombinationen (1-3).

In der vorliegenden Nutzenbewertung soll nun das Ausmaß des Zusatznutzens der neuen Festkombination Dolutegravir/Rilpivirin (Juluca) bewertet werden. Dolutegravir/Rilpivirin ist indiziert zur Behandlung von Erwachsenen, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen.

Als Ergebnis der Beratung zur zweckmäßigen Vergleichstherapie beim G-BA am 24. August 2017 (4, 5) hat der G-BA die Notwendigkeit gesehen, innerhalb des Anwendungsgebietes A zwei Teilpopulationen anhand des Grundes für einen Therapiewechsel separat zu betrachten. Diese beiden Teilpopulationen sind demnach wie folgt definiert:

- Antiretroviral vorbehandelte Patienten mit Umstellindikation (d.h. Patienten, die aufgrund von virologischem Versagen oder Nebenwirkungen oder Unverträglichkeiten die Vortherapie wechseln) bzw.

- antiretroviral vorbehandelte Patienten ohne Umstellindikation.

Patientenpopulation:

Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen.

Die Patientenpopulation wird dabei wie zuvor ausgeführt in zwei Teilpopulationen aufgeteilt und diese werden separat betrachtet:

- Antiretroviral vorbehandelte Patienten mit Umstellindikation (d.h. Patienten die aufgrund von virologischem Versagen oder Nebenwirkungen oder Unverträglichkeiten die Vortherapie wechseln) und

- antiretroviral vorbehandelte Patienten ohne Umstellindikation.

Intervention:

Dolutegravir/Rilpivirin 50 mg/25 mg ist einmal täglich als Tablette einzunehmen.

Vergleichstherapie:

Die Vergleichstherapie für das Anwendungsgebiet A besteht in einer individuellen antiretroviralen Therapie in Abhängigkeit der Vortherapie(n) und ggf. unter Berücksichtigung von Nebenwirkungen (4).

Bezüglich der beiden Teilpopulationen (mit oder ohne Umstellindikation) ergeben sich folgende zweckmäßige Vergleichstherapien:

- **Patienten mit Umstellindikation**: Die zweckmäßige Vergleichstherapie (zVT) entspricht der Festlegung der zVT für das gesamte Anwendungsgebiet A, da alle in Frage kommenden Wirkstoffkombinationen gleichermaßen zweckmäßig sind (4).

- **Patienten ohne Umstellindikation**: Die zVT besteht in der Fortführung der bisherigen Therapie (4).

Endpunkte:

Folgende patientenrelevante Endpunkte, die auch in den eingeschlossenen randomisierten, kontrollierten Studien (RCT) erhoben wurden, werden in diesem Nutzendossier dargestellt:

Morbidität:

- Viruslast < 50 Kopien/ml zu Woche 48

Lebensqualität:

- EQ-5D
  - Utility Score
  - VAS
- Symptom Distress Module zu Woche 48
  - Symptom Count Score
  - Symptom Bother Score
- HIVTSQ Total Score

Mortalität:

- Anzahl der Todesfälle über 48 Wochen

Sicherheit/Verträglichkeit:

- Unerwünschte Ereignisse bis Woche 48
- Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48
- Schwerwiegende unerwünschte Ereignisse bis Woche 48
- Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten
- Unerwünschte Ereignisse von besonderem Interesse:
  - Herzerkrankungen
  - Erkrankungen des Gastrointestinaltrakts
  - Leber- und Gallenerkrankungen
  - Hypersensitivität
  - Stoffwechsel- und Ernährungsstörungen
  - Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen
  - Erkrankungen des Nervensystems
  - Psychiatrische Erkrankungen
  - Erkrankungen der Nieren und Harnwege
  - Erkrankungen der Haut und des Unterhautzellgewebes
- Knochenmineraldichte zu Woche 48
- Knochenspezifische Biomarker
  - Alkalische Phosphatase (BSAP)
  - Prokollagen 1-N-terminales Propeptid (P1NP)

- CTx
- Osteocalcin

Studientypen:

Es wurden zwei randomisierte, kontrollierte Studien (RCT) zur Bewertung des Zusatznutzens eingeschlossen: SWORD-1 und SWORD-2.

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. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister erfolgen, während ein Kongressabstract allein in der Regel nicht für eine Studienbewertung ausreicht.

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.

Das Ausmaß des Zusatznutzens von Dolutegravir/ Rilpivirin wird anhand der Studien dargestellt, die die Kriterien in der nachfolgenden Tabelle 4-5 erfüllen. Auf Grundlage dieser festgelegtenKriterien erfolgte der Ein- bzw. Ausschluss von Studien. Im Anschluss an Tabelle 4-5 werden die einzelnen Kriterien begründet.

Tabelle 4-5: Ein- und Ausschlusskriterien bei der Studienselektion

	Einschlusskriterien	**E**	**Ausschlusskriterien**	**A**
Patientenpopulation	Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen.	E1	Von E1 abweichende Patientenpopulation oder andere Altersgruppe der unter E1 genannten Patientenpopulation	A1
Intervention	Dolutegravir/Rilpivirin 50 mg/25 mg einmal täglich	E2	andere Intervention bzw. abweichende Dosierung	A2
Vergleichstherapie	individuelle antiretrovirale Therapie in Abhängigkeit der Vortherapie(n) und ggf. unter Berücksichtigung von Nebenwirkungen (bei Patienten ohne Umstellindikation wird diese Vergleichstherapie durch Fortführung ihrer bisherigen Therapie realisiert)	E3	andere Vergleichstherapie	A3
Endpunkte	Ergebnisse zu mind. einem patientenrelevanten Endpunkt bzgl. • Morbidität • Lebensqualität • Mortalität • Sicherheit/Verträglichkeit sind berichtet	E4	Ergebnisse zu keinem der nebenstehend genannten Endpunkte sind berichtet	A4
Studientyp	randomisierte, kontrollierte Studie (RCT)	E5	nicht randomisierte Studie	A5
Publikationstyp	Primärpublikation oder Sekundärpublikation mit neuen Daten zu den unter E4 genannten Endpunkten oder ausführlicher Studienbericht	E6	Sekundärpublikation ohne neue Daten oder ohne Angaben zu patientenrelevanten Endpunkten	A6
Studiendauer	mindestens 48 Wochen	E7	kürzer als 48 Wochen	A7

Begründung der einzelnen Kriterien:

Patientenpopulation:

Die Patientenpopulation entspricht der Zielpopulation, d.h. der Population, die durch die Zulassung des zu bewertenden Arzneimittels umfasst ist. Dies sind laut Indikation Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen. Zusätzlich wird eine Unterscheidung vorgenommen, ob identifizierte relevante RCT Aussagen für die Teilpopulation mit oder ohne Umstellindikation liefern.

Intervention:

Die Intervention entspricht dem Zulassungsstatus und dem bestimmungsgemäßen Gebrauch des zu bewertenden Arzneimittels für die zugelassene Indikation. Demnach ist die orale Einnahme der Festkombination aus Dolutegravir/Rilpivirin 50 mg/25 mg einmal täglich vorgesehen.

Vergleichstherapie:

Als Vergleichstherapie gilt die zweckmäßige Vergleichstherapie, wie sie vom G-BA festgelegt wurde: Individuelle antiretrovirale Therapie in Abhängigkeit der Vortherapie(n) und ggf. unter Berücksichtigung von Nebenwirkungen (4). Es kommen hierfür alle Wirkstoffe und Wirkstoffkombinationen einer antiretroviralen Therapie entsprechend ihres bestimmungsgemäßen Gebrauchs in Frage, die zur Behandlung einer Infektion mit HIV-1 zugelassen sind. Dies gilt prinzipiell gleichermaßen für beide Teilpopulationen, mit oder ohne Umstellindikation. Für die Teilpopulation ohne Umstellindikation besteht die zweckmäßige Vergleichstherapie in der Fortführung der bisherigen Therapie (4).

Endpunkte:

Die vorliegende Nutzenbewertung basiert auf den Ergebnissen zu patientenrelevanten Endpunkten, welche die Morbidität, Lebensqualität, Mortalität sowie Sicherheit und Verträglichkeit des zu bewertenden Arzneimittels im vorliegenden Indikationsgebiet beschreiben. Eine ausführliche Liste der Endpunkte ist in Abschnitt 4.2.5.2 gegeben.

Studientyp/Publikationstyp:

Es wird ausschließlich nach randomisierten, kontrollierten Studien gesucht, um den höchst möglichen Evidenzgrad zur Beantwortung der Fragestellung heranzuziehen. Auch Substudien von identifizierten RCT werden eingeschlossen, sofern sie einen Beitrag zur Bewertung des Zusatznutzens leisten. Der Fokus bei der Studienselektion liegt auf Primärpublikationen als

Datenquelle. Sekundärpublikationen werden darüber hinaus als Quellen einbezogen, sofern diese zusätzliche Erkenntnisse zu patientenrelevanten Endpunkten liefern.

Studiendauer:

Studien mit einer Dauer von mindestens 48 Wochen werden für das vorliegende Indikationsgebiet als ausreichend angesehen, um insbesondere das Absenken der Viruslast zur Beurteilung der Wirksamkeit adäquat abbilden zu können. Dies entspricht auch der Vorgabe in den Leitlinen der europäischen Zulassungsbehörde. (8)

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 erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ 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. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Am 03.04.2018 wurde eine systematische Literaturrecherche zur Suche nach Studien mit dem zu bewertenden Arzneimittel im entsprechenden Indikationsgebiet durchgeführt. Hierfür wurde in den folgenden Datenbanken recherchiert:

- MEDLINE
- EMBASE
- *Cochrane Central Register of Controlled Trials*

In den Datenbanken MEDLINE und EMBASE wurde jeweils ein validierter Filter (9) für die Suche nach RCT angewendet.

Es wurde in jeder Datenbank einzeln und mit einer jeweils adaptierten Suchstrategie gesucht. Die detaillierten Suchstrategien pro Datenbank sind in Anhang 4-A hinterlegt.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern immer dann

durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), Klinische Prüfungen PharmNet.Bund (http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: http://apps.who.int/trialsearch/) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

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. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

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

Am 03.04.2018 erfolgte eine Recherche in den Studienregistern zur Suche nach Studien mit dem zu bewertenden Arzneimittel im entsprechenden Indikationsgebiet. Diese wurde in den folgenden Studienregistern durchgeführt:

- ClinicalTrials.gov
- *International Clinical Trials Registry Platform Search Portal* der WHO (ICTRP)
- *EU Clinical Trials Register* (EU-CTR)
- PharmNet.Bund.

Die jeweiligen Suchstrategien sind in Anhang 4-B hinterlegt.

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 Rechercheschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die aus der systematischen Literaturrecherche identifizierten Treffer wurden unabhängig von zwei Personen anhand der Angaben in Titel und Abstract bewertet, auf ihre Relevanz für die vorliegende Fragestellung (siehe Ein- und Ausschlusskriterien der Studienrecherche) hin beurteilt und als Volltext beschafft. Im Falle von abweichenden Bewertungen der beiden Reviewer wurden die fraglichen Treffer wiederholt bewertet und im Zweifelsfall als relevant für die Beschaffung der Volltexte eingestuft.

Nach Vorliegen der Volltexte wurden diese erneut unabhängig von zwei Personen bezüglich ihrer Relevanz durchgesehen und für die Nutzenbewertung herangezogen, falls sie als relevant eingestuft wurden. Eine Liste der im Volltext gesichteten und ausgeschlossenen Studien ist mit Angabe des spezifischen Ausschlussgrunds in Anhang 4-C aufgeführt.

Dasselbe Vorgehen wurde für die Treffer aus den Studienregistern angewendet. Diese zusätzlichen Quellen wurden in einem zweiten Schritt ebenfalls – sofern verfügbar – gesichtet und auf vorhandene Ergebnisdarstellungen in den Studienregistern untersucht. Eine Liste der ausgeschlossenen Studien ist mit Angabe des spezifischen Ausschlussgrunds in Anhang 4-D aufgeführt.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

<u>A: Verzerrungsaspekte der Ergebnisse auf Studienebene</u>

– 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 der behandelnden Personen
– ergebnisgesteuerte Berichterstattung
– sonstige Aspekte

<u>B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene</u>

– 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 der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Analyse des Verzerrungspotenzials erfolgte anhand der Kriterien der Bewertungsbögen zur Einschätzung von Verzerrungsaspekten gemäß Anhang 4-F.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)[1]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer

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

Beobachtungsstudien soll mindestens den Anforderungen des TREND-[2] bzw. STROBE-Statements[3] 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.

Die Beschreibung des Designs und der Methodik der eingeschlossenen randomisierten, kontrollierten Studien erfolgte gemäß den Vorgaben des CONSORT-Statements und ist in Anhang 4-E 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 patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. 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). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

[2] Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.
[3] Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

Charakteristika der Studienpopulationen:

Die Patientencharakteristika werden anhand demographischer Angaben sowie der Krankheitslast jeweils zu Studienbeginn beschrieben. Folgende Charakteristika werden dargestellt:

- Geschlecht
- Alter
- Abstammung
- Ethnie
- Region
- Baseline-Viruslast
- Baseline CD4+ Zellzahl
- CDC Kategorie zu Studienbeginn
- bestehende ART zum Zeitpunkt des Screenings
- Therapiedauer vor Studienbeginn

Zusätzlich werden die Studienpopulationen dahingehend klassifiziert, ob es sich um Patienten mit oder ohne Umstellindikation handelt.

Ergebnisse zu patientenrelevanten Endpunkten:

Als patientenrelevante Endpunkte werden Ergebnisse zur Morbidität, Lebensqualität, Sicherheit/Verträglichkeit und Mortalität dargestellt. Folgende patientenrelevante Endpunkte werden berichtet:

Morbidität:

- Viruslast < 50 Kopien/ml zu Woche 48

  Patientenrelevanz: Wie schon in vorangegangenen Nutzenbewertungen im Indikationsgebiet HIV belegt, gilt das virologische Ansprechen (Absinken der Viruslast unter 50 Kopien/ml) als patientenrelevanter Endpunkt. Dies wurde vom G-BA in den Beschlüssen zu den Nutzenbewertungen insbesondere der Einzelsubstanz Dolutegravir (10) sowie der antiretroviralen Substanz Rilpivirin (11) und weiteren Bewertungen im vorliegenden Anwendungsgebiet anerkannt (z.B. kürzlich durchgeführte Bewertungen zu Emtricitabin/Tenofoviralafenamid oder Elvitegravir/Cobicistat/Emtricitabin/Tenofoviralafenamid) (12, 13). Das Absinken der Viruslast gilt als allgemein anerkannter Surrogatparameter für das Auftreten von AIDS-definierenden Ereignissen oder Tod (8, 14, 15).

  Beschreibung und Validität: Die klinischen Studien im Anwendungsgebiet der HIV-Infektion sind generell seit Verfügbarkeit von hocheffektiven antiretroviralen Therapien

(HAART) nicht mehr dafür ausgelegt, einen Effekt hinsichtlich verhinderter AIDS-definierender Ereignisse zu belegen. Hierfür sind die Ereignishäufigkeiten bei diesem langsam fortschreitenden Krankheitsverlauf im Rahmen einer Studiendauer von beispielsweise 48 Wochen nicht ausreichend. Insbesondere im für die hier zu bewertende Indikation vorliegenden Studientyp treten im eingeschlossenen Patientenkollektiv AIDS-definierende Ereignisse und/oder Todesfälle besonders selten auf. Der in klinischen Studien erhobene Endpunkt „Viruslast < 50 Kopien/ml“ besitzt einen deutlich höheren prädiktiven Wert bzgl. der Verhinderung von AIDS-definierenden Ereignissen und Todesfällen, da er über die unmittelbare Studiendauer hinaus validierte Vorhersagen erlaubt. Die Ergebnisse zur Auftretenshäufigkeit von AIDS-definierender Ereignisse bzw. Todesfällen in den dargestellten Studien werden im Anschluss an die Ausführungen zum Endpunkt Viruslast summarisch aufgeführt. Auf Einzelstudienbasis liegen alle Details zusätzlich in den Studienberichten vor.

In der Darstellung der Ereignisse und der Einschätzung des Zusatznutzens wird im vorliegenden Nutzendossier somit nicht zusätzlich auf den Endpunkt AIDS-definierende Ereignisse/Tod eingegangen, da er nach Auffassung des pharmazeutischen Unternehmers angesichts der hier relevanten Therapiesituation und der sich daraus ergebenden Patientencharakteristika wegen der Seltenheit des Auftretens keinen relevanten zusätzlichen, über die Viruslast hinausgehenden Beitrag zur Nutzenbewertung liefert.

Die CD4+ Zellzahl ist ein gut validierter Surrogatendpunkt für den Immunstatus des Patienten. Da die zusätzliche Bedeutung dieses Parameters über die Viruslast hinaus für den patientenrelevanten Zusatznutzen von Dolutegravir/Rilpivirin im Indikationsgebiet als eingeschränkt betrachtet wird, wird dieser Parameter nicht separat dargestellt. Die Ergebnisse wurden jedoch in den dargestellten Studien erfasst und liegen auf Basis der Einzelstudien durch die angefügten Studienberichte vor.

Lebensqualität:

- EQ-5D

Patientenrelevanz: Der EQ-5D ist unmittelbar patientenrelevant, da er vom Patienten selbst berichtet wird und dessen Lebensqualität widerspiegelt.

Beschreibung und Validität: Der EQ-5D ist ein allgemein einsetzbares Standardinstrument zur Evaluierung der gesundheitsbezogenen Lebensqualität (http://www.euroqol.org). Die Lebensqualität wird anhand von fünf Dimensionen (Mobilität, Fähigkeit, für sich selbst zu sorgen, allgemeine Tätigkeiten, Schmerzen/körperliche Beschwerden, Angst/Depression) abgefragt. Die Lebensqualität wird durch einen gewichteten Summenscore der fünf Dimensionen (*Utility Score*) (16) und mittels einer visuellen Analogskala (VAS), die den gesamtheitlichen Gesundheitszustand misst (17), abgebildet.

- Symptom Distress Module (SDM)

  Patientenrelevanz: Unmittelbar patientenrelevant, da vom Patienten selbst berichtet und dessen Lebensqualität betreffend; kein Surrogatendpunkt. Das SDM wurde bereits in früheren Nutzenbewertungen im vorliegenden Anwendungsgebiet als patientenrelevant eingestuft (18, 19).

  Beschreibung und Validität: Das Symptom Distress Module (auch als HIV Symptom Index bezeichnet) ist ein Messinstrument, das die Symptomatik spezifisch für HIV und dessen Therapie anhand von 20 Items abbildet. Der Score beschreibt somit die indikationsspezifische Belastung durch die Symptomatik, die durch die Erkrankung oder die Therapie verursacht wird. Das SDM besteht insgesamt aus zwei Scores: Der Symptom Count, der die Anzahl der vorhandenen Symptome angibt und der Symptom Bother Score, mittels dessen der Patient seine Belastung durch die jeweiligen Symptome einschätzt. Der Symptom Bother Score ist somit im Kern eine Selbsteinschätzung der gesundheitsbezogenen Lebensqualität des Patienten. Die Entwicklung des Scores wurde in Abstimmung mit der *AIDS Clinical Trials Group of the US National Institute of Allergy and Infectious Diseases* vorgenommen. Das SDM ist validiert und Assoziationen mit physischen und psychischen gesundheitsbezogenen Scores zur Lebensqualität wurden berichtet (20, 21).

  Das SDM wurde in zurückliegenden Nutzenbewertungen alternativ der Morbidität zugeordnet. Diese Interpretation ist hinsichtlich des Symptom Counts eher begründbar. Wegen des in vergangenen Nutzenbewertungen stärkeren Gewichts, das der Interpretation des Symtom Bother Scores zukam, wurde das SDM hier im Dossier als der Lebensqualität zurgehörig eingestuft.

- HIVTSQ (*HIV Treatment Satisfaction Questionnaire*)

  Patientenrelevanz: Der HIVTSQ ist unmittelbar patientenrelevant, da er vom Patienten selbst berichtet wird und somit die Patientenzufriedenheit mit der Therapie widerspiegelt.

  Beschreibung und Validität: Es handelt sich hierbei um einen validierten Score, der die generelle Patientenzufriedenheit mit der Behandlung abfragt (22). In dieser Nutzenbewertung wird der *Total Score* dargestellt, der sich aus den Antworten zu zehn Fragen (u.a. zu Anwendungsfreundlichkeit, Flexibilität) zusammensetzt.

Sicherheit/Verträglichkeit:

Patientenrelevanz: Alle Endpunkte zur Arzneimittelsicherheit sind unmittelbar patientenrelevant und ermöglichen im Zusammenhang mit den Wirksamkeitsendpunkten die Nutzen-Schaden-Abwägung einer Therapie.

- Mortalität bis Woche 48
- Unerwünschte Ereignisse bis Woche 48
- Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48
- Schwerwiegende unerwünschte Ereignisse bis Woche 48
- Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten
- Unerwünschte Ereignisse von speziellem Interesse bis Woche 48

Die folgenden Ereignisse von speziellem Interesse bzgl. der Sicherheit/Verträglichkeit wurden ausgewählt, um einerseits das Sicherheits-/Verträglichkeitsprofil von Dolutegravir/Rilpivirin abbilden zu können und andererseits relevante Ereigniskategorien der zweckmäßigen Vergleichstherapie berücksichtigen zu können:

- Herzerkrankungen
- Erkrankungen des Gastrointestinaltrakts
- Leber- und Gallenerkrankungen
- Hypersensitivität
- Stoffwechsel- und Ernährungsstörungen
- Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen
- Erkrankungen des Nervensystems
- Psychiatrische Erkrankungen
- Erkrankungen der Nieren und Harnwege
- Erkrankungen der Haut und des Unterhautzellgewebes

Knochenmineraldichte:

Die Knochenmineraldichte (*bone mineral density*, BMD) als Surrogat für Knochenfrakturen ist patientenrelevant. Die Knochenmineraldichte wird an der Hüfte und an den Lendenwirbeln bestimmt. Es besteht ein Zusammenhang zwischen verminderter Knochenmineraldichte und inzidenter Frakturen bei mit HIV-infizierten Patienten (23).

Knochenspezifische Biomarker:

Ergebnisse zu den folgenden Biomarkern werden dargestellt:

- alkalische Phosphatase (BSAP)
- Prokollagen 1-N-terminales Propeptid (P1NP)
- type 1 collagen cross-linked C-telopeptide (CTx)
- Osteocalcin

Alle diese Biomarker sind mit einem erhöhten Risiko des Knochenumsatzes assoziiert, welcher seinerseits mit einem erhöhten Risiko des Verlusts an Knochenmineraldichte in Zusammenhang steht.

Darstellung der Effektschätzer:

Sowohl innerhalb als auch zwischen den beiden eingeschlossenen Studien werden einheitliche Operationalisierungen zur Darstellung der Ergebnisse verwendet. Die Definition der Operationalisierung für die einzelnen Endpunkte pro Studie befindet sich jeweils in den entsprechenden Abschnitten der Ergebnisdarstellung in Abschnitt 4.3.1.

Als Effektschätzer auf Studienebene wird für dichotome Endpunkte das relative Risiko inklusive des zugehörigen 95%-Konfidenzintervalls und des p-Werts berichtet. Zudem wird das Odds Ratio sowie die absolute Risikoreduktion (inkl. 95%-Konfidenzintervall und p-Wert) angegeben. Die Interpretation der Ergebnisse findet primär anhand der relativen Risiken statt.

Kontinuierliche Endpunkte, die jeweils zu Studienbeginn und -ende erhoben wurden, für die geeignete validierte Relevanzschwellen auf Gruppenebene vorliegen, werden durch Mittelwertsdifferenzen (und deren 95%-Konfidenzintervalle) beschrieben. Im Falle des Nichtvorhandenseins von geeigneten Relevanzschwellen auf Gruppenebene werden für solche Endpunkte die Ergebnisse anstatt mittels der Mittelwertsdifferenzen mit Hilfe des standardisierten Hedges‘ g (und dessen 95%-Konfidenzintervall) als Schätzer für die Effektstärke dargestellt. Kontinuierliche Endpunkte werden weitestgehend mittels Kovarianzanalysen (*Analysis of Covariance*, ANCOVA) ausgewertet. Welche Methodik im Einzelnen jeweils angewendet wird und nach welchen Variablen im Einzelfall adjustiert wird, ist in den jeweiligen Operationalisierungen angegeben.

4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. 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[4] 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

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

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 sowohl anhand des Odds Ratios als auch des Relativen Risikos 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[5] werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest-Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität[6] 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.

Für jeden Endpunkt werden studienübergreifend Meta-Analysen durchgeführt, sofern zu mehr als einer RCT Ergebnisse vorliegen und diese eine zumindest weitgehend identische Operationalisierung aufweisen. Ergebnisse zu dichotomen Endpunkten werden in den Meta-Analysen mittels relativer Risiken und den zugehörigen 95%-Konfidenzintervallen angegeben und interpretiert. Kontinuierliche Endpunkte werden – sofern geeignete Relevanzschwellen vorliegen – durch Mittelwertsdifferenzen (inklusive deren 95%-Konfidenzintervalle) meta-analytisch zusammengefasst. Bei Fehlen geeigneter Relevanzschwellen werden die Ergebnisse der Meta-Analysen mittels des standardisierten Effektmaßes Hedges' g (und dessen 95%-Konfidenzintervall) dargestellt.

Die Meta-Analysen werden sowohl mit festen als auch mit zufälligen Effekten berechnet. Die Modelle mit festen Effekten folgen der gängigen Inversen-Varianz Methodik, Modelle mit zufälligen Effekten der Methodik von DerSimonian & Laird (24). Die Analyse erfolgte mit der Software *Comprehensive Meta Analysis* (Biostat™, USA), Version 2.2.064. In Fällen, in denen in einer Behandlungsgruppe keinerlei Ereignisse bezüglich des jeweiligen Endpunkts aufgetreten sind, wird 0,5 als Kontinuitätskorrektur für die betreffenden Zellen addiert. Liegen

[5] *Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2008. S. 243-296.*
[6] *Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.*

für alle Studienarme keine Ereignisse vor, wird für den entsprechenden Endpunkt keine Meta-Analyse berechnet.

Als Basis für die Berechnung der Meta-Analysen werden die adjustierten Effektschätzer und Konfidenzintervallgrenzen herangezogen, die aus den eigens für diese Nutzenbewertung angestellten Zusatzanalysen resultieren. In den wenigen Fällen, in denen kein relatives Risiko durch die angewandten, statistischen Modelle berechnet werden konnte, werden die puren Ereigniszahlen (ohne Adjustierung, ggf. mit Kontinuitätskorrektur) zur Berechnung der Meta-Analyse verwendet. Abweichungen zwischen den Schätzern auf Studienebene zwischen den Forest-Plots und den Zusatzanalysen können sich durch gerundete Angaben ergeben, die bei der Berechnung der Meta-Analysen verwendet werden und als Folgeerscheinung unvermeidbar sind.

Das Ausmaß der statistischen Heterogenität wird anhand der I^2-Statistik beurteilt. Hierbei wird ein $I^2 \leq 50$ % als gering, > 50 – 70 % als mäßig und ≥ 70 % als hoch heterogen interpretiert (25).

Gemäß den Vorgaben des Methodenpapiers 5.0 des IQWiG (26) werden bei Vorliegen von wenigen Studien (hier n=2), die zudem homogen bzgl. Studiendesign, Patientenkollektiv und Methodik sind, zur Interpretation des Zusatznutzens die Meta-Analysen mit festen Effekten herangezogen. Im Falle von starker Heterogenität wird zusätzlich zur Berechnung nach DerSimonian & Laird, das Konfidenzintervall nach der Knapp-Hartung-Methode (27) angegeben, welches sich jedoch nur für die nicht-adjustierten relativen Risiken berechnen lässt. Dies ist dem Umstand geschuldet, dass nötige Detailangaben in den Ergebnissen der Zusatzanalysen für die adjustierten Effektschätzer nicht vorliegen. Ergebnisse der Knapp-Hartung-Methode wurden mittels der Software R (https://www.R-project.org/), Version 3.4.1, und dem zugehörigen Programm-Paket „metafor“ berechnet.

Für die geschilderte Situation (Vorliegen von zwei homogenen und spezifisch dafür geplanten RCT) erfolgt in der Situation von Meta-Analysen mit hoher Heterogenität die Interpretation nicht auf Basis der Meta-Analyse, sondern durch eine zusammenfassende Diskussion der Evidenz beider Einzelstudien. Methodisch ist es angezeigt, den eventuellen Übergang von der Betrachtung einer Meta-Analyse mit festen Effekten zu einer mit zufälligen Effekten nicht von der konkreten Heterogenität, sondern von der Studiensituation abhängig zu machen. Zusätzlich ist es jedoch aus methodischen Überlegungen heraus ebenfalls unumstritten, dass die Interpretation einer Meta-Analyse mit festen Effekten bei hoher Heterogenität nicht adäquat ist. (26)

Mittels Forest-Plots werden die Ergebnisse der Meta-Analysen grafisch dargestellt. In diesen Darstellungen sind sowohl die Einzelstudienergebnisse als auch die gepoolten Schätzer (mit festen und zufälligen Effekten) abgebildet.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. 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.

Für den Endpunkt Viruslast < 50 Kopien/ml zu Woche 48 werden zwei Sensitivitätsanalysen zur primären Studienauswertung mittels des Snapshot Algorithmus berechnet und berichtet.

Sensitivitätsanalyse 1:

Patienten, die zu Woche 48 fehlende Werte für die Viruslast haben, werden aus der Analyse ausgeschlossen. Als Effektschätzer werden adjustierte (nach Alter und *Baseline Third Agent*) relative Risiken, Odds Ratios und Risikodifferenzen inkl. der zugehörigen 95%-Konfidenzintervalle sowie p-Werte angegeben.

Sensitivitätsanalyse 2:

Für Patienten, die keine Angabe zur Viruslast zu Woche 48 haben oder die die Studie vorzeitig beendet hatten (außer aufgrund von unerwünschten Ereignissen oder Tod) und die einen Viruslastwert < 50 Kopien/ml zu Woche 36 dokumentiert hatten, werden Werte entsprechend der Responderraten (ohne diese Patienten) in den jeweiligen Behandlungsarmen imputiert. Die Varianz berechnet sich hierfür nach der *data-set re-sizing* Methode W3 basierend auf Higgins et al. 2008 (28). Als Effektschätzer werden relative Risiken, Odds Ratios und Risikodifferenzen inkl. der zugehörigen 95%-Konfidenzintervalle sowie p-Werte angegeben.

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 sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

– Geschlecht
– Alter
– Krankheitsschwere bzw. –stadium
– Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen 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.

Folgende Subgruppenfaktoren wurden analysiert:

- Alter (< 50 Jahre / ≥ 50 Jahre)
- Geschlecht (männlich / weiblich)
- Ethnizität (weiß / nicht-weiß)
- Region (Nordamerika / Europa / Sonstige)
- Baseline CD4+ Zellzahl (< 500 / ≥ 500 Zellen pro mm^3)
- *Baseline Third Agent* (PI / NNRTI / INI)

Diese Subgruppen waren mit Ausnahme der Regionen *a priori* im statistischen Analyseplan festgelegt, „Alter“ und „*Baseline Third Agent*“ waren zudem stratifizierende Randomisierungsvariablen. Der Subgruppenfaktor „Region“ wurde gemäß den Dossieranforderungen zusätzlich mit in die Analysen aufgenommen. Die Baseline CD4+ Zellzahl gilt als Faktor für den Schweregrad der Erkrankung.

Der Trennpunkt des quantitativen Merkmals „Alter“ (50 Jahre) wurde ebenfalls *a priori* im statistischen Analyseplan festgelegt. Für Endpunkte, für die keine Subgruppenanalysen *a priori* geplant waren und die nicht im Studienbericht enthalten waren, wurden die Subgruppenanalysen nachträglich im Rahmen von Zusatzanalysen für die Dossiererstellung

durchgeführt. Für Endpunkte mit sehr seltenen Ereignissen (siehe Allgemeine Methoden 5.0 des IQWiG, (26)) werden keine Subgruppenergebnisse berichtet.

Interaktionstests werden studienübergreifend mittels des p-Werts der Q-Statistik über die Meta-Analysen berechnet. Dies erfolgte mit Hilfe der Software *Comprehensive Meta Analysis* (Biostat™, USA), Version 2.2.064.

Subgruppenergebnisse werden in diesem Dossier dargestellt, falls im entsprechenden Fall ein Beleg für Effektmodifikation vorliegt, d.h. der Interaktionstest einen p-Wert von < 0,05 zeigt (26). Die p-Werte der Interaktionstest für jeden untersuchten Subgruppenfaktor (pro Endpunkt) werden jedoch vollständig im Abschnitt 4.3.1.3.25 angegeben.

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

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist[11].

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in

[7] *Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. Health Technol Assess 2005; 9(26): 1-148.*
[8] *Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.*
[9] *Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005; 331(7521): 897-900.*
[10] *Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. Stat Methods Med Res 2008;17(3): 279-301.*
[11] *B. Schöttker, D. Lühmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

Frage kommen, vorliegen. Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen[12].

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

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

Es wurde kein indirekter Vergleich durchgeführt.

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

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*

- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, 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“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

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

Studie	**Zulassungsstudie (ja/nein)**	**Sponsor (ja/nein)**	**Status (abgeschlossen / abgebrochen / laufend)**	**Studiendauer**	**Therapiearme**
201636 SWORD-1	ja	ja	laufend	148 Wochen	Dolutegravir/Rilpivirin, Fortführen der bisherigen Therapie (*Current Antiretroviral Regimen*, CAR)
201637 SWORD-2	ja	ja	laufend	148 Wochen	Dolutegravir/Rilpivirin, Fortführen der bisherigen Therapie (*Current Antiretroviral Regimen*, CAR)
202094[a] DEXA Substudie	ja	ja	laufend	148 Wochen	Dolutegravir/Rilpivirin, Fortführen der bisherigen Therapie (*Current Antiretroviral Regimen*, CAR)
[a] Keine RCT im eigentlichen Sinne. In dieser Substudie werden bereits in die beiden Zulassungsstudien (201636 und 201637) randomisierte Patienten eingeschlossen, die zusätzliche bestimmte Kriterien erfüllen (insbesondere das Vorhandensein von Tenofovirdisoproxilfumarat als ART-Komponente). Da sie auch im Zulassungsantrag als eigenständige Studie geführt wird, wird sie hier extra aufgeführt, obwohl es für die Fragestellung keine eigenständige RCT ist.					

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

Obige Tabelle gibt den Studienstatus zum 03.04.2018 wider.

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

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

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
-	-

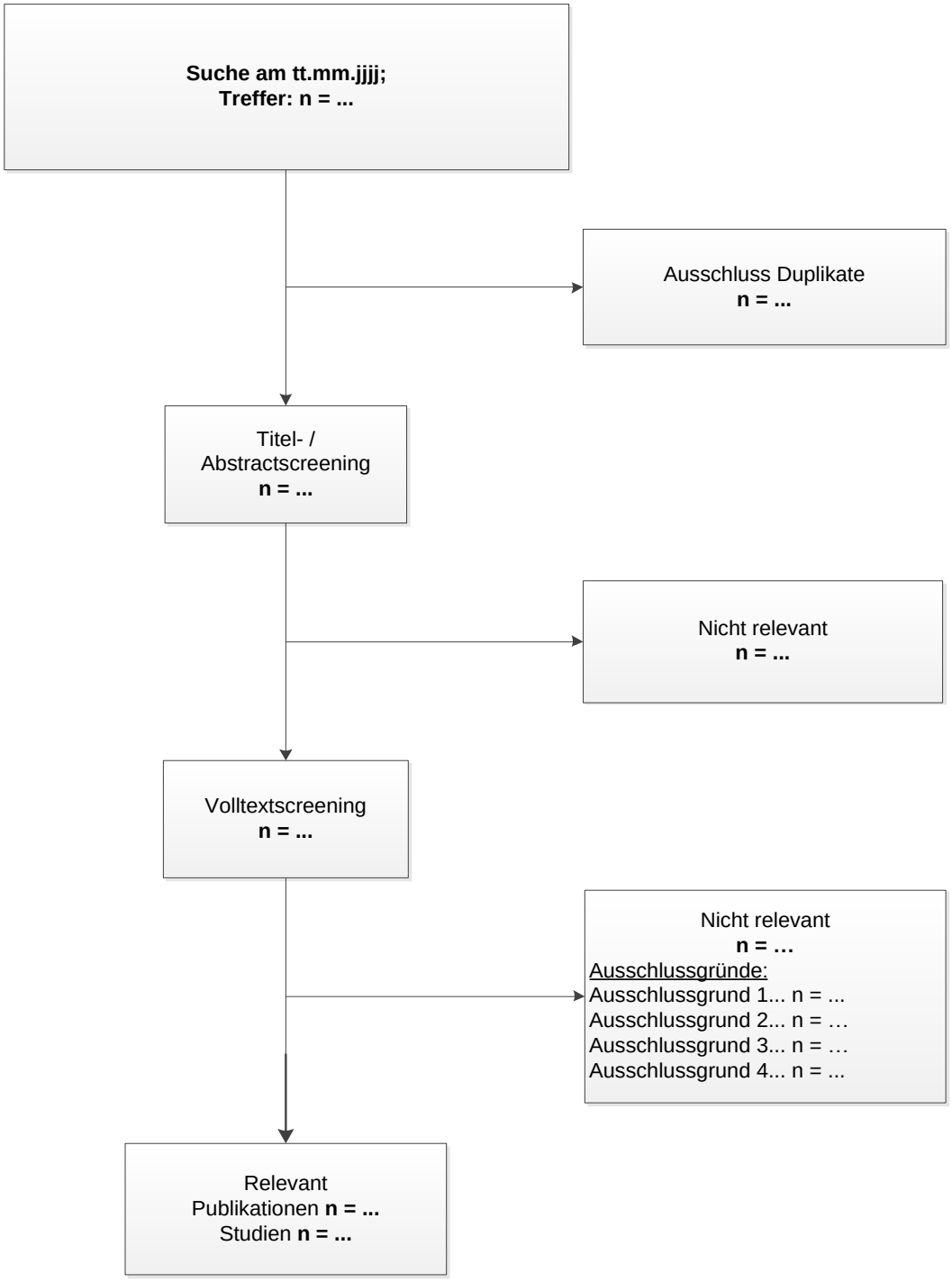
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. 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 Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.

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

Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

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


Suche am tt.mm.jjjj;
Treffer: n = ...
Ausschluss Duplikate
n = ...
Titel- / Abstractscreening
n = ...
Nicht relevant
n = ...
Volltextscreening
n = ...
Nicht relevant
n = …
Ausschlussgründe:
Ausschlussgrund 1... n = ...
Ausschlussgrund 2... n = …
Ausschlussgrund 3... n = …
Ausschlussgrund 4... n = ...
Relevant
Publikationen n = ...
Studien n = ...

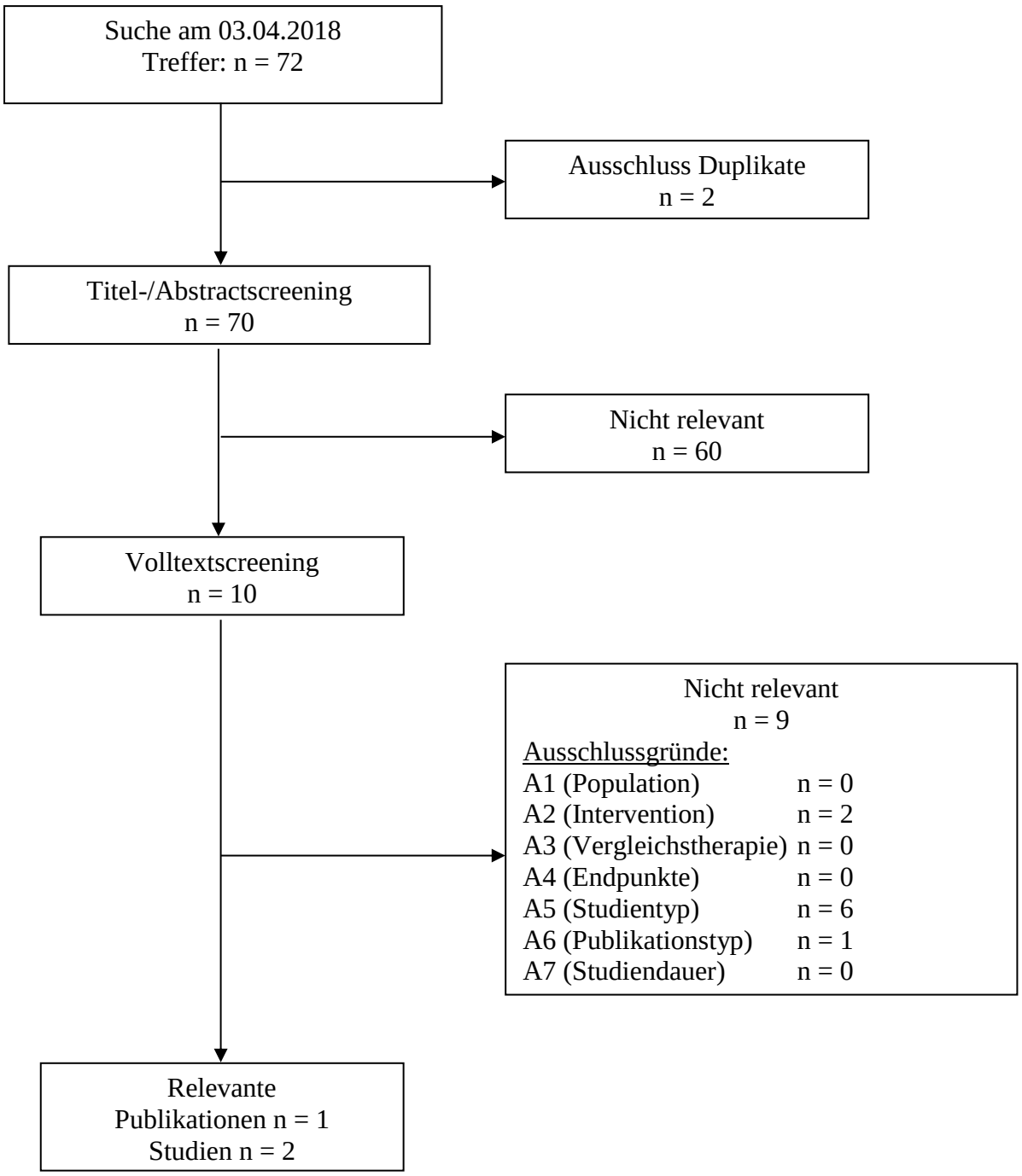


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern

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

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

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

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

Studie	**Identifikationsorte (Name des Studienregisters und Angabe der Zitate[a])**	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)**	**Status (abgeschlossen/ abgebrochen/ laufend)**
201636 SWORD-1	clinicaltrials.gov [NCT02429791] (29) ICTRP (30, 31) EUCTR (32) Pharmnet.Bund (33)	ja	ja (34)	laufend
201637 SWORD-2	clinicaltrials.gov [NCT02422797] (35) ICTRP (36, 37) EUCTR (38) Pharmnet.Bund (39)	ja	ja (34)	laufend
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.				

Geben Sie an, welchen Stand die Information in Tabelle 4-8 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Recherche in den Studienregistern wurde am 03.04.2018 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 (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

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

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

Studie	**Studienkategorie**			**verfügbare Quellen[a]**		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	**gesponserte Studie[b] (ja/nein)**	**Studie Dritter (ja/nein)**	**Studien-bericht (ja/nein [Zitat])**	**Register-eintrag[c] (ja/nein [Zitat])**	**Publikation (ja/nein [Zitat])**
201636 SWORD-1	ja	ja	nein	ja (40)	ja (29-33)	(34)
201637 SWORD-2	ja	ja	nein	ja (41)	ja (35-39)	(34)
202094[d] DEXA Substudie	ja	ja	nein	ja (42)	ja (43, 44)	(45)

a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.7 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden.

b: Studie, für die der Unternehmer Sponsor war.

c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

d: Die DEXA Substudie wird hier der Vollständigkeit halber mit aufgeführt. Es handelt sich nicht um eine RCT im eigentlichen Sinn, sondern um eine Substudie, in die Patienten der beiden Zulassungsstudien (201636 und 201637) eingeschlossen wurden.

Die Patientenpopulationen aller in Tabelle 4-9 aufgeführten Studien erfüllen die Kriterien des G-BA für die **Teilpopulation ohne Umstellindikation.**

Für die **Teilpopulation mit Umstellindikation** (nach den G-BA-Kriterien) konnte keine RCT-Evidenz identifiziert werden. Somit ist für diese Teilpopulation keine evidenzbasierte Darstellung zur Ableitung eines Zusatznutzens möglich.

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-E zu hinterlegen.

In diesem Abschnitt werden die beiden Zulassungsstudien SWORD-1 und SWORD-2 beschrieben.

Eine Beschreibung der DEXA-Substudie und die Charakterisierung des in dieser Studie eingeschlossenen Patientenkollektivs erfolgen in Abschnitt 4.3.1.3.20. Dort werden die Ergebnisse zu dem einzigen aus dieser Studie dargestellten Endpunkt, der Knochenmineraldichte, dargestellt. Die Knochenmineraldichte wurde ausschließlich in der DEXA-Substudie erhoben.

Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation)

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**
201636 (SWORD-1)	RCT, offen, parallel, Nichtunterlegenheits-studie	Erwachsene HIV-1 Patienten, die virologisch supprimiert unter ihrer aktuellen ART sind.	DTG + RPV (n=252) CAR (n=256)	Randomisierte Phase (*Early Switch Phase*, Tag 1 bis Woche 52): Auswertung des primären Endpunkts zu Woche 48 *Late Switch Phase*: Woche 52 bis Woche 148 (Wechsel auf bzw. Weiterführung von DTG+RPV für alle Patienten mit Viruslast < 50 Kopien/ml)	66 Zentren in 13 Ländern weltweit: Argentinien, Australien, Belgien, Kanada, Frankreich, Deutschland, Italien, Niederlande, Russland, Spanien, Taiwan, UK und USA seit 04/2015 laufend (Datenbankschluss für 48 Wochen-Auswertung: 11/2016)	Anteil Patienten mit Viruslast <50 Kopien/ml zu 48 Wochen; Krankheitsprogression bzgl. CDC-Kategorien, Resistenzentwicklung, CD4-Zellzahl, Lebensqualität, unerwünschte Ereignisse

201637 (SWORD-2)	RCT, offen, parallel, Nichtunterlegenheits-studie	Erwachsene HIV1-Patienten, die virologisch supprimiert unter ihrer aktuellen ART sind.	DTG + RPV (n=261) CAR (n=255)	Randomisierte Phase (*Early Switch Phase* Tag 1 bis Woche 52): Auswertung des primären Endpunkts zu Woche 48 *Late Switch Phase*: Woche 52 bis Woche 148 (Wechsel auf bzw. Weiterführung von DTG+RPV für alle Patienten mit Viruslast < 50 Kopien/ml)	60 Zentren in 11 Ländern weltweit: Spanien, USA, Kanada, Frankreich, Deutschland, Russland, Taiwan, Argentinien, Australien, Italien, UK. seit 04/2015 laufend (Datenbankschluss für 48 Wochen-Auswertung: 11/2016)	Anteil Patienten mit Viruslast <50 Kopien/ml zu 48 Wochen; Krankheitsprogression bzgl. CDC-Kategorien, Resistenzentwicklung, CD4-Zellzahl, Lebensqualität, unerwünschte Ereignisse
Abkürzungen: DTG = Dolutegravir, RPV = Rilpivirin, CAR = Current Antiretroviral Regime, ART = antiretrovirale Therapie						

Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation)

Studie	**DTG + RPV**	**CAR**	***ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
201636 (SWORD-1)	DTG 50 mg einmal täglich eine Tablette + RPV 25 mg einmal täglich eine Tablette zu einer Mahlzeit	2 NRTI + INI oder 2 NRTI + NNRTI oder 2 NRTI + PI nach den jeweiligen Fachinformationen anzuwenden	Die Patienten, die nicht dem DTG+RPV Arm zugeteilt wurden, führten ihre bestehende Therapie (*Current Antiretroviral Therapie*, CAR) fort.
201637 (SWORD-2)	DTG 50 mg einmal täglich eine Tablette + RPV 25 mg einmal täglich eine Tablette zu einer Mahlzeit	2 NRTI + INI oder 2 NRTI + NNRTI oder 2 NRTI + PI nach den jeweiligen Fachinformationen anzuwenden	Die Patienten, die nicht dem DTG+RPV Arm zugeteilt wurden, führten ihre bestehende Therapie (*Current Antiretroviral Therapie*, CAR) fort.
Abkürzungen: DTG = Dolutegravir, RPV = Rilpivirin, CAR = Current Antiretroviral Regime, NRTI = nukleosidaler Reverse-Transkriptase-Inhibitor, NNRTI = Nicht-Nukleosidaler Reverse-Transkriptase-Inhibitor, PI = Protease-Inhibitor, INI = Integrase-Inhibitor			

Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation) [ITT-E Population]

Studie	**201636 SWORD-1**		**201637 SWORD-2**	
	DTG + RPV N=252	**CAR N=256**	**DTG + RPV N=261**	**CAR N=255**
Geschlecht, n (%)				
männlich	194 (77)	205 (80)	199 (76)	198 (78)
weiblich	58 (23)	51 (20)	62 (24)	57 (22)
Alter [Jahre], Median [Min - Max]	43,0 [23-78]	43,0 [22-76]	43,0 [21-79]	43,0 [22-69]
Alter, n (%)				
< 35 Jahre	53 (21)	60 (23)	64 (25)	55 (22)
35 bis < 50 Jahre	130 (52)	125 (49)	119 (46)	129 (51)
≥ 50 Jahre	69 (27)	71 (28)	78 (30)	71 (28)
Abstammung, n (%)				
hispanisch/lateinamerikanisch	30 (12)	34 (13)	37 (14)	48 (19)
andere	222 (88)	222 (87)	224 (86)	207 (81)
Ethnie, n (%)				
weiß	198 (79)	188 (73)	223 (85)	210 (82)
schwarz/afrikanisch-amerikanisch	24 (10)	27 (11)	13 (5)	20 (8)
asiatisch	25 (10)	34 (13)	13 (5)	16 (6)
andere	5 (2)	7 (3)	12 (5)	9 (4)
Region, n (%)				
Nordamerika	63 (25,0)	51 (19,9)	36 (13,8)	42 (16,5)
Europa	161 (63,9)	164 (64,1)	188 (72,0)	175 (68,6)
andere	28 (11,1)	41 (16,0)	37 (14,2)	38 (14,9)
Baseline Viruslast, n (%)				

< 50 Kopien/ml	247 (98)	253 (99)	259 (99)	251 (98)
≥ 50 Kopien/ml	5 (2)	3 (1)	2 (1)	4 (2)
Baseline CD4+ Zellzahl [pro mm^3], Median [Min-Max]	611 [30-1585]	638 [98-1445]	610 [115-1778]	628 [107-1660]
CDC-Kategorie, n (%)				
A	203 (81)	198 (77)	197 (75)	187 (73)
B	20 (8)	35 (14)	35 (13)	33 (13)
C	29 (12)	23 (9)	29 (11)	34 (13)
unbekannt	0 (0)	0 (0)	0 (0)	1 (<1)
ART zum Screening-Zeitpunkt, n (%)				
NRTI	252 (100)	256 (100)	261 (100)	255 (100)
NNRTI[$]	131 (52)	134 (52)	144 (55)	144 (56)
PI[$]	75 (30)	74 (29)	58 (22)	62 (24)
INI[$]	46 (18)	48 (19)	59 (23)	49 (19)
Therapiedauer vor Studienbeginn [Monate], Median [Min-Max]	49,9 [8-221]	50,9 [9-208]	51,4 [10-205]	54,0 [9-270]

Abkürzungen: DTG = Dolutegravir, RPV = Rilpivirin, CAR = Current Antiretroviral Regime, ART = antiretrovirale Therapie, NRTI = Nukleosidaler Reverse-Transkriptase-Inhibitor, NNRTI = Nicht-Nukleosidaler Reverse-Transkriptase-Inhibitor, PI = Protease-Inhibitor, INI = Integrase-Inhibitor, Min = Minimum, Max = Maximum

[$] jeweils zuzüglich weiterer NRTI

Beschreiben Sie die Studien zusammenfassend. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

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

Die Studienpopulationen sind bzgl. ihrer Patientencharakteristika und ihres Schweregrads zu Studienbeginn sehr gut balanciert zwischen beiden Studien und beiden Therapiearmen. Dies ergibt sich aus dem identischen Studiendesign inkl. Ein- und Ausschlusskriterien der beiden Zulassungsstudien SWORD-1 und SWORD-2. Die Patienten in der CAR-Gruppe setzten ihre seit mindestens 6 Monaten bestehende, effektive (Viruslast < 50 Kopien/ml) und stabil ohne Nebenwirkungen und Toxizitäten verlaufende ART fort, wobei die mediane Dauer der bestehenden Therapie war sogar noch sehr viel größer als die geforderten 6 Monate war. Es handelt sich zudem um Patienten, die überwiegend CDC-Kategorie A aufwiesen und mit der ersten oder zweiten antiretroviralen Kombinationstherapie behandelt wurden, welche ein NNRTI, PI oder INI in Kombination mit 2 NRTI beinhaltet. Es dürfen bei den Studienpatienten bisher noch keine Resistenzen auf ART aufgetreten sein.

Aus den beschriebenen Charakteristika der Studienpopulationen der beiden SWORD-Studien (sowie der DEXA-Substudie, die sich aus den Patienten der SWORD-Studien rekrutiert) ergibt sich, dass die Studien in Tabelle 4-9 (SWORD-1, SWORD-2, DEXA-Substudie) hinsichtlich der Zusatznutzenbewertung ausschließlich und komplett für die **Teilpopulation ohne Umstellindikation** (entsprechend der Definition des G-BA, d.h. Patienten, die nicht wegen virologischen Versagens oder Nebenwirkungen oder Toxizitäten ihre ART wechseln) relevant sind.

Da die Studiendurchführung zum Großteil in Europa und Nordamerika stattfand und da das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht wurden, ist von einer uneingeschränkten Übertragbarkeit auf den deutschen Versorgungskontext auszugehen.

Es liegen mittlerweile auch die Studienberichte für die Verlängerungsphase der beiden SWORD-Studien nach 100 Wochen Behandlung vor. Da zwischen Woche 48 und Woche 52 jedoch alle Patienten der CAR-Arme auf die Kombinationstherapie Dolutegravir + Rilpivirin umgestellt wurden (die Patienten der DTG/RPV-Arme werden unverändert weiter behandelt), trägt diese einarmige „*late switch*“ Phase keine zusätzlichen direkt vergleichenden Daten im Rahmen einer RCT-Bewertung bei. In diesem Dossier werden daher ausschließlich die Ergebnisse der Auswertung nach 48 Wochen dargestellt und zur Beurteilung des Zusatznutzens herangezogen.

Für die **Teilpopulation** der Zielpopulation der **Patienten mit Umstellindikation** (lt. Definition des G-BA) liegen hingegen keine RCT vor, so dass für diese Teilpopulation keine Evidenzbewertung möglich ist.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

Tabelle 4-13: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation)

Studie	Adäquate Erzeugung der Randomisierungs-sequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
SWORD-1	ja	ja	nein	nein	ja	ja	niedrig
SWORD-2	ja	ja	nein	nein	ja	ja	niedrig
DEXA-Substudie	ja	ja	nein	nein	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei beiden Studien handelt es sich um randomisierte, kontrollierte Studien. Die Zuordnung der Patienten zu den Behandlungsgruppen erfolgte jeweils mittels eines zuvor generierten zufälligen Randomisierungscodes. Beide Studien waren nicht verblindet, sowohl der behandelnde Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation. Des Weiteren existieren keine Hinweise auf eine ergebnisabhängige Berichterstattung sowie keine sonstigen Aspekte, die für ein zusätzliches hohes Verzerrungspotenzial ursächlich wären.

Die offene Studiendurchführung allein führt nicht dazu, dass die Studien generell als mit hohem Verzerrungspotenzial behaftet eingestuft werden sollten. Sofern sich dies auf Endpunktebene anders verhält, wird dies individuell bei den entsprechenden Endpunkten sowie in Anhang 4-F jeweils diskutiert.

Die Substudie besitzt, basierend auf der Einstufung der beiden Hauptstudien SWORD-1 und SWORD-2, auf Studienebene ebenfalls niedriges Verzerrungspotenzial.

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 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel - Teilpopulation ohne Umstellindikation

Endpunkt	**untersucht in SWORD-1**	**untersucht in SWORD-2**
Viruslast < 50 Kopien/ml zu Woche 48	ja	ja
EQ-5D zu Woche 48	ja	ja
Symptom Distress Module zu Woche 48	ja	ja
HIVTSQ Total Score zu Woche 48	ja	ja
Mortalität bis Woche 48		
Unerwünschte Ereignisse bis Woche 48	ja	ja
Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48	ja	ja
Schwerwiegende unerwünschte Ereignisse bis Woche 48	ja	ja
Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege	ja	ja
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes	ja	ja
Knochenmineraldichte	im Rahmen der DEXA Substudie untersucht	

Knochenspezifischer Biomarker: BSAP zu Woche 48	ja	ja
Knochenspezifischer Biomarker: P1NP zu Woche 48	ja	ja
Knochenspezifischer Biomarker: CTx zu Woche 48	ja	ja
Knochenspezifischer Biomarker: Osteocalcin zu Woche 48	ja	ja

Alle folgenden Darstellungen der Evidenz für die in Tabelle 4-14 genannten Endpunkte sind für die **Teilpopulation ohne Umstellindikation** relevant. Dies wird der besseren Lesbarkeit halber nicht bei jedem Endpunkt (und bei den dazugehörigen Subgruppenanalysen ebenfalls nicht) nochmals explizit erwähnt.

Da für die **Teilpopulation mit Umstellindikation** keine RCT-Evidenz identifiziert werden konnte, erfolgt auch keine Evidenzdarstellung.

4.3.1.3.1 Viruslast < 50 Kopien/ml zu Woche 48 – 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 sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler
– Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden.

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen 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). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

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

Tabelle 4-15: Operationalisierung von Viruslast < 50 Kopien/ml zu Woche 48

Studie	Operationalisierung
SWORD-1	Die Responderrate zu Woche 48 wurde nach dem Snapshot-Algorithmus (der FDA) bestimmt. Die Klassifizierung Responder (virologischer Erfolg) / Non-Responder (virologisches Versagen) erfolgt durch den letzten verfügbaren Wert der Viruslast im jeweiligen Studienintervall (< 50 Kopien/ml oder ≥ 50 Kopien/ml), während der Patient unter Behandlung steht. Liegen in einem Studienintervall keine Daten zur Viruslast vor, kann ein Patient nicht als Responder eingestuft werden. Abhängig vom Grund des Fehlens des Werts wird der Patient als Non-Responder bzw. als Missing (z.B. aufgrund unerwünschter Ereignisse) gezählt. Patienten, die aus anderen Gründen als wegen unerwünschter Ereignisse die Studie abbrechen und deren Viruslast zu dem Zeitpunkt ≥ 50 Kopien/ml war, werden als Non-Responder gezählt. Abhängig von einer Umstellung der Backbone-Therapie kann ein Patient auch als Non-Responder gezählt werden, wenn • die Backbone-Therapie laut Protokoll nicht zulässig ist oder • die Backbone-Therapie laut Protokoll zulässig ist, die Entscheidung für die Umstellung jedoch vor oder zum ersten Studienzeitpunkt (unter Studienmedikation) mit Messung der Viruslast getroffen wurde. Im Anschluss an diese Operationalisierung ist der Snapshot-Algorithmus zusätzlich zum besseren Verständnis grafisch dargestellt. Die Analysen basieren auf der ITT-E Population. Die dargestellten Effektschätzer sind jeweils adjustiert für Alter (< 50, ≥ 50) und *Baseline Third Agent* (INI, PI, NNRTI).
SWORD-2	Die Responderrate zu Woche 48 wurde nach dem Snapshot-Algorithmus (der FDA) bestimmt. Die Klassifizierung Responder (virologischer Erfolg) / Non-Responder (virologisches Versagen) erfolgt durch den letzten verfügbaren Wert der Viruslast im jeweiligen Studienintervall (< 50 Kopien/ml oder ≥ 50 Kopien/ml), während der Patient unter Behandlung steht. Liegen in einem Studienintervall keine Daten zur Viruslast vor, kann ein Patient nicht als Responder eingestuft werden. Abhängig vom Grund des Fehlens des Werts wird der Patient als Non-Responder bzw. als Missing (z.B. aufgrund unerwünschter Ereignisse) gezählt. Patienten, die aus anderen Gründen als wegen unerwünschter Ereignisse die Studie abbrechen und deren Viruslast zu dem Zeitpunkt ≥ 50 Kopien/ml war, werden als Non-Responder gezählt. Abhängig von einer Umstellung der Backbone-Therapie kann ein Patient auch als Non-Responder gezählt werden, wenn • die Backbone-Therapie laut Protokoll nicht zulässig ist oder • die Backbone-Therapie laut Protokoll zulässig ist, die Entscheidung für die Umstellung jedoch vor oder zum ersten Studienzeitpunkt (unter Studienmedikation) mit Messung der Viruslast getroffen wurde. Im Anschluss an diese Operationalisierung ist der Snapshot-Algorithmus zusätzlich zum besseren Verständnis grafisch dargestellt. Die Analysen basieren auf der ITT-E Population. Die dargestellten Effektschätzer sind jeweils adjustiert für Alter (< 50, ≥ 50) und *Baseline Third Agent* (INI, PI, NNRTI).

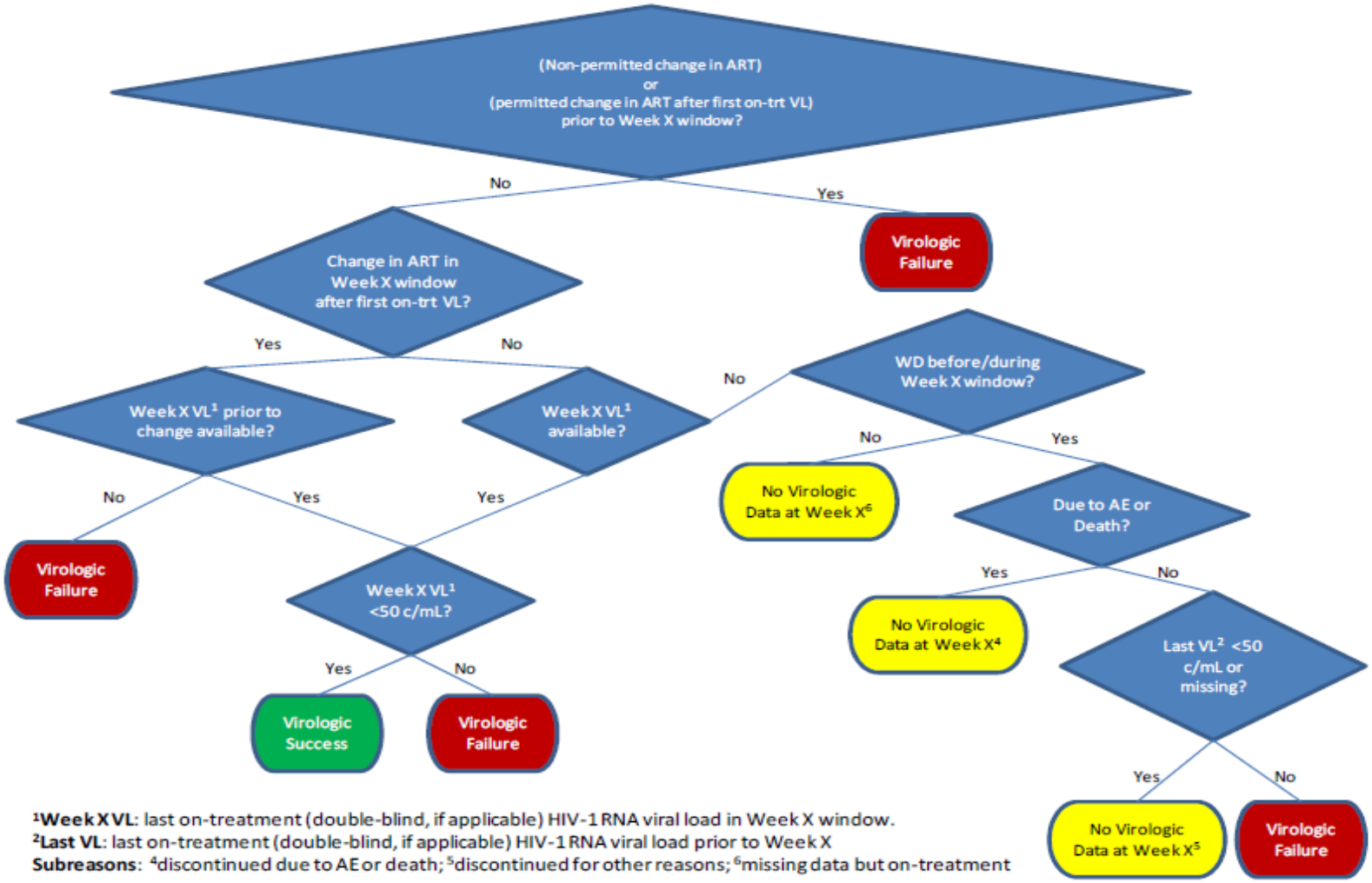


Abbildung 2: Snapshot-Algorithmus

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

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

Tabelle 4-16: Bewertung des Verzerrungspotenzials für Viruslast < 50 Kopien/ml zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
SWORD-1	niedrig	ja	ja	ja	ja	niedrig
SWORD-2	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde verblindet in einem zentralen Labor analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt somit keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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 4-17: Ergebnisse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - ITT-E Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	240 (95)	245 (96)	1,00 [0,97; 1,03], p=0,982 0,89 [0,38; 2,07], p=0,790 -0,65% [-4,25%; 2,96%], p=0,726
SWORD-2	261	255	246 (94)	240 (94)	0,99 [0,95; 1,04], p=0,771 1,02 [0,49; 2,14], p=0,961 0,18% [-3,89%; 4,25%], p=0,931

Quellen: Tabellen 2.1 und 2.2 in den Zusatzanalysen

Effektschätzer adjustiert für Alter und *Baseline Third Agent*.

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

Die Raten der Patienten mit einer Viruslast < 50 Kopien/ml sind in beiden Studien und in beiden Behandlungsarmen nach 48 Wochen jeweils vergleichbar und mit über 90 % sehr hoch. Es liegen jeweils keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

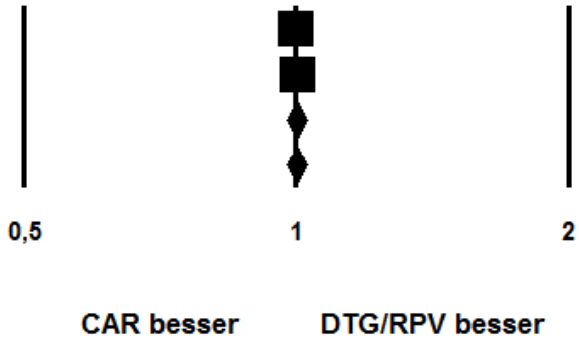
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 0 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. Im Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-18: Meta-Analyse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	1,00 [0,98 – 1,03] p=0,970
Modell mit zufälligen Effekten	1,00 [0,98 – 1,03] p=0,970
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI= Konfidenzintervall	

Auch in der Meta-Analyse liegt kein statistisch signifikanter Unterschied vor.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	1,000	0,970	1,030	1,000
	SWORD2	total	1,001	0,959	1,045	0,948
Fixed			1,000	0,976	1,025	0,970
Random			1,000	0,976	1,025	0,970



Heterogenität: I Quadrat = 0%

Abbildung 3: Meta-Analyse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Sensitivitätsanalysen:

Tabelle 4-19: Sensitivitätsanalyse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Sensitivitätsanalyse 1: Patienten ohne Angabe zur Viruslast ausgeschlossen

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	242	247	240 (99)	245 (99)	1,00 [0,98; 1,02], p=0,984 0,90 [0,12; 6,59], p=0,920 -0,08% [-1,71%; 1,55%], p=0,924
SWORD-2	247	244	246 (99,6)	240 (98)	1,01 [0,99; 1,03], p=0,176 3,88 [0,43; 35,21], p=0,228 1,17% [-0,62%; 2,97%], p=0,201

Quellen: Tabellen 2.3 und 2.4 in den Zusatzanalysen sowie eigene Berechnungen

Effektschätzer adjustiert für Alter und *Baseline Third Agent*.

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In der Sensitivitätsanalyse 1, bei der die Patienten aus der Analyse ausgeschlossen werden, die keine Angabe zur Viruslast zu Woche 48 haben, bestätigt sich das oben dargestellte Ergebnis des Snapshot-Algorithmus. Die Raten der Patienten mit einer Viruslast < 50 Kopien/ml liegen insgesamt sehr hoch, fast alle Patienten erfüllen das Kriterium, d.h. tatsächliches virologisches Versagen kommt über 48 Wochen Behandlung nur extrem selten vor (unter DTG/RPV nur zwei- bzw. einmal und unter CAR nur zwei- bzw. viermal - jeweils SWORD-1 bzw. SWORD-2). Es liegen keine statistisch signifikanten Unterschiede vor.

Tabelle 4-20: Sensitivitätsanalyse für Viruslast < 50 Kopien/ml zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Sensitivitätsanalyse 2: Imputation mittels Responderrate ohne fehlender Werte

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	240 (95)	245 (96)	1,00 [0,96; 1,03], p=0,801 0,90 [0,39; 2,07], p=0,801 -0,47% [-4,08%; 3,15%], p=0,801
SWORD-2	261	255	246 (94)	246 (96)	0,98 [0,94; 1,01], p=0,231 0,60 [0,26; 1,40], p=0,236 -2,22% [-5,84%; 1,40%], p=0,230

Quellen: Tabellen 2.5 und 2.6 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

Auch in der Sensitivitätsanalyse 2, bei der fehlende Werte gemäß der Responderrate der Patienten ohne fehlende Werte imputiert werden, ergibt sich ein ähnliches Bild wie beim Ergebnis des Snapshot-Algorithmus. Die Raten der Patienten mit einer Viruslast < 50 Kopien/ml liegen vergleichbar sehr hoch. Auch hier liegen keine statistisch signifikanten Unterschiede vor.

Beide durchgeführten Sensitivitätsanalysen zeigen sehr deutlich, dass die sehr hohe Übereinstimmung bezüglich der Raten der Patienten mit einer Viruslast < 50 Kopien/ml, die sich im Snapshot-Algorithmus gezeigt hatte, bestehen bleibt und lediglich der Absolutwert sich durch die unterschiedlichen Ersetzungsstrategien der beiden Sensitivitätsanalysen vom Snapshot-Algorithmus unterscheidet. Insbesondere durch die auch hinsichtlich der absoluten Erfolgsrate gezeigte, sehr hohe Übereinstimmung zwischen Snapshot-Algorithmus und Sensititvitätsanalyse 2 konnte gezeigt werden, dass „fehlende Werte zu Woche 48“ keine Verzerrung des Ergebnisses des Snapshot-Algorithmus nach sich gezogen haben.

Es sei hier ergänzend auf die Resultate bezüglich des Auftretens **AIDS-definierender Ereignisse (CDC-Klasse C)** eingegangen. Die **Mortalität** wird separat im Abschnitt 4.3.1.3.5 dargestellt. Es gab in beiden SWORD-Studien nur jeweils einen Todesfall pro Studienarm (SWORD-1: ein Todesfall im CAR-Arm; SWORD-2: ein Todesfall im DTG/RPV-Arm).

Bei Baseline befand sich kein Patient der beiden SWORD-Studien in der CDC-Klasse C. Im Studienverlauf über 48 Wochen traten Progressionen zur CDC-Klasse C bei insgesamt zwei Patienten (DTG/RPV) und einem Patienten (CAR) auf (SWORD-1), in der Studie SWORD-2 war es insgesamt nur ein Patient, der unter DTG/RPV eine Progression zur CDC-Klasse C aufwies. Die Seltenheit des Auftretens von Progressionen zur CDC-Klasse C in beiden SWORD-Studien insgesamt und die nur geringen, numerischen Unterschiede zwischen den Behandlungsgruppen unterstreichen, dass in der hier relevanten Teilpopulation der

Zielpopulation sowohl eine stabile virologische Suppression über die Studiendauer erhalten blieb (was der hier beschriebene Endpunkt „Virologischer Erfolg“ beschreibt). Auch die Seltenheit von unerwünschten Ereignissen über 48 Wochen, die Ereignisse umfasst hätten, die eine Einstufung der Patienten in die CDC-Klasse C nach sich gezogen hätten, unterstützen dies. Damit wird die Validität – insbesondere in der hier relevanten Patientenpopulation – des Endpunktes „Viruslast < 50 Kopien/ml“ auch durch die klinische Einschätzung deutlich untermauert.

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

4.3.1.3.2 EQ-5D zu Woche 48 – 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 EQ-5D zu Woche 48

Studie	Operationalisierung
SWORD-1	Der EQ-5D Fragebogen wurde jeweils zu Baseline, zu Woche 4, Woche 24 und Woche 48 von den Patienten ausgefüllt. **Utility Score:** Dargestellt wird hier die mittlere Änderung des gewichteten Summenscores (Utiltiy Score) zu Woche 48 gegenüber Baseline. Die Werte des Summenscores liegen zu den erhobenen Zeitpunkten im Allgemeinen zwischen 0 (Tod) und 1 (perfekte Gesundheit). **VAS/Thermometer Score:** Des Weiteren werden die Ergebnisse der visuellen Analogskala (VAS) des EQ-5D (auch als Thermometer Score bezeichnet) dargestellt. Diese gibt auf einer Skala von 0 (denkbar schlechtester Gesundheitszustand) bis 100 (bester denkbarer Gesundheitszustand) die Einschätzung des Patienten an, wie gut oder schlecht dieser seinen Gesundheitszustand einschätzt. Dargestellt werden die mittleren Änderungen gegenüber Baseline zu Woche 48. Die Analysen basieren jeweils auf der ITT-E Population. Fehlende Werte wurden mittels LOCF ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Geschlecht, Ethnizität, Alter, *Baseline Third Agent* und Baseline EQ-5D. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.
SWORD-2	Der EQ-5D Fragebogen wurde jeweils zu Baseline, zu Woche 4, Woche 24 und Woche 48 von den Patienten ausgefüllt. **Utility Score:** Dargestellt wird hier die mittlere Änderung des gewichteten Summenscores (Utiltiy Score) zu Woche 48 gegenüber Baseline. Die Werte des Summenscores liegen zu den erhobenen Zeitpunkten im Allgemeinen zwischen 0 (Tod) und 1 (perfekte Gesundheit). **VAS/Thermometer Score:** Des Weiteren werden die Ergebnisse der visuellen Analogskala (VAS) des EQ-5D (auch als Thermometer Score bezeichnet) dargestellt. Diese gibt auf einer Skala von 0 (denkbar schlechtester Gesundheitszustand) bis 100 (bester denkbarer Gesundheitszustand) die Einschätzung des Patienten an, wie gut oder schlecht dieser seinen Gesundheitszustand einschätzt. Dargestellt werden die mittleren Änderungen gegenüber Baseline zu Woche 48. Die Analysen basieren jeweils auf der ITT-E Population. Fehlende Werte wurden mittels LOCF ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Geschlecht, Ethnizität, Alter, *Baseline Third Agent* und Baseline EQ-5D. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.

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

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

Tabelle 4-22: Bewertung des Verzerrungspotenzials für EQ-5D zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunktherheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde von den Patienten selbst erhoben, die unverblindet bzgl. der Behandlungsgruppe waren. Es kann davon ausgegangen werden, dass trotz des offenen Studiendesigns der Patient in seiner individuellen Bewertung nicht von den Bewertungen anderer Patienten oder insbesondere anderer Patienten der alternativen Behandlungsgruppe beeinflusst wurde. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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 EQ-5D zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – ITT-E Population

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
EQ-5D Utility Score							
SWORD-1						0,004	0,047
DTG/RPV	252	0,946±0,09	252	0,945±0,10	-0,001±0,09	[-0,01; 0,02]	[-0,13; 0,22]
CAR	255	0,936±0,11	255	0,934±0,12	-0,002±0,09	p=0,603	p=0,600
SWORD-2						-0,002	-0,020
DTG/RPV	261	0,963±0,08	257	0,956±0,10	-0,008±0,09	[-0,02; 0,01]	[-0,20; 0,15]
CAR	252	0,951±0,09	253	0,949±0,11	-0,002±0,09	p=0,821	p=0,819
EQ-5D VAS							
SWORD-1						0,786	0,089
DTG/RPV	252	88,095±10,81	252	91,694±9,95	3,599±11,44	[-0,76; 2,33]	[-0,09; 0,26]
CAR	255	87,561±12,58	255	90,733±10,58	3,165±10,85	p=0,318	p=0,315
SWORD-2						-0,917	-0,099
DTG/RPV	261	89,885±10,21	257	91,132±11,41	1,249±10,35	[-2,55; 0,72]	[-0,27; 0,08]
CAR	252	87,482±11,26	253	91,115±9,48	3,572±11,64	p=0,270	p=0,267

Quelle: Tabellen 6.1-6.4 und 6.7-6.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse, VAS = Visuelle Analogskala

ANCOVA: adjustiert für Geschlecht, Ethnizität, Alter, *Baseline Third Agent* und Baseline EQ-5D

Die Werte des **EQ-5D Utility Score** sind bereits zu Baseline sehr hoch und verändern sich kaum bis zu Woche 48. Dies ist für beide Studien und beide Behandlungsarme zu beobachten. Es liegen keine statistisch signifikanten Unterschiede vor.

Auch bzgl. der **EQ-5D VAS** sind die Baseline-Werte relativ hoch mit knapp 90 (von 100) Score-Punkten im Mittel. Die Werte verbessern sich insgesamt noch ein wenig im Studienverlauf. Die geringen Unterschiede zwischen den Studienarmen sind nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel, sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

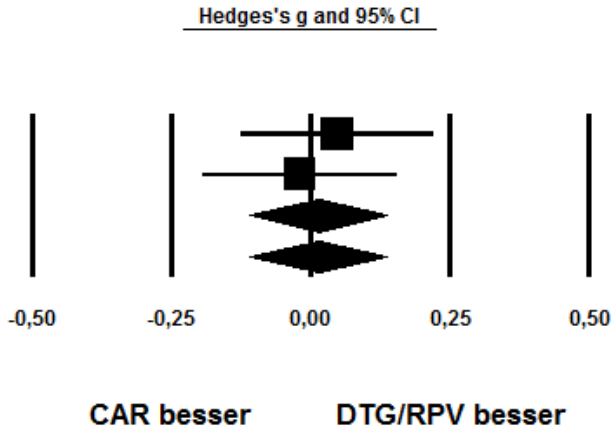
Für die vorliegenden Endpunkte liegt maximal mäßige Heterogenität vor (I^2 = 0 % bzw. I^2 = 55,2 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand der Modelle mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels Forest-Plots dargestellt.

Tabelle 4-24: Meta-Analyse für EQ-5D zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
EQ-5D Utility Score	
Heterogenität: I^2=0%	
Modell mit festen Effekten	0,014 [-0,110 – 0,137] p=0,830
Modell mit zufälligen Effekten	0,014 [-0,110 – 0,137] p=0,830
EQ-5D VAS	
Heterogenität: I^2=55,2%	
Modell mit festen Effekten	-0,005 [-0,128 – 0,118] p=0,937
Modell mit zufälligen Effekten	-0,005 [-0,189 – 0,179] p=0,958
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI= Konfidenzintervall, VAS = visuelle Analogskala	

Für beide Scores des EQ-5D ergeben sich in der meta-analytischen Betrachtung keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	0,047	-0,127	0,221	0,597
	SWORD-2	total	-0,020	-0,194	0,154	0,822
Fixed			0,014	-0,110	0,137	0,830
Random			0,014	-0,110	0,137	0,830

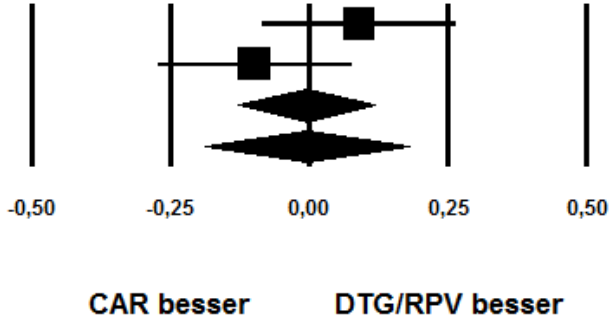


Heterogenität: I Quadrat = 0%

Abbildung 4: Meta-Analyse für EQ-5D Utility Score zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	0,089	-0,085	0,263	0,317
	SWORD-2	total	-0,099	-0,273	0,075	0,266
Fixed			-0,005	-0,128	0,118	0,937
Random			-0,005	-0,189	0,179	0,958



Heterogenität: I Quadrat = 55,2%

Abbildung 5: Meta-Analyse für EQ-5D VAS zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.3 Symptom Distress Module zu Woche 48 – RCT

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

Tabelle 4-25: Operationalisierung von SDM zu Woche 48

Studie	Operationalisierung
SWORD-1	Der SDM Fragebogen wurde jeweils zu Baseline, zu Woche 4, Woche 24 und Woche 48 von den Patienten ausgefüllt. **Symptom Bother Score:** Der Symptom Bother Score kann Werte im Bereich von 0 (keine Symptome) bis 80 (maximale Belastung durch die Symptome) annehmen. Dargestellt wird die mittlere Änderung des Scores zu Woche 48 gegenüber Baseline. **Symptom Count Score:** Symptom Count kann Werte im Bereich von 0 (keine Symptome) bis 20 (alle Symptome) annehmen. Dargestellt wird die mittlere Änderung des Scores zu Woche 48 gegenüber Baseline. Die Analysen basieren jeweils auf der ITT-E Population. Fehlende Werte wurden mittels LOCF ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Geschlecht, Ethnizität, Alter, *Baseline Third Agent* und Baseline SDM. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.
SWORD-2	Der SDM Fragebogen wurde jeweils zu Baseline, zu Woche 4, Woche 24 und Woche 48 von den Patienten ausgefüllt. **Symptom Bother Score:** Der Symptom Bother Score kann Werte im Bereich von 0 (keine Symptome) bis 80 (maximale Belastung durch die Symptome) annehmen. Dargestellt wird die mittlere Änderung des Scores zu Woche 48 gegenüber Baseline. **Symptom Count Score:** Symptom Count kann Werte im Bereich von 0 (keine Symptome) bis 20 (alle Symptome) annehmen. Dargestellt wird die mittlere Änderung des Scores zu Woche 48 gegenüber Baseline. Die Analysen basieren jeweils auf der ITT-E Population. Fehlende Werte wurden mittels LOCF ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Geschlecht, Ethnizität, Alter, *Baseline Third Agent* und Baseline SDM. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.

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

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

Tabelle 4-26: Bewertung des Verzerrungspotenzials für SDM zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde von den Patienten selbst erhoben, die unverblindet bzgl. der Behandlungsgruppe waren. Es kann davon ausgegangen werden, dass trotz des offenen Studiendesigns der Patient in seiner individuellen Bewertung der vorhandenen Symptome und der durch sie bestehenden Belastung nicht von den Bewertungen anderer Patienten oder insbesondere anderer Patienten der alternativen Behandlungsgruppe beeinflusst wurde. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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

Tabelle 4-27: Ergebnisse für SDM zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – ITT-E Population

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
SDM Symptom Count Score							
SWORD-1						-0,660	-0,146
DTG/RPV	214	5,4±4,96	215	5,0±5,24	-0,4±5,52	[-1,54; 0,22]	[-0,34; 0,05]
CAR	202	5,9±5,24	204	6,0±5,38	0,0±4,49	p=0,141	p=0,138
SWORD-2						-0,383	-0,092
DTG/RPV	232	5,2±4,80	229	4,6±5,17	-0,5±4,33	[-1,15; 0,38]	[-0,28; 0,09]
CAR	232	5,9±5,15	234	5,5±5,30	-0,4±4,82	p=0,327	p=0,324
SDM Symptom Bother Score							
SWORD-1						-1,569	-0,21
DTG/RPV	214	9,5±9,62	215	8,1±8,64	-1,4±8,32	[-3,05; -0,09]	[-0,40; -0,01]
CAR	202	11,8±12,2	204	11,2±11,37	-0,7±9,03	p=0,038	p=0,037
SWORD-2						-1,037	-0,13
DTG/RPV	232	9,7±10,42	229	8,2±10,52	-1,5±7,97	[-2,50; 0,43]	[-0,32; 0,05]
CAR	232	10,4±10,33	234	9,7±10,94	-0,7±9,30	p=0,164	p=0,162

Quelle: Tabellen 6.13-6.18, 6.23 und 6.24 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse, SDM = Symptom Distress Module, VAS = Visuelle Analogskala

ANCOVA: adjustiert für Geschlecht, Ethnizität, Alter, *Baseline Third Agent* und Baseline SDM Score

Für beide Scores liegen im Mittel in beiden Studien und Behandlungarmen jeweils Verbesserungen zu Woche 48 gegenüber Baseline vor. Die Verbesserungen sind im DTG/RPV-Arm für beide Scores und in beiden Studien jeweils numerisch deutlicher als im CAR-Arm. Die beobachteten Unterschiede sind für den Symptom Bother Score in Studie SWORD-1 statistisch signifikant im Sinne einer deutlicheren Verbesserung im Studienverklauf unter DTG/RPV als unter CAR, für alle anderen Endpunkte jedoch nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

Für die vorliegenden Endpunkte liegt jeweils keine Heterogenität vor ($I^2 = 0$ %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand der Modelle mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels Forest-Plots dargestellt.

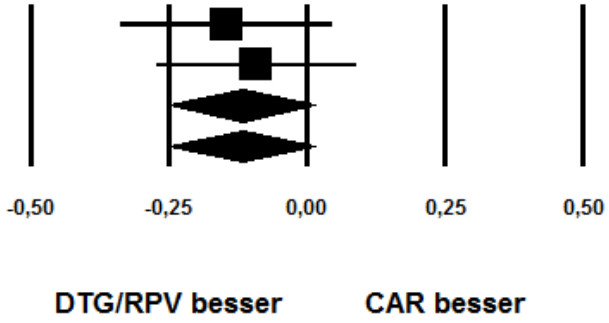
Tabelle 4-28: Meta-Analyse für SDM zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
SDM Symptom Count Score	
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,118 [-0,250; 0,015] p=0,081
Modell mit zufälligen Effekten	-0,118 [-0,250; 0,015] p=0,081
SDM Symptom Bother Score	
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,167 [-0,299; -0,034] p=0,014
Modell mit zufälligen Effekten	-0,167 [-0,299; -0,034] p=0,014
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall, SDM = Symptom Distress Module, VAS = visuelle Analogskala	

Für den **Symptom Count Score** zeigt das Ergebnis der Meta-Analyse einen numerischen Trend zugunsten von DTG/RPV, welcher jedoch nicht statistisch signifikant ist.

Die Meta-Analyse für den **Symptom Bother Score** führt hingegen zu einem statistisch signifikanten Ergebnis zugunsten von DTG/RPV (p = 0,014).

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	-0,146	-0,338	0,046	0,136
	SWORD-2	total	-0,092	-0,274	0,090	0,323
Fixed			-0,118	-0,250	0,015	0,081
Random			-0,118	-0,250	0,015	0,081

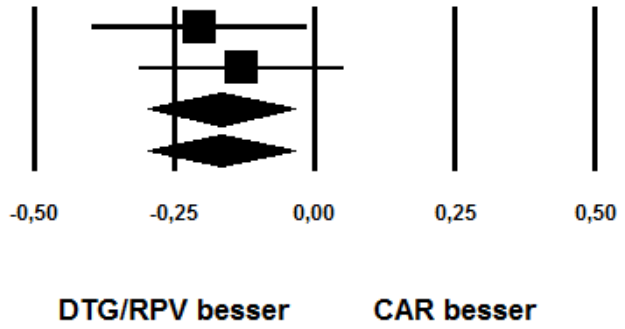


Heterogenität: I Quadrat = 0%

Abbildung 6: Meta-Analyse für SDM Symptom Count Score zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	-0,206	-0,398	-0,014	0,036
	SWORD-2	total	-0,131	-0,313	0,051	0,159
Fixed			-0,167	-0,299	-0,034	0,014
Random			-0,167	-0,299	-0,034	0,014



Heterogenität: I Quadrat = 0%

Abbildung 7: Meta-Analyse für SDM Symptom Bother Score zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.4 HIVTSQ Total Score zu Woche 48 – RCT

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

Tabelle 4-29: Operationalisierung von HIVTSQ Total Score zu Woche 48

Studie	Operationalisierung
SWORD-1	Der HIVTSQ (*HIV Treatment Satisfaction Questionnaire*) Fragebogen wurde jeweils zu Baseline, zu Woche 4, Woche 24 und Woche 48 von den Patienten ausgefüllt. Der HIVTSQ Total Score misst anhand von 10 Fragen die Patientenzufriedenheit. Dabei handelt es sich um allgemeine und spezifische Fragen, die jeweils Werte zwischen 0 und 6 annehmen. Der Total Score ist die Summe über die Werte dieser 10 Fragen und nimmt somit Werte zwischen 0 und 60 an; je höher der Wert, desto besser die Zufriedenheit. Dargestellt wird die mittlere Änderung des HIVTSQ Total Scores zu Woche 48 gegenüber Baseline. Die Analysen basieren jeweils auf der ITT-E Population. Fehlende Werte wurden mittels LOCF ersetzt. Zusätzlich zu den mittleren Änderungen wird Hedges‘ g als standardisiertes Effektmaß angegeben.
SWORD-2	Der HIVTSQ (*HIV Treatment Satisfaction Questionnaire*) Fragebogen wurde jeweils zu Baseline, zu Woche 4, Woche 24 und Woche 48 von den Patienten ausgefüllt. Der HIVTSQ Total Score misst anhand von 10 Fragen die Patientenzufriedenheit. Dabei handelt es sich um allgemeine und spezifische Fragen, die jeweils Werte zwischen 0 und 6 annehmen. Der Total Score ist die Summe über die Werte dieser 10 Fragen und nimmt somit Werte zwischen 0 und 60 an; je höher der Wert, desto besser die Zufriedenheit. Dargestellt wird die mittlere Änderung des HIVTSQ Total Scores zu Woche 48 gegenüber Baseline. Die Analysen basieren jeweils auf der ITT-E Population. Fehlende Werte wurden mittels LOCF ersetzt. Zusätzlich zu den mittleren Änderungen wird Hedges‘ g als standardisiertes Effektmaß angegeben.

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

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

Tabelle 4-30: Bewertung des Verzerrungspotenzials für HIVTSQ Total Score zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde von den Patienten selbst erhoben, die unverblindet bzgl. der Behandlungsgruppe waren. Es kann davon ausgegangen werden, dass trotz des offenen Studiendesigns der Patient in seiner individuellen Bewertung seiner Therapiezufriedenheit nicht von den Bewertungen anderer Patienten oder insbesondere anderer Patienten der alternativen Behandlungsgruppe beeinflusst wurde. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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 4-31: Ergebnisse für HIVTSQ Total Score zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – ITT-E Population

Studie	**Baseline**		**Woche 48**		**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**			
SWORD-1						1,00	0,15
DTG/RPV	252	53,9±6,64	252	55,8±6,35	1,8±7,30	[-0,13; 2,13]	[-0,02; 0,33]
CAR	255	54,0±6,48	255	54,7±6,88	0,8±5,55	p=0,083	p=0,083
SWORD-2						1,20	0,18
DTG/RPV	261	54,8±6,18	257	55,9±5,85	1,2±6,64	[0,06; 2,34]	[0,01; 0,36]
CAR	252	53,8±6,67	253	53,9±7,65	0,0±6,49	p=0,039	p=0,040

Quelle: Tabelle 6.21 und 6.22 in den Zusatzanalysen sowie eigene Berechnungen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse, HIVTSQ = HIV Treatment Satisfaction Questionnaire

In beiden Studien verbessert sich der Score im DTG/RPV-Arm im Mittel mehr als im CAR-Arm. In SWORD-2 ist dieser Unterschied zudem statistisch signifikant (p = 0,040), in SWORD-1 liegt keine statistische Signifikanz vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

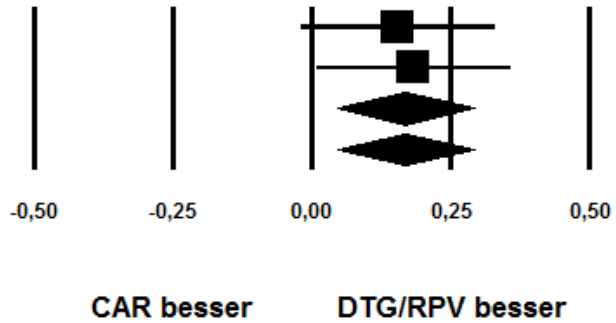
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 0 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-32: Meta-Analyse für HIVTSQ Total Score zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	0,168 [0,045 – 0,291] p=0,007
Modell mit zufälligen Effekten	0,168 [0,045 – 0,291] p=0,007
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI= Konfidenzintervall, HIVTSQ = HIV Treatment Satisfaction Questionnaire	

Das Ergebnis der Meta-Analyse zeigt einen statistisch signifikanten Unterschied zugunsten von DTG/RPV gegenüber CAR (p = 0,007).

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	0,154	-0,020	0,328	0,083
	SWORD-2	total	0,182	0,008	0,357	0,040
Fixed			0,168	0,045	0,291	0,007
Random			0,168	0,045	0,291	0,007



Heterogenität: I Quadrat = 0%

Abbildung 8: Meta-Analyse für HIVTSQ Total Score zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.5 Mortalität bis Woche 48 – RCT

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

Tabelle 4-33: Operationalisierung von Mortalität bis Woche 48

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, die als Todesfälle im Rahmen der Studie bis Woche 48 gemeldet wurden. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.
SWORD-2	Anzahl der Patienten, die als Todesfälle im Rahmen der Studie bis Woche 48 gemeldet wurden. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.

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

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

Tabelle 4-34: Bewertung des Verzerrungspotenzials für Mortalität bis Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Diese Tatsache hat jedoch bei diesem objektiv erhobenen Endpunkt keine Auswirkungen auf ein mögliches erhöhtes Verzerrungspotenzial. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren

existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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

Tabelle 4-35: Ergebnisse für Mortalität bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	0 (0)	1 (<1)	0,34 [0,01; 8,27], p=0,507 0,34 [0,01; 8,32], p=0,506 -0,39% [-1,47%; 0,69%], p=0,478
SWORD-2	261	255	1 (<1)	0 (0)	2,93 [0,12; 71,62], p=0,510 2,94 [0,12; 72,57], p=0,509 0,38% [-0,75%; 1,45%], p=0,479

Quellen: Tabellen 31 (SWORD-1) und 32 (SWORD-2) in den jeweiligen Studienberichten sowie eigene Berechnungen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien wurde jeweils nur ein Todesfall berichtet, einer unter CAR (SWORD-1) und einer unter DTG/RPV (SWORD-2). Es liegen keine statistisch signifikanten Unterschiede vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Aufgrund der Seltenheit der Ereignisse wurde für diesen Endpunkt keine Meta-Analyse durchgeführt.

4.3.1.3.6 Unerwünschte Ereignisse bis Woche 48 – RCT

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

Tabelle 4-36: Operationalisierung von unerwünschte Ereignisse bis Woche 48

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mindestens ein unerwünschtes Ereignis (UE) im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.
SWORD-2	Anzahl der Patienten, für die mindestens ein unerwünschtes Ereignis (UE) im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.

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

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

Tabelle 4-37: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse bis Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Es existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten. Bei dem Endpunkt „jedwedes

unerwünschtes Ereignis“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens sechs Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt potentiell dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-38: Ergebnisse für unerwünschte Ereignisse bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	200 (79)	190 (74)	1,07 [0,97; 1,18], p=0,159 1,34 [0,88; 2,02], p=0,172 5,24% [-2,15%; 12,64%], p=0,165
SWORD-2	261	255	195 (75)	174 (68)	1,09 [0,97; 1,21], p=0,134 1,36 [0,92; 2,00], p=0,478 6,25% [-1,54%; 14,04%], p=0,116

Quellen: Tabellen 3.1 und 3.2 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien sind die Ereignisraten etwas höher im DTG/RPV-Arm als im CAR-Arm (79 % vs. 74 % bzw. 75 % vs. 68 %). Diese beobachteten Unterschiede sind jedoch nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

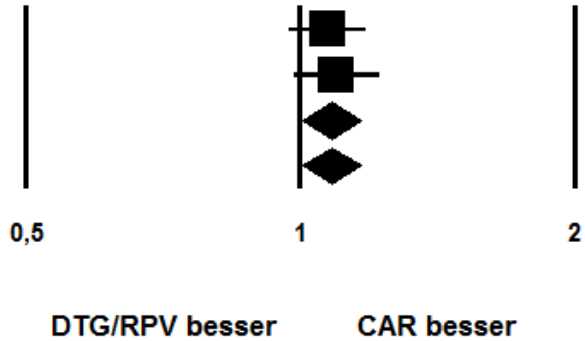
Für den vorliegenden Endpunkt liegt keine Heterogenität vor ($I^2 = 0$ %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-39: Meta-Analyse für unerwünschte Ereignisse bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	1,08 [1,01 – 1,16] p=0,037
Modell mit zufälligen Effekten	1,08 [1,01 – 1,16] p=0,037
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI= Konfidenzintervall	

Der in den Einzelstudien beobachtete Trend, der auf Studienebene jeweils nicht statistisch signifikant war, ist im Ergebnis der Meta-Analyse mit einem relativen Risiko von 1,08 zuungunsten von DTG/RPV statistisch signifikant (p=0,037).

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	1,070	0,970	1,180	0,176
	SWORD2	total	1,095	0,981	1,222	0,105
Fixed			1,081	1,005	1,163	0,037
Random			1,081	1,005	1,163	0,037



Heterogenität: I Quadrat = 0%

Abbildung 9: Meta-Analyse für unerwünschte Ereignisse bis Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.7 Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 – RCT

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

Tabelle 4-40: Operationalisierung von unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mindestens ein unerwünschtes Ereignis (UE) von Grad 3 oder 4 im Rahmen der Studie bis Woche 48 gemeldet wurde. Die Klassifizierung des Schweregrads erfolgte durch die Beurteilung des Prüfarztes anhand der DAIDS (*Division of AIDS*) Toxizitätsskala. Grad 3 bzw. Grad 4 entspricht demnach dem Schweregrad „schwer" bzw. „potenziell lebensbedrohlich“. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.
SWORD-2	Anzahl der Patienten, für die mindestens ein unerwünschtes Ereignis (UE) von Grad 3 oder 4 im Rahmen der Studie bis Woche 48 gemeldet wurde. Die Klassifizierung des Schweregrads erfolgte durch die Beurteilung des Prüfarztes anhand der DAIDS (*Division of AIDS*) Toxizitätsskala. Grad 3 bzw. Grad 4 entspricht demnach dem Schweregrad „schwer" bzw. „potenziell lebensbedrohlich“. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.

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

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

Tabelle 4-41: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten. Trotz der bereits erläuterten a priori bestehenden Ungleichheit zwischen beiden Behandlungsgruppen hinsichtlich der Beurteilung von unerwünschten Ereignissen bei dem offenem Studiendesign der beiden SWORD-Studien, wird bei unerwünschten Ereignissen vom Schweregrad 3 oder 4 nicht von einem erhöhten Verzerrungspotenzial ausgegangen, da hier eine im Vergleich zu weniger schweren unerwünschten Ereignissen noch genauere kausale Berücksichtigung der Umstände bei der Beurteilung durch den Prüfarzt sowie das Einbeziehen von Plausibilitätsargumenten vorausgesetzt werden kann.

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 4-42: Ergebnisse für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	15 (6)	15 (6)	0,98 [0,49; 1,95], p=0,955 1,01 [0,48; 2,13], p=0,979 0,28% [-3,82%; 4,38%], p=0,893
SWORD-2	261	255	17 (7)	5 (2)	3,28 [1,23; 8,76], p=0,018 3,48 [1,26; 9,60], p=0,016 4,51% [1,06%; 7,97%], p=0,010

Quellen: Tabelle 3.3 und 3.4 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In SWORD-1 sind die Ereignisraten in beiden Studienarmen ausgeglichen (jeweils 6 %), es besteht kein statistisch signifikanter Unterschied. In SWORD-2 ist die Rate im DTG/RPV-Arm etwas höher (7 % vs. 2 %), was zu einem statistisch signifikant höheren Risiko in dieser Behandlungsgruppe führt (p = 0,018).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

Für den vorliegenden Endpunkt liegt hohe Heterogenität vor (I^2 = 74,3 %). Die Ergebnisse dieses Endpunktes werden daher auf Basis der Einzelstudien zusammenfassend interpretiert. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

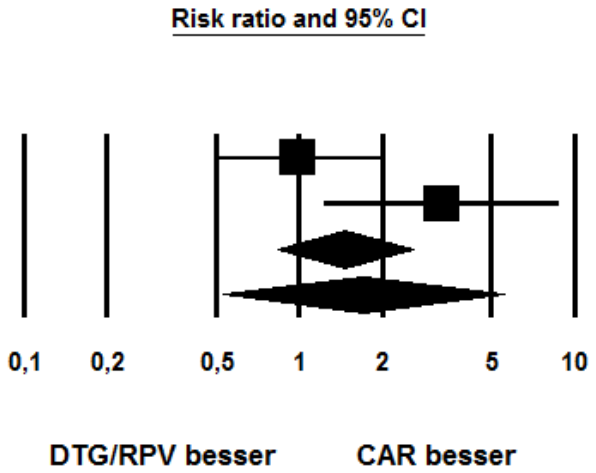
Tabelle 4-43: Meta-Analyse für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=74,3%	
Modell mit festen Effekten	1,46 [0,83 – 2,57] p=0,188
Modell mit zufälligen Effekten	1,70 [0,52 – 5,53] p=0,377
Modell mit zufälligen Effekten[a] (Knapp-Hartung-Methode)	1,74 [0,001 – 3139,73] p=0,519
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall [a] basierend auf den nicht-adjustierten Schätzern	

Im meta-analytischen Modell mit zufälligen Effekten zeigt sich kein statistisch signifikanter Unterschied.

Zur Interpretation der Evidenz zu „unerwünschten Ereignissen vom Grad 3 oder 4 bis Woche 48“ ist zu beachten, dass bei hoher Heterogenität der Meta-Analyse die Ergebnisse der beiden Einzelstudien keinen einheitlichen Trend abzuleiten gestatten. Während in der Studie SWORD-1 keinerlei relevanter Trend zugunsten einer Behandlung auszumachen ist und das Ergebnis folgerichtig auch nicht statistisch signifikant ist, liegt nur in der Studie SWORD-2 ein Trend zu einer höheren Anzahl von unerwünschten Ereignissen des Grades 3 oder 4 vor, der auch statistisch signifikant ist (p=0,018). Indizien, weshalb sich die beiden Studien hinsichtlich dieses Endpunktes unterscheiden, konnten nicht identifiziert werden. Es fällt auch auf, dass der hauptsächliche Unterschied zwischen den beiden Studienergebnissen sehr deutlich von einer abweichend sehr viel niedrigeren relativen Häufigkeit von unerwünschten Ereignissen vom Grad 3 oder 4 im CAR-Arm der SWORD-2 im Vergleich zur SWORD-1 (nur noch ein Drittel des Wertes bei SWORD-1) abzuhängen scheint, die relativen Auftretenshäufigkeiten im DTG/RPV-Arm jedoch über beide Studien hinweg sehr gut reproduzierbar sind. Insofern ist diese Situation so zu interpretieren, dass es sich hierbei höchstwahrscheinlich um Schwankungen im Rahmen von zwei Einzelstudien handelt und damit in der Gesamtschau keine Aussage zu einem generellen Trend oder gar einem statistisch signifikanten Ergebnis in die eine oder andere Richtung möglich ist. Diese Interpretation wird auch durch das statistisch nicht signifikante Ergebnis der Meta-Analyse mit zufälligen Effekten gestützt, ohne dass diese primär zur Interpretation direkt herangezogen werden sollte.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	0,980	0,491	1,955	0,954
	SWORD2	total	3,280	1,229	8,753	0,018
Fixed			1,462	0,831	2,572	0,188
Random			1,701	0,523	5,533	0,377



Heterogenität: I Quadrat = 74,3%

Abbildung 10: Meta-Analyse für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.8 Schwerwiegende unerwünschte Ereignisse bis Woche 48 – RCT

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

Tabelle 4-44: Operationalisierung von schwerwiegende unerwünschte Ereignisse bis Woche 48

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mindestens ein schwerwiegendes unerwünschtes Ereignis (SUE) im Rahmen der Studie bis Woche 48 gemeldet wurde. Ein SUE ist definiert als UE, für welches mindestens eines der folgenden Umstände zutrifft: • Tod (s. unten) • unmittelbar lebensbedrohliche Situation • nötige Hospitalisierung bzw. Verlängerung eines bestehenden Krankenhausaufenthalts • resultierende Behinderung bzw. Invalidität • kongenitale Anomalie / Geburtsfehler • klinisch relevantes Ereignis (z.B. medizinische Intervention, um eine der obigen Konsequenzen zu verhindern) • assoziierte Leberverletzung und verschlechterte Leberfunktion Todesfälle, die im Rahmen der Studienteilnahme aufgetreten sind, werden in der vorliegenden Operationalisierung nicht gezählt (s. hierzu eigenen Endpunkt zur Mortalität). Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.
SWORD-2	Anzahl der Patienten, für die mindestens ein schwerwiegendes unerwünschtes Ereignis (SUE) im Rahmen der Studie bis Woche 48 gemeldet wurde. Ein SUE ist definiert als UE, für welches mindestens eines der folgenden Umstände zutrifft: • Tod (s. unten) • unmittelbar lebensbedrohliche Situation • nötige Hospitalisierung bzw. Verlängerung eines bestehenden Krankenhausaufenthalts • resultierende Behinderung bzw. Invalidität • kongenitale Anomalie / Geburtsfehler • klinisch relevantes Ereignis (z.B. medizinische Intervention, um eine der obigen Konsequenzen zu verhindern) • assoziierte Leberverletzung und verschlechterte Leberfunktion Todesfälle, die im Rahmen der Studienteilnahme aufgetreten sind, werden in der vorliegenden Operationalisierung nicht gezählt (s. hierzu eigenen Endpunkt zur Mortalität). Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.

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

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

Tabelle 4-45: Bewertung des Verzerrungspotenzials für schwerwiegende unerwünschte Ereignisse bis Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten. Trotz der bereits erläuterten a priori bestehenden Ungleichheit zwischen beiden Behandlungsgruppen hinsichtlich der Beurteilung von unerwünschten Ereignissen bei dem offenem Studiendesign der beiden SWORD-Studien, wird bei schwerwiegenden unerwünschten Ereignissen nicht von einem erhöhten Verzerrungspotenzial ausgegangen, da hier eine im Vergleich zu weniger schweren unerwünschten Ereignissen noch genauere kausale Berücksichtigung der Umstände bei der Beurteilung durch den Prüfarzt sowie das Einbeziehen von Plausibilitätsargumenten vorausgesetzt werden kann.

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 4-46: Ergebnisse für schwerwiegende unerwünschte Ereignisse bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	9 (4)	11 (4)	0,81 [0,34; 1,90], p=0,623 0,82 [0,33; 2,02], p=0,662 -0,59% [-4,01%; 2,83%], p=0,736
SWORD-2	261	255	18 (7)	9 (4)	1,95 [0,89; 4,26], p=0,094 2,03 [0,89; 4,61], p=0,092 3,34% [-0,52%; 7,19%], p=0,090

Quellen: Tabellen 3.5 und 3.6 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In SWORD-1 sind die Ereignisraten in beiden Studienarmen ausgeglichen (jeweils 4 %), in SWORD-2 ist die Rate im DTG/RPV-Arm etwas höher (7 % vs. 4 %). Es besteht jeweils kein statistisch signifikanter Unterschied.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

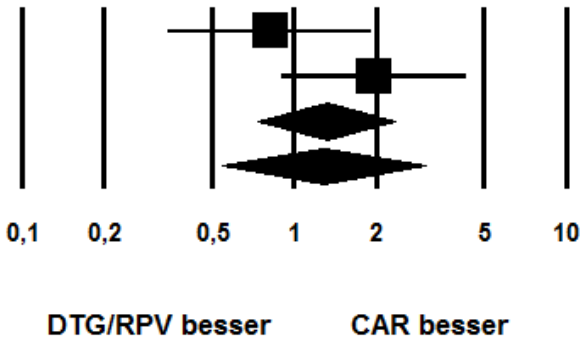
Für den vorliegenden Endpunkt liegt mäßige Heterogenität vor (I^2 = 54,4%). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-47: Meta-Analyse für schwerwiegende unerwünschte Ereignisse bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=54,4%	
Modell mit festen Effekten	1,31 [0,73 – 2,34] p=0,361
Modell mit zufälligen Effekten	1,28 [0,54 – 3,03] p=0,573
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Das Ergebnis der Meta-Analyse zeigt keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungsarmen.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	0,810	0,343	1,915	0,631
	SWORD2	total	1,950	0,891	4,266	0,095
Fixed			1,310	0,734	2,337	0,361
Random			1,281	0,542	3,027	0,573



Heterogenität: I Quadrat = 54,4%

Abbildung 11: Meta-Analyse für schwerwiegende unerwünschte Ereignisse bis Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.9 Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten – RCT

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

Tabelle 4-48: Operationalisierung von Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mindestens ein unerwünschtes Ereignis (UE), das zum Therapieabbruch geführt hatte, im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.
SWORD-2	Anzahl der Patienten, für die mindestens ein unerwünschtes Ereignis (UE), das zum Therapieabbruch geführt hatte, im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen auf der *Safety Population*.

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

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

Tabelle 4-49: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunktherheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine

ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass für diesen Endpunkt ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse, die zum Studienabbruch führen, potentiell häufiger im DTG/RPV-Arm berichtet werden. Insbesondere der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger im Zusammenhang mit einem unerwünschten Ereignis einen Therapieabbruch durchzuführen, als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-50: Ergebnisse für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	9 (4)	2 (<1)	4,49 [0,98; 20,49], p=0,053 4,73 [1,01; 22,24], p=0,049 2,88% [0,33%; 5,42%], p=0,027
SWORD-2	261	255	12 (5)	1 (<1)	11,69 [1,53; 89,25], p=0,018 12,22 [1,58; 94,81], p=0,017 4,20% [1,54%; 6,87%], p=0,002

Quellen: Tabellen 3.7 und 3.8 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien sind die Ereignisraten etwas höher im DTG/RPV-Arm als im CAR-Arm (4 % vs. < 1 % bzw. 5 % vs. < 1 %). Diese beobachteten Unterschiede sind hinsichtlich des relativen Risikos in SWORD-2 statistisch signifikant (p = 0,018), in SWORD-1 liegt für das relative Risiko jedoch keine statistische Signifikanz vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

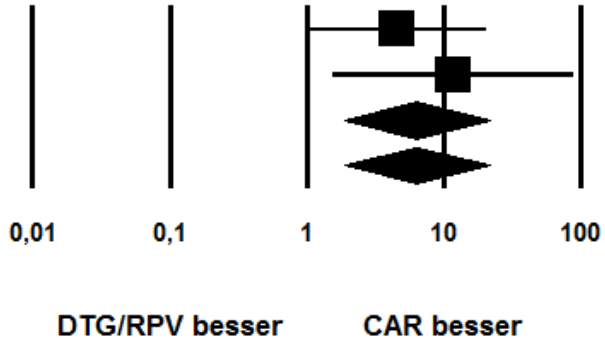
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 0 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-51: Meta-Analyse für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	6,33 [1,87 – 21,38] p=0,003
Modell mit zufälligen Effekten	6,33 [1,87 – 21,38] p=0,003
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Das Ergebnis der Meta-Analyse ist mit einem relativen Risiko von 6,33 zuungunsten von DTG/RPV statistisch signifikant (p = 0,003).

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	4,490	0,982	20,531	0,053
	SWORD2	total	11,690	1,531	89,284	0,018
Fixed			6,328	1,873	21,378	0,003
Random			6,328	1,873	21,378	0,003



Heterogenität: I Quadrat = 0%

Abbildung 12: Meta-Analyse für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.10 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen – RCT

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

Tabelle 4-52: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit einem der folgenden MedDRA *Preferred Terms* (PT) im Rahmen der Studie bis Woche 48 gemeldet wurde: • *palpitation* • *pulmonary embolism* • *atrial fibrillation* • *peripheral arterial occlusive disease* • *venous thrombosis*. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit einem der folgenden MedDRA *Preferred Terms* (PT) im Rahmen der Studie bis Woche 48 gemeldet wurde: • *palpitation* • *pulmonary embolism* • *atrial fibrillation* • *peripheral arterial occlusive disease* • *venous thrombosis*. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-53: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse - Herzerkrankungen“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-54: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	1 (<1)	1 (<1)	0,99 [0,06; 15,71], p=0,995 0,99 [0,06; 16,13], p=0,996 0,05% [-1,03%; 1,14%], p=0,924
SWORD-2	261	255	1 (<1)	1 (<1)	1,03 [0,06; 16,27], p=0,985 1,03 [0,06; 16,59], p=0,985 0,02% [-1,07%; 1,11%], p=0,965

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien wurden pro Studienarm nur für jeweils einen Patienten Ereignisse gemeldet. Es liegen keine statistisch signifikanten Unterschiede vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

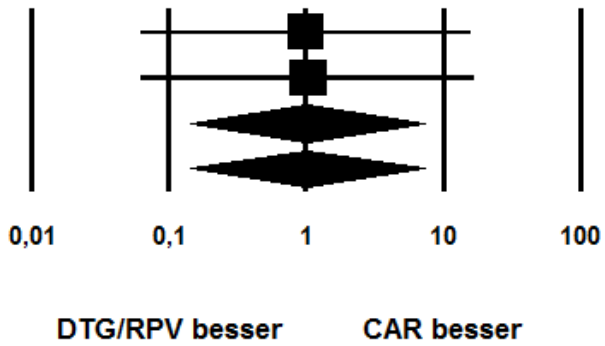
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 0%). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-55: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	1,01 [0,14 – 7,27] p=0,992
Modell mit zufälligen Effekten	1,01 [0,14 – 7,27] p=0,992
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Im meta-analytischen Modell zeigt sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	0,990	0,061	16,019	0,994
	SWORD2	total	1,030	0,063	16,961	0,984
Fixed			1,010	0,140	7,274	0,992
Random			1,010	0,140	7,274	0,992



Heterogenität: I Quadrat = 0%

Abbildung 13: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.11 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts – RCT

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

Tabelle 4-56: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*gastrointestinal disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*gastrointestinal disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-57: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Erkrankungen des Gastroinstestinaltraktes“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens sechs Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-58: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	71 (28)	56 (22)	1,28 [0,95; 1,74], p=0,104 4,41 [0,94; 2,11], p=0,101 6,63% [-0,82%; 14,08%], p=0,081
SWORD-2	261	255	58 (22)	26 (10)	2,16 [1,41; 3,32], p<0,001 2,49 [1,15; 4,11], p<0,001 11,90% [5,62%; 18,17%], p<0,001

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien sind die Ereignisraten etwas höher im DTG/RPV-Arm als im CAR-Arm (28 % vs. 22 % bzw. 22 % vs. 10 %). Diese beobachteten Unterschiede sind in SWORD-2 statistisch signifikant (p < 0,001), in SWORD-1 liegt jedoch keine statistische Signifikanz vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

Für den vorliegenden Endpunkt liegt hohe Heterogenität vor (I^2 = 73,9 %). Die Ergebnisse des Endpunktes werden daher auf Basis der Einzelstudien interpretiert. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

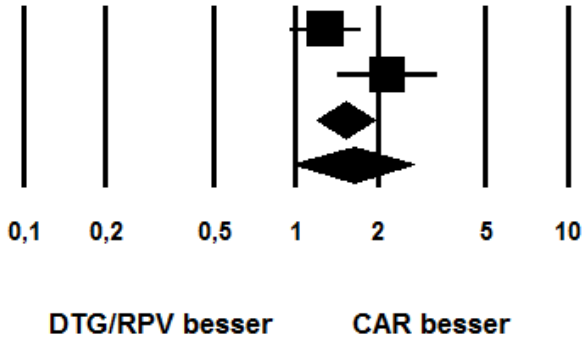
Tabelle 4-59: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=73,9%	
Modell mit festen Effekten	1,52 [1,19 – 1,95] p=0,001
Modell mit zufälligen Effekten	1,63 [0,98 – 2,71] p=0,062
Modell mit zufälligen Effekten[a] (Knapp-Hartung-Methode)	1,64 [0,06 – 45,73] p=0,311
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI= Konfidenzintervall [a] basierend auf den nicht-adjustierten Schätzern	

Im meta-analytischen Modell mit zufälligen Effekten zeigt sich kein statistisch signifikanter Unterschied.

Zur Interpretation der Evidenz zu „unerwünschten Ereignissen von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts“ ist zu beachten, dass bei hoher Heterogenität der Meta-Analyse die Ergebnisse in beiden Einzelstudien ein Trend zu einer höheren Ereignishäufigkeit unter DTG/RPV zu beobachten ist, der jedoch nur in der Studie SWORD-2 statistisch signifikant ist. Indizien, weshalb sich die beiden Studien hinsichtlich dieses Endpunktes unterscheiden, konnten nicht identifiziert werden. Es ist hervorzuheben, dass die beiden Einzelstudien sich sehr dominant durch die nur etwas weniger als halb so hohe relative Häufigkeit von Ereignissen unter CAR in SWORD-2 im Vergleich zu SWORD-1 unterscheiden, wohingegen die Ereignisrate unter DTG/RPV in beiden Studien sehr gut vergleichbar war. Da keine Gründe für diesen Unterschied zu identifizieren waren, sollte jedoch berücksichtigt werden, dass die hohe Heterogenität der Meta-Analyse hauptsächlich auf diesen Unterschied im CAR-Arm zurückzuführen zu sein scheint. Insofern ist diese Situation so zu interpretieren, dass es sich - unter Beachtung dieser Abweichungen zwischen den Studien - zwar um einen Trend zuungunsten von DTG/RPV handelt, dieser jedoch in der Gesamtschau derzeit keine klare Aussage erlaubt. Diese Interpretation wird auch durch das statistisch nicht signifikante Ergebnis der Meta-Analyse mit zufälligen Effekten gestützt, ohne dass diese primär zur Interpretation direkt herangezogen werden sollte.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	1,280	0,946	1,732	0,110
	SWORD2	total	2,160	1,408	3,314	0,000
Fixed			1,524	1,190	1,951	0,001
Random			1,625	0,975	2,709	0,062



Heterogenität: I Quadrat = 73,9%

Abbildung 14: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.12 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen – RCT

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

Tabelle 4-60: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*hepatobiliary disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*hepatobiliary disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-61: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Leber- und Gallenerkrankungen“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens sechs Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-62: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	3 (1)	1 (<1)	3,02 [0,32; 28,84], p=0,336 3,06 [0,32; 29,70], p=0,335 0,79% [-0,75%; 2,34%], p=0,315
SWORD-2	261	255	2 (<1)	1 (<1)	2,06 [0,19; 22,45], p=0,552 2,05 [0,18; 22,96], p=0,559 0,39% [-0,93%; 1,72%], p=0,561

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien wurden pro Studienarm nur für sehr wenige Patienten Ereignisse festgestellt. Es liegen keine statistisch signifikanten Unterschiede vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

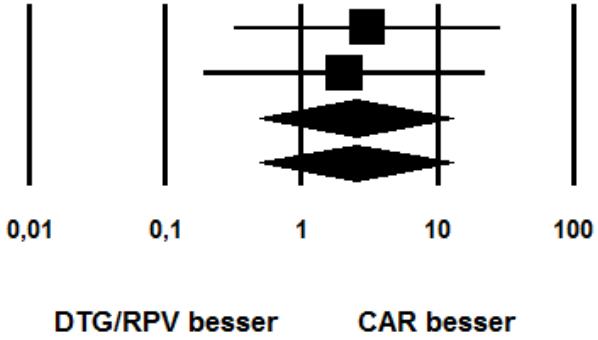
Für den vorliegenden Endpunkt liegt keine Heterogenität vor ($I^2 = 0$ %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-63: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	2,52 [0,49 – 12,97] p=0,268
Modell mit zufälligen Effekten	2,52 [0,49 – 12,97] p=0,268
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Im meta-analytischen Modell zeigt sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	3,020	0,318	28,670	0,336
	SWORD2	total	2,060	0,190	22,392	0,553
Fixed			2,522	0,491	12,966	0,268
Random			2,522	0,491	12,966	0,268



Heterogenität: I Quadrat = 0%

Abbildung 15: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.13 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität – RCT

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

Tabelle 4-64: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit einem der folgenden MedDRA *Preferred Terms* (PT) im Rahmen der Studie bis Woche 48 gemeldet wurde: • *hypersensitivity* • *drug hypersensitivity* • *allergic reaction.* Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit einem der folgenden MedDRA *Preferred Terms* (PT) im Rahmen der Studie bis Woche 48 gemeldet wurde: • *hypersensitivity* • *drug hypersensitivity* • *allergic reaction.* Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-65: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Hypersensitivität“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-66: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	1 (<1)	0 (0)	3,05 [0,13; 74,46], p=0,494 3,06 [0,12; 75,46], p=0,494 0,38% [-0,38%; 1,15%], p=0,325
SWORD-2	261	255	0 (0)	1 (<1)	0,33 [0,01; 7,96], p=0,491 0,32 [0,01; 8,00], p=0,491 -0,39% [-1,16%; 0,38%], p=0,318

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen sowie eigene Berechnungen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien wurde das Ereignis jeweils nur für einen Patienten berichtet, einer unter DTG/RPV (SWORD-1) und einer unter CAR (SWORD-2). Es liegen keine statistisch signifikanten Unterschiede vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Aufgrund der Seltenheit der Ereignisse, wurde für diesen Endpunkt keine Meta-Analyse berechnet.

4.3.1.3.14 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen – RCT

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

Tabelle 4-67: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*metabolism and nutrition disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*metabolism and nutrition disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-68: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunktherheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Stoffwechsel- und Ernährungsstörungen“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-69: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	2 (<1)	2 (<1)	1,00 [0,14; 7,01], p=0,997 1,00 [0,14; 7,20], p=0,998 -0,00% [-1,57%; 1,56%], p=0,996
SWORD-2	261	255	7 (3)	6 (2)	1,06 [0,36; 3,10], p=0,910 1,06 [0,35; 3,25], p=0,914 0,16% [-2,49%; 2,80%], p=0,908

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien sind die Ereignisraten der beiden Studienarme ausgeglichen (< 1 % vs. < 1 % bzw. 3 % vs. 2 %), es besteht jeweils kein statistisch signifikanter Unterschied.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

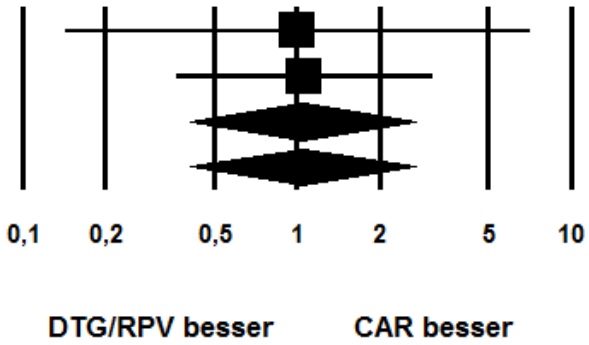
Für den vorliegenden Endpunkt liegt keine Heterogenität vor ($I^2 = 0$ %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-70: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	1,05 [0,41 – 2,69] p=0,926
Modell mit zufälligen Effekten	1,05 [0,41 – 2,69] p=0,926
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Im meta-analytischen Modell zeigt sich wie in den Einzelstudien kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	1,000	0,141	7,076	1,000
	SWORD2	total	1,060	0,361	3,111	0,916
Fixed			1,046	0,407	2,686	0,926
Random			1,046	0,407	2,686	0,926



Heterogenität: I Quadrat = 0%

Abbildung 16: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.15 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen – RCT

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

Tabelle 4-71: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*musculoskeletal and connective tissue disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*musculoskeletal and connective tissue disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-72: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Skelettmuskulatur-, Bindegewens- und Knochenerkrankungen“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-73: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	35 (14)	50 (20)	0,71 [0,48; 1,05], p=0,088 0,66 [0,41; 1,07], p=0,089 -5,72% [-12,15%; 0,72%], p=0,082
SWORD-2	261	255	43 (16)	32 (13)	1,33 [0,87; 2,03], p=0,184 1,40 [0,85; 2,30], p=0,188 4,09% [-1,94%; 10,13%], p=0,184

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In SWORD-1 ist die Ereignisrate etwas höher im CAR-Arm als im DTG/RPV-Arm (20 % vs. 14 %), in SWORD-2 ist das Verhältnis umgekehrt (13 % vs. 16 %). Diese beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

Für den vorliegenden Endpunkt liegt hohe Heterogenität vor (I^2 = 78,0 %). Die Ergebnisse zu diesem Endpunkt werden daher für die Einzelstudien interpretiert. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

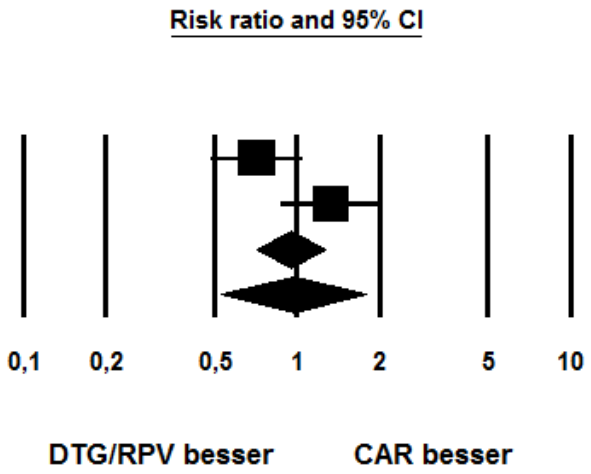
Tabelle 4-74: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=78,0%	
Modell mit festen Effekten	0,95 [0,71 – 1,26] p=0,715
Modell mit zufälligen Effekten	0,97 [0,52 – 1,79] p=0,913
Modell mit zufälligen Effekten[a] (Knapp-Hartung-Methode)	0,96 [0,02 – 47,25] p=0,919
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall [a] basierend auf den nicht-adjustierten Schätzern	

Im meta-analytischen Modell mit zufälligen Effekten zeigt sich kein statistisch signifikanter Unterschied.

Zur Interpretation der Evidenz zu „unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen“ ist zu beachten, dass bei hoher Heterogenität der Meta-Analyse die Ergebnisse der beiden Einzelstudien keinen einheitlichen Trend abzuleiten gestatten. Während in der Studie SWORD-1 ein Trend zugunsten von DTG/RPV zu erkennen ist, ergibt sich ein Trend in die entgegengesetzte Richtung in SWORD-2. Keines der Studienergebnisse erreicht statistische Signifikanz, ebenso liegt für die Meta-Analysen keine statistische Signifikanz vor. Insofern ist diese Situation klarer Weise so zu interpretieren, dass es sich hierbei höchstwahrscheinlich lediglich um zufällige Schwankungen im Rahmen von zwei Einzelstudien handelt und damit in der Gesamtschau keine Aussage zu einem generellen Trend oder gar einem statistisch signifikanten Ergebnis in die eine oder andere Richtung möglich ist.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	0,710	0,480	1,050	0,086
	SWORD2	total	1,330	0,871	2,032	0,187
Fixed			0,948	0,711	1,264	0,715
Random			0,966	0,523	1,788	0,913



Heterogenität: I Quadrat = 78,0%

Abbildung 17: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.16 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems – RCT

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

Tabelle 4-75: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*nervous system disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*nervous system disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-76: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Erkrankungen des Nervensystems“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-77: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	37 (15)	28 (11)	1,36 [0,86;2,14], p=0,193 1,41 [0,83;2,38], p=0,205 3,98% [-1,85%;9,82%], p=0,181
SWORD-2	261	255	40 (15)	14 (5)	2,78 [1,55;4,98], p<0,001 3,09 [1,64;5,84], p<0,001 9,78% [4,59%;14,98%], p<0,001

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien sind die Ereignisraten etwas höher im DTG/RPV-Arm als im CAR-Arm (15 % vs. 11 % bzw. 15 % vs. 5 %). Diese beobachteten Unterschiede sind in SWORD-2 zudem statistisch signifikant (p < 0,001), in SWORD-1 liegt keine statistische Signifikanz vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

Für den vorliegenden Endpunkt liegt hohe Heterogenität vor (I^2 = 72,1 %). Die Ergebnisse zu diesem Endpunkt werden daher auf Einzelstudienebene interpretiert. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-78: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=72,1%	
Modell mit festen Effekten	1,78 [1,25 – 2,55] p=0,002
Modell mit zufälligen Effekten	1,90 [0,94 – 3,82] p=0,072
Modell mit zufälligen Effekten[a] (Knapp-Hartung-Methode)	1,89 [0,02 – 196,23] p=0,331

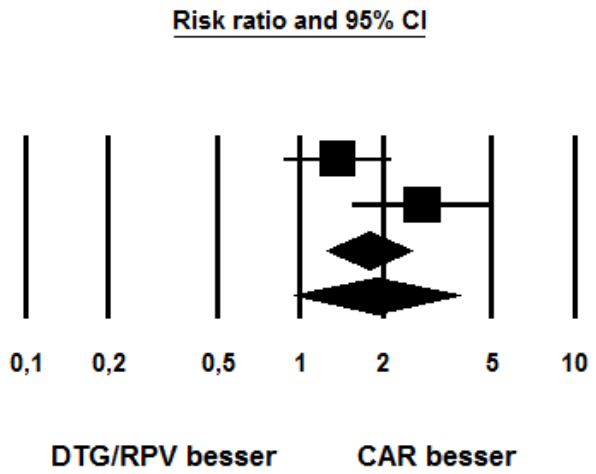
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI= Konfidenzintervall

[a] basierend auf den nicht-adjustierten Schätzern

Im meta-analytischen Modell mit zufälligen Effekten zeigt sich kein statistisch signifikanter Unterschied.

Zur Interpretation der Evidenz zu „unerwünschten Ereignissen von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems“ ist zu beachten, dass bei hoher Heterogenität der Meta-Analyse die Ergebnisse in beiden Einzelstudien ein Trend zu einer höheren Ereignishäufigkeit unter DTG/RPV zu beobachten ist, der jedoch nur in der Studie SWORD-2 statistisch signifikant ist. Indizien, weshalb sich die beiden Studien hinsichtlich dieses Endpunktes unterscheiden, konnten nicht identifiziert werden. Es ist hervorzuheben, dass die beiden Einzelstudien sich sehr dominant durch die nur etwas weniger als halb so hohe relative Häufigkeit von Ereignissen unter CAR in SWORD-2 im Vergleich zu SWORD-1 unterscheiden, wohingegen die Ereignisrate unter DTG/RPV in beiden Studien sehr gut vergleichbar war. Da keine Gründe für diesen Unterschied zu identifizieren waren, sollte jedoch berücksichtigt werden, dass die hohe Heterogenität der Meta-Analyse hauptsächlich auf diesen Unterschied im CAR-Arm zurückzuführen zu sein scheint. Insofern ist diese Situation so zu interpretieren, dass es sich - unter Beachtung dieser Abweichungen zwischen den Studien - zwar um einen Trend zuungunsten von DTG/RPV handelt, dieser jedoch in der Gesamtschau derzeit keine klare Aussage erlaubt. Diese Interpretation wird auch durch das statistisch nicht signifikante Ergebnis der Meta-Analyse mit zufälligen Effekten gestützt, ohne dass diese primär zur Interpretation direkt herangezogen werden sollte.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	1,360	0,862	2,145	0,186
	SWORD2	total	2,780	1,551	4,983	0,001
Fixed			1,783	1,245	2,554	0,002
Random			1,898	0,943	3,818	0,072



Heterogenität: I Quadrat = 72,1%

Abbildung 18: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.17 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen – RCT

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

Tabelle 4-79: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*psychiatric disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*psychiatric disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-80: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Psychiatrische Erkrankungen“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-81: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	34 (13)	18 (7)	1,92 [1,11; 3,30], p=0,019 2,08 [1,14; 3,79], p=0,017 6,40% [1,13%; 11,68%], p=0,017
SWORD-2	261	255	27 (10)	14 (5)	1,83 [0,98; 3,41], p=0,057 1,94 [0,99; 3,80], p=0,053 4,75% [0,08%; 9,43%], p=0,046

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien sind die Ereignisraten etwas höher im DTG/RPV-Arm als im CAR-Arm (13 % vs. 7 % bzw. 10 % vs. 5 %). Diese beobachteten Unterschiede sind in SWORD-1 zudem statistisch signifikant (p = 0,019), in SWORD-2 liegt keine statistische Signifikanz vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

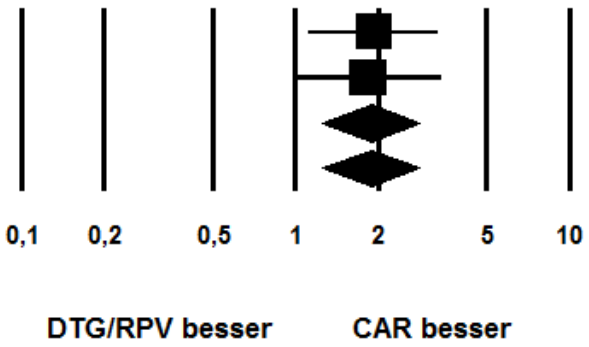
Für den vorliegenden Endpunkt liegt keine Heterogenität vor ($I^2 = 0$ %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-82: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	1,88 [1,25 – 2,83] p=0,003
Modell mit zufälligen Effekten	1,88 [1,25 – 2,83] p=0,003
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Das Ergebnis der Meta-Analyse ist mit einem relativen Risiko von 1,88 zuungunsten von DTG/RPV statistisch signifikant (p = 0,003).

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	1,920	1,114	3,311	0,019
	SWORD2	total	1,830	0,981	3,414	0,057
Fixed			1,881	1,248	2,834	0,003
Random			1,881	1,248	2,834	0,003



Heterogenität: I Quadrat = 0%

Abbildung 19: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.18 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege – RCT

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

Tabelle 4-83: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*renal and urinary disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*renal and urinary disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-84: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunktherheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Erkrankungen der Nieren und Harnwege“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-85: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	5 (2)	5 (2)	1,02 [0,30; 3,47], p=0,980 1,02 [0,29; 3,57], p=0,981 -0,12% [-2,49%; 2,25%], p=0,922
SWORD-2	261	255	10 (4)	4 (2)	2,41 [0,77; 7,60], p=0,132 2,48 [0,77; 8,03], p=0,129 2,34% [-0,46%; 5,15%], p=0,102

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen sowie eigene Berechnungen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien sind die Ereignisraten in beiden Studienarmen ausgeglichen (2 % vs. 2 % bzw. 4 % vs. 2 %), es besteht jeweils kein statistisch signifikanter Unterschied.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

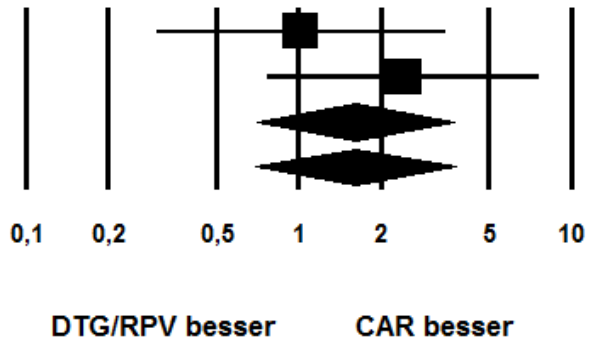
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 1,7 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-86: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=1,7%	
Modell mit festen Effekten	1,61 [0,70 – 3,72] p=0,263
Modell mit zufälligen Effekten	1,61 [0,69 – 3,75] p=0,268
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Im meta-analytischen Modell zeigt sich wie in den Einzelstudien kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	1,016	0,298	3,466	0,980
	SWORD2	total	2,410	0,767	7,571	0,132
Fixed			1,612	0,698	3,724	0,263
Random			1,612	0,693	3,750	0,268



Heterogenität: I Quadrat = 1,7%

Abbildung 20: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.19 Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes – RCT

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

Tabelle 4-87: Operationalisierung von unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes

Studie	Operationalisierung
SWORD-1	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*skin and subcutaneous tissue disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.
SWORD-2	Anzahl der Patienten, für die mind. ein unerwünschtes Ereignis mit der MedDRA *System Organ Class* (SOC) „*skin and subcutaneous tissue disorders*“ im Rahmen der Studie bis Woche 48 gemeldet wurde. Als Effektschätzer für dichotome Endpunkte werden das relative Risiko, Odds Ratio und die Risikodifferenz mit den jeweiligen 95%-Konfidenzintervallen und p-Werten dargestellt. Die Analysen bzw. Darstellungen beruhen jeweils auf dem *Safety Set*.

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

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

Tabelle 4-88: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunktherheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	hoch
SWORD-2	niedrig	nein	ja	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Bei dem Endpunkt „unerwünschte Ereignisse von besonderem Interesse – Erkrankungen der Haut und des Unterhautzellgewebes“ werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diesen Endpunkt gesehen.

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 4-89: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	252	256	30 (12)	22 (9)	1,39 [0,83; 2,34], p=0,215 1,44 [0,81; 2,57], p=0,219 2,98% [-2,26%; 8,23%], p=0,265
SWORD-2	261	255	38 (15)	23 (9)	1,59 [0,98; 2,60], p=0,061 1,72 [0,99; 2,98], p=0,055 5,67% [0,13%; 11,21%], p=0,045

Quellen: Tabellen 3.9 und 3.10 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

In beiden Studien sind die Ereignisraten etwas höher im DTG/RPV-Arm als im CAR-Arm (12 % vs. 9 % bzw. 15 % vs. 9 %). Diese beobachteten Unterschiede sind für das RR nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

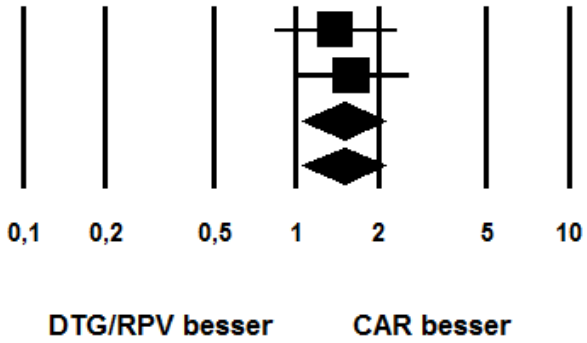
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 0 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-90: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes aus RCT mit dem zu bewertenden Arzneimittel

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	1,49 [1,05 – 2,13] p=0,027
Modell mit zufälligen Effekten	1,49 [1,05 – 2,13] p=0,027
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Anders als in den Einzelstudien ist das Ergebnis der Meta-Analyse mit einem relativen Risiko von 1,49 zuungunsten von DTG/RPV statistisch signifikant (p = 0,027).

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	total	1,390	0,828	2,334	0,213
	SWORD2	total	1,590	0,976	2,590	0,062
Fixed			1,493	1,046	2,129	0,027
Random			1,493	1,046	2,129	0,027



Heterogenität: I Quadrat = 0%

Abbildung 21: Meta-Analyse für unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.20 Knochenmineraldichte zu Woche 48 – RCT

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

Tabelle 4-91: Operationalisierung von Knochenmineraldichte zu Woche 48

Studie	Operationalisierung
SWORD-1	Dieser Endpunkt wurde im Rahmen einer Substudie (DEXA Substudie, s. unten) erhoben.
SWORD-2	Dieser Endpunkt wurde im Rahmen einer Substudie (DEXA Substudie, s. unten) erhoben.
DEXA Substudie	In diese Studien wurden Patienten aus den beiden Zulassungsstudien SWORD-1 und SWORD-2 eingeschlossen. Im Rahmen dieser Substudie wurde keine weitere Studienmedikation verabreicht als derjenigen, die durch die Hauptstudien SWORD-1 und SWORD-2 zugeteilt war. Haupteinschlusskriterium für die Substudie war die Teilnahme an einer der Hauptstudien und der Erhalt von Tenofovirdisoproxilfumarat als Bestandteil der ART. Insgesamt wurden ca. 10% der Patienten aus den beiden Hauptstudien in die Substudie eingeschlossen. Die Knochenmineraldichte wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein höherer Wert entspricht einem besseren Knochenzustand. In diesem Abschnitt werden die Ergebnisse von zwei verschiedenen Lokalisationen beschrieben: Hüfte (inkl. Oberschenkelhals, Trochanter und intertrochantäre Bereiche) und Lendenwirbel (L1 bis L4). Die Knochenmineraldichte wird in g/cm^2 gemessen. Mittels der DEXA-Software wird die für die Hüfte die gesamte zu messende Hüftregion der Knochenmineralgehalt in Gramm gemessen und ins Verhältnis zum gesamten Knochenbereich in cm^2 gesetzt. Analog erfolgt die Messung des Bereichs der Lendenwirbel L1 bis L4. Die DEXA Aufnahmen wurden zentral, verblindet beurteilt. Im Falle von nicht beurteilbaren Aufnahmen mussten diese wiederholt werden. Der zentrale Provider, der für die Beurteilung zuständig war, stellte auch die Kalibrierung der einzelnen DEXA Scanner in den jeweiligen Studienzentren sowie zentrumsübergreifend sicher. Dargestellt wird hier die mittlere prozentuale Änderung der Knochenmineraldichte der Hüfte (Minimum der Einzelmessungen) sowie der Lendenwirbel (Mittelwert der Einzelmessungen) zu Woche 48 gegenüber Baseline. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht und Baseline Knochenmineraldichte.

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

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

Tabelle 4-92: Bewertung des Verzerrungspotenzials für Knochenmineraldichte zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
DEXA-Substudie	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Substudie besitzt basierend auf der Einstufung der beiden Hauptstudien SWORD-1 und SWORD-2 auf Studienebene ebenfalls niedriges Verzerrungspotenzial. Der Endpunkt wurde verblindet von einem zentralen Begutachter analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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

Tabelle 4-93: Patientencharakteristika der DEXA Substudie

Studie 202094 **DEXA Substudie**	**DTG + RPV** **N=53**	**CAR** **N=49**
Geschlecht, n (%)		
männlich	26 (49)	23 (47)
weiblich	27 (51)	26 (53)
Alter, n (%)		
< 35 Jahre	14 (26)	10 (20)
35 bis < 50 Jahre	24 (45)	23 (47)
≥ 50 Jahre	15 (28)	16 (33)
BMI [kg/m^2], Mittelwert ± SD	25,17±3,93	25,79±4,83
Raucherstatus, n (%)		
Nie-/Nichtraucher	40 (75)	36 (74)
< 1 *pack year*[a]	10 (19)	8 (16)
≥ 1 *pack year*[a]	3 (6)	5 (10)

Alkoholkonsum, n (%)		
keiner	37 (70)	30 (61)
< 14 Einheiten/Woche[b]	15 (28)	17 (35)
≥ 14 Einheiten/Woche[b]	1 (2)	2 (4)
Abstammung, n (%)		
hispanisch/lateinamerikanisch	17 (32)	18 (37)
andere	36 (68)	31 (63)
Ethnie, n (%)		
weiß	44 (83)	40 (82)
schwarz/afrikanisch-amerikanisch	3 (6)	7 (14)
asiatisch	2 (4)	1 (2)
andere	4 (8)	1 (2)
Baseline CD4+ Zellzahl [pro mm^3], n (%)		
< 200	4 (8)	1 (2)
200 bis < 350	2 (4)	7 (14)
350 bis < 500	15 (28)	6 (12)
≥ 500	29 (55)	26 (53)
Vitamin D/Kalzium als Begleitmedikation, n (%)		
keines von beidem	38 (72)	37 (76)
Vitamin D und Kalzium	12 (23)	8 (16)
Kalzium	3 (6)	4 (8)
ART zum Screening-Zeitpunkt, n (%)		
NRTI	53 (100)	49 (100)
NNRTI[§]	32 (60)	33 (67)
PI[§]	12 (23)	11 (22)
INI[§]	9 (17)	5 (10)
Therapiedauer vor Studienbeginn [Monate], Median [Min-Max]	44,5 [10-152]	51,8 [11-175]

Abkürzungen: SD = Standardabweichung, ART = Antiretrovirale Therapie

[a] 1 *pack year* ist definiert als 20 Zigaretten, die über den Zeitraum eines Jahres täglich geraucht werden.

[b] Eine Einheit entspricht einem halben Glas Bier, einem Glas Wein oder einem Schnaps.

[§] jeweils zuzüglich weiterer NRTI

Die Patientenkollektive in den beiden Studienarmen sind gut vergleichbar hinsichtlich ihrer Patientencharakteristika. Im Folgenden werden schließlich die Ergebnisse der Knochenmineraldichte dargestellt.

Tabelle 4-94: Ergebnisse für Knochenmineraldichte (g/cm^2) zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – DEXA Substudie

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**
	Baseline		**Woche 48**		**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere prozentuale Änderung zu Woche 48 [%]** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**		
Hüfte (Knochenmineraldichte in g/cm^2)						
DEXA-Substudie						1,29
DTG/RPV	50	0,964±0,146	46	NA	0,012±0,023	[0,27; 2,31]
CAR	40	0,974±0,115	35	NA	0,000±0,017	p=0,014
Lendenwirbel (Knochenmineraldichte in g/cm^2)						
DEXA-Substudie						1,32
DTG/RPV	52	1,063±0,161	46	NA	0,015±0,030	[0,07; 2,57]
CAR	42	1,086±0,150	35	NA	0,000±0,032	p=0,039

Quelle: Tabelle 2.200, 2.240, 2.260 und 2.280 im Studienbericht der DEXA-Substudie

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse, NA = nicht angegeben

ANCOVA: adjustiert für Alter, Baseline BMI und Baseline Knochenmineraldichte

Sowohl für die **Hüfte** als auch für die **Lendenwirbel** liegt ein statistisch signifikanter Effekt zugunsten von DTG/RPV gegenüber CAR vor (p = 0,014 bzw. p = 0,039). Während sich die mittlere Knochenmineraldichte im CAR-Arm kaum verändert, steigt sie im DTG/RPV-Arm deutlich an.

Zusätzlich zur Knochenmineraldichte seien hier noch die Studienergebnisse zu im Studienverlauf aufgetretenen Frakturen erwähnt. Insgesamt wurden in der SOC „Verletzungen, Vergiftungen und durch Eingriffe bedingte Komplikationen“ (und sonst in keiner anderen SOC) in beiden SWORD-Studien 5 Patienten mit Frakturen im Studienverlauf (5 unterschiedliche Lokalisationen) unter DTG/RPV beobachtet und 8 Patienten mit Frakturen im CAR-Arm (sechs unterschiedliche Lokalisationen). Bezüglich des Auftretens von Frakturen ergab sich demnach über 48 Wochen noch kein Unterschied zwischen den Behandlungsgruppen, obwohl die Knochemineraldichte bei den beobachteten statistisch signifikanten Verbesserungen und bei insbesondere bei längerer Behandlungsdauer eine positive Prognose zur Reduktion der Knochenbrüche gestattet.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Da dieser Endpunkt in der einzelnen Substudie der beiden Zulassungsstudien erhoben wurde, kann keine Meta-Analyse durchgeführt werden.

4.3.1.3.21 Knochenspezifischer Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 – RCT

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

Tabelle 4-95: Operationalisierung von knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48

Studie	Operationalisierung
SWORD-1	Die alkalische Phosphatase wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein niedrigerer Wert entspricht eines besseren Knochenzustands. Dargestellt wird hier die mittlere Änderung zu Woche 48 gegenüber Baseline. Die Analysen basieren auf der Safety Population. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.
SWORD-2	Die alkalische Phosphatase wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein niedrigerer Wert entspricht eines besseren Knochenzustands. Dargestellt wird hier die mittlere Änderung zu Woche 48 gegenüber Baseline. Die Analysen basieren auf der Safety Population. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.

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

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

Tabelle 4-96: Bewertung des Verzerrungspotenzials für knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde verblindet in einem zentralen Labor analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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 4-97: Ergebnisse für knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
SWORD-1						-3,97	-1,09
DTG/RPV	251	15,56±6,06	235	12,36±4,21	-2,89±4,02	[-4,63; -3,31]	[-1,28; -0,90]
CAR	255	15,52±5,78	245	16,72±6,22	0,90±4,13	p<0,001	p<0,001
SWORD-2						-4,23	-0,95
DTG/RPV	260	16,25±7,22	246	13,11±4,92	-3,18±5,68	[-5,03; -3,43]	[-1,14; -0,76]
CAR	250	16,66±7,15	240	17,46±7,57	0,92±4,63	p<0,001	p<0,001

Quelle: Tabelle 3.13, 3.14, 3.19 und 3.20 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse

ANCOVA: adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker

In beiden Studien ist die Veränderung zu Woche 48 stärker zugunsten von DTG/RPV ausgeprägt. Die beobachteten Unterschiede sind auch in beiden Fällen statistisch signifikant ($p < 0,001$).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

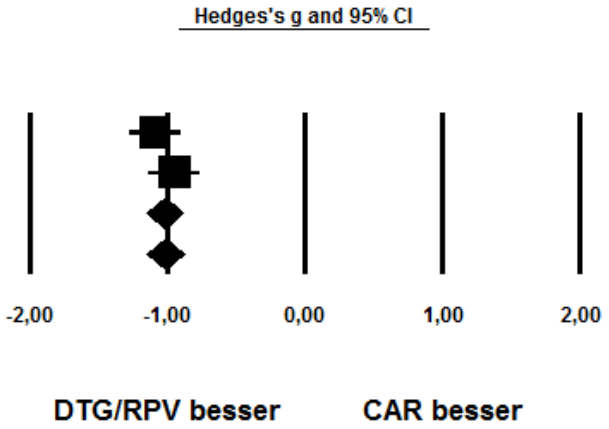
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 0 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-98: Meta-Analyse für knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	-1,020 [-1,154; -0,886] p<0,001
Modell mit zufälligen Effekten	-1,020 [-1,156; -0,884] p<0,001
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Das Ergebnis der Meta-Analyse bestätigt den deutlichen Effekt aus den Einzelstudien und zeigt einen statistisch signifikanten Unterschied zugunsten von DTG/RPV gegenüber CAR (p < 0,001).

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	-1,091	-1,283	-0,899	0,000
	SWORD-2	total	-0,952	-1,140	-0,764	0,000
Fixed			-1,020	-1,154	-0,886	0,000
Random			-1,020	-1,156	-0,884	0,000



Heterogenität: I Quadrat = 2,6%

Abbildung 22: Meta-Analyse für knochenspezifischen Biomarker: alkalische Phosphatase (BSAP) zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.22 Knochenspezifischer Biomarker: Prokollagen 1-N-terminales Propeptid (P1NP) zu Woche 48 – RCT

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

Tabelle 4-99: Operationalisierung von Knochenspezifischen Biomarker: P1NP zu Woche 48

Studie	Operationalisierung
SWORD-1	P1NP wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein niedrigerer Wert entspricht eines besseren Knochenzustands. Dargestellt wird hier die mittlere Änderung zu Woche 48 gegenüber Baseline. Die Analysen basieren auf der Safety Population. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.
SWORD-2	P1NP wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein niedrigerer Wert entspricht eines besseren Knochenzustands. Dargestellt wird hier die mittlere Änderung zu Woche 48 gegenüber Baseline. Die Analysen basieren auf der Safety Population. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.

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

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

Tabelle 4-100: Bewertung des Verzerrungspotenzials für Knochenspezifischen Biomarker: P1NP zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde verblindet in einem zentralen Labor analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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 4-101: Ergebnisse für Knochenspezifischen Biomarker: P1NP zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
SWORD-1						-9,70	-0,64
DTG/RPV	252	51,7±23,42	234	41,7±16,43	-9,1±20,34	[-12,45; -6,94]	[-0,82; -0,46]
CAR	253	54,6±21,39	245	53,0±19,47	-1,4±18,95	p<0,001	p<0,001
SWORD-2						-6,95	-0,42
DTG/RPV	260	54,2±22,87	246	48,2±20,24	-5,8±20,00	[-9,91; -3,99]	[-0,61; -0,24]
CAR	249	56,0±24,28	242	55,9±21,82	0,3±19,28	p<0,001	p<0,001

Quelle: Tabelle 3.15, 3.16, 3.19 und 3.20 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse

ANCOVA: adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker

In beiden Studien ist die Veränderung zu Woche 48 stärker zugunsten von DTG/RPV ausgeprägt. Die beobachteten Unterschiede sind auch in beiden Studien jeweils statistisch signifikant und zwar sowohl bezüglich der absoluten mittleren Änderung als auch Hedges' g ($p < 0,001$).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

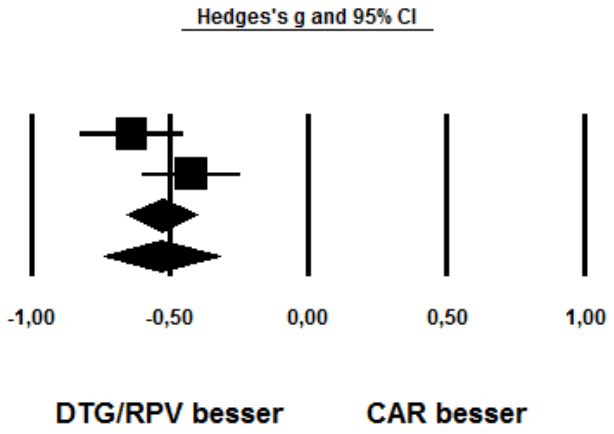
Für den vorliegenden Endpunkt liegt mäßige Heterogenität vor (I^2 = 62,6 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-102: Meta-Analyse für Knochenspezifischen Biomarker: P1NP zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=62,6%	
Modell mit festen Effekten	-0,529 [-0,658; -0,400] p<0,001
Modell mit zufälligen Effekten	-0,531 [-0,741; -0,320] p<0,001
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Das Ergebnis der Meta-Analyse bestätigt den deutlichen Effekt aus den Einzelstudien und zeigt einen statistisch signifikanten Unterschied in Hedges‘ g zugunsten von DTG/RPV gegenüber CAR (p < 0,001).

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	-0,639	-0,823	-0,455	0,000
	SWORD-2	total	-0,424	-0,604	-0,244	0,000
Fixed			-0,529	-0,658	-0,400	0,000
Random			-0,531	-0,741	-0,320	0,000



Heterogenität: I Quadrat = 62,6%

Abbildung 23: Meta-Analyse für Knochenspezifischen Biomarker: P1NP zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.23 Knochenspezifischer Biomarker: type 1 collagen cross-linked C-telopeptide (CTx) zu Woche 48 – RCT

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

Tabelle 4-103: Operationalisierung von Knochenspezifischen Biomarker: CTx zu Woche 48

Studie	Operationalisierung
SWORD-1	CTx wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein niedrigerer Wert entspricht eines besseren Knochenzustands. Dargestellt wird hier die mittlere Änderung zu Woche 48 gegenüber Baseline. Die Analysen basieren auf der Safety Population. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.
SWORD-2	CTx wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein niedrigerer Wert entspricht eines besseren Knochenzustands. Dargestellt wird hier die mittlere Änderung zu Woche 48 gegenüber Baseline. Die Analysen basieren auf der Safety Population. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.

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

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

Tabelle 4-104: Bewertung des Verzerrungspotenzials für Knochenspezifischen Biomarker: CTx zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde verblindet in einem zentralen Labor analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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 4-105: Ergebnisse für Knochenspezifischen Biomarker: CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
SWORD-1						-0,163	-0,197
DTG/RPV	252	0,643±0,355	234	0,445±0,262	-0,182±0,307	[-0,13; -0,01]	[-0,38; -0,02]
CAR	254	0,675±0,346	243	0,633±1,149	-0,044±1,160	p=0,034	p=0,033
SWORD-2						-0,085	-0,307
DTG/RPV	260	0,670±0,319	243	0,520±0,336	-0,149±0,313	[-0,14; -0,04]	[-0,49; -0,13]
CAR	250	0,709±0,387	242	0,626±0,331	-0,087±0,344	p<0,001	p<0,001

Quelle: Tabelle 3.17, 3.18, 3.19 und 3.20 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse

ANCOVA: adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker

In beiden Studien ist die Veränderung zu Woche 48 stärker zugunsten von DTG/RPV ausgeprägt. Die beobachteten Unterschiede sind auch in beiden Studien statistisch signifikant ($p = 0{,}033$ bzw. $p < 0{,}001$).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

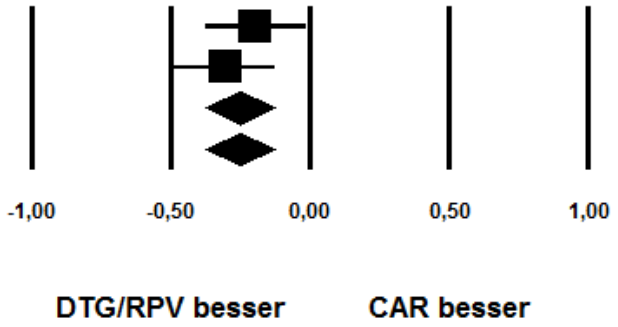
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 0 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-106: Meta-Analyse für Knochenspezifischen Biomarker: CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,252 [-0,380; -0,124] p<0,001
Modell mit zufälligen Effekten	-0,252 [-0,380; -0,124] p<0,001
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall	

Das Ergebnis der Meta-Analyse bestätigt den deutlichen Effekt aus den Einzelstudien und zeigt einen statistisch signifikanten Unterschied in Hedges‘ g zugunsten von DTG/RPV gegenüber CAR (p < 0,001).

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	-0,197	-0,377	-0,017	0,032
	SWORD-2	total	-0,307	-0,487	-0,127	0,001
Fixed			-0,252	-0,380	-0,124	0,000
Random			-0,252	-0,380	-0,124	0,000



Heterogenität: I Quadrat = 0%

Abbildung 24: Meta-Analyse für Knochenspezifischen Biomarker: CTx zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.24 Knochenspezifischer Biomarker: Osteocalcin zu Woche 48 – RCT

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

Tabelle 4-107: Operationalisierung von Knochenspezifischen Biomarker: Osteocalcin zu Woche 48

Studie	Operationalisierung
SWORD-1	Osteocalcin wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein niedrigerer Wert entspricht eines besseren Knochenzustands. Dargestellt wird hier die mittlere Änderung zu Woche 48 gegenüber Baseline. Die Analysen basieren auf der Safety Population. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.
SWORD-2	Osteocalcin wurde jeweils zu Baseline und zu Woche 48 gemessen. Ein niedrigerer Wert entspricht eines besseren Knochenzustands. Dargestellt wird hier die mittlere Änderung zu Woche 48 gegenüber Baseline. Die Analysen basieren auf der Safety Population. Fehlende Werte wurden nicht ersetzt. Für die mittleren Änderungen wurden Kovarianzanalysen (ANCOVA) durchgeführt, adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker. Zusätzlich wird Hedges‘ g als standardisiertes Effektmaß angegeben.

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

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

Tabelle 4-108: Bewertung des Verzerrungspotenzials für Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SWORD-1	niedrig	nein	ja	ja	ja	niedrig
SWORD-2	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die beiden Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial. Der Endpunkt wurde verblindet in einem zentralen Labor analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

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

Tabelle 4-109: Ergebnisse für Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Safety Population

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
SWORD-1						-4,102	-0,723
DTG/RPV	251	22,87±9,54	234	18,06±6,82	-4,40±7,61	[-5,13; -3,07]	[-0,91; -0,54]
CAR	253	23,42±8,39	245	22,69±7,65	-0,68±6,58	p<0,001	p<0,001
SWORD-2						-4,015	-0,745
DTG/RPV	260	24,66±9,89	245	19,42±7,37	-5,11±7,33	[-4,99; -3,04]	[-0,93; -0,56]
CAR	248	24,50±9,53	241	23,33±8,68	-1,14±6,02	p<0,001	p<0,001

Quelle: Tabelle 3.19, 3.20, 3.27 und 3.28 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse

ANCOVA: adjustiert für Alter, Geschlecht, BMI, Raucherstatus, *Baseline Third Agent* und Baseline Biomarker

In beiden Studien ist die Veränderung zu Woche 48 stärker zugunsten von DTG/RPV ausgeprägt. Die beobachteten Unterschiede sind auch in beiden Studien und hinsichtlich der absoluten mittleren Änderung und im Hedges' g statistisch signifikant ($p < 0,001$).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Beide eingeschlossenen Studien (SWORD-1 und SWORD-2) sind homogen bzgl. Studiendesign, Patientenkollektiv und Methodik. Eine meta-analytische Zusammenfassung ist daher gerechtfertigt.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist gegeben, da keine Hinweise existieren, dass die Studienkollektive bzgl. ihrer Patientencharakteristika dem widersprechen würden. Das zu bewertende Arzneimittel sowie die Therapien der zweckmäßigen Vergleichstherapie wurden jeweils gemäß ihres bestimmungsgemäßen Gebrauchs verabreicht. Die Studienergebnisse können somit auf den deutschen Versorgungskontext übertragen werden.

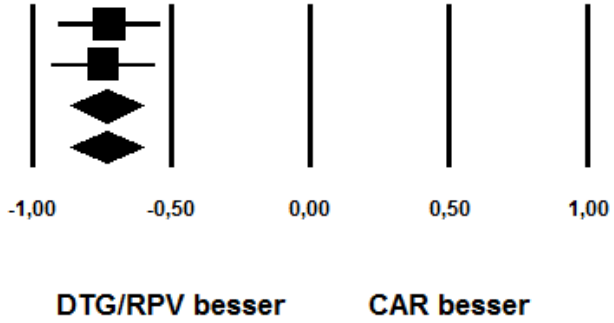
Für den vorliegenden Endpunkt liegt keine Heterogenität vor (I^2 = 0 %). Die Ergebnisse der Einzelstudien können daher meta-analytisch anhand des Modells mit festen Effekten kombiniert und interpretiert werden. In Folgenden werden zunächst die meta-analytischen Schätzwerte tabellarisch präsentiert und anschließend grafisch mittels eines Forest-Plots dargestellt.

Tabelle 4-110: Meta-Analyse für Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,734 [-0,866; -0,602] p<0,001
Modell mit zufälligen Effekten	-0,734 [-0,866; -0,602] p<0,001
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI= Konfidenzintervall	

Das Ergebnis der Meta-Analyse bestätigt den deutlichen Effekt aus den Einzelstudien und zeigt einen statistisch signifikanten Unterschied in Hedges‘ g zugunsten von DTG/RPV gegenüber CAR (p < 0,001).

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	total	-0,723	-0,909	-0,537	0,000
	SWORD-2	total	-0,745	-0,931	-0,559	0,000
Fixed			-0,734	-0,866	-0,602	0,000
Random			-0,734	-0,866	-0,602	0,000



Heterogenität: I Quadrat = 0%

Abbildung 25: Meta-Analyse für Knochenspezifischen Biomarker: Osteocalcin zu Woche 48 aus RCT; DTG/RPV versus CAR

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.25 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. 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.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch 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.

Es wurden für beide Studien, SWORD-1 und SWORD-2, folgende Subgruppenfaktoren analysiert:

- Alter (< 50 Jahre / ≥ 50 Jahre)
- Geschlecht (männlich / weiblich)
- Ethnizität (weiß / nicht-weiß)
- Region (Nordamerika / Europa / Sonstige)
- Baseline CD4+ Zellzahl (< 500 / ≥ 500 Zellen pro mm^3)
- *Baseline Third Agent* (PI / NNRTI / INI)

Diese Subgruppen waren mit Ausnahme der Regionen *a priori* im statistischen Analyseplan festgelegt, „Alter“ und „*Baseline Third Agent*“ waren zudem stratifizierende Randomisierungsvariablen. Der Subgruppenfaktor „Region“ wurde gemäß den Dossieranforderungen zusätzlich mit in die Analysen aufgenommen. Die Baseline CD4+ Zellzahl gilt als Faktor für den Schweregrad der Erkrankung.

Subgruppenergebnisse wurden für alle Endpunkte mit Ausnahme solcher, für die nur sehr wenige Ereignisse vorliegen (weniger als 10 Ereignisse in einer Subgruppe, siehe (26)), dargestellt. Auch für den Endpunkt Knochenmineraldichte, der ausschließlich in der DEXA-Substudie erhoben wurde, werden wegen der insgesamt kleinen Patientenzahl in dieser Substudie keine Subgruppenanalysen dargestellt. Im zugehörigen Studienbericht sind Subgruppenanalysen enthalten (42). Die Subgruppenergebnisse werden in diesem Abschnitt berichtet, falls im entsprechenden Fall ein Beleg für Effektmodifikation vorliegt, d.h. der

Interaktionstest einen p-Wert von < 0,05 zeigt (26). Die p-Werte der Interaktionstest für jeden untersuchten Subgruppenfaktor (pro Endpunkt) werden jedoch vollständig im Folgenden angegeben.

Wegen des Vorhandenseins von zwei bewertungsrelevanten RCT (SWORD-1 und SWORD-2) wird für die Interaktionstests die Q-Statistik der Meta-Analyse herangezogen und diese im Falle von signifikanten Interaktions-p-Werten berichtet. Zusätzlich sind in Tabelle 4-111 die p-Werte der Interaktionstests pro Endpunkt auf Einzelstudienebene zusammengetragen, die einen Überblick geben, welche Subgruppenergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet werden – zusätzlich zu den o.g. Meta-Analysen, die die Grundlage für die Interpretation des Zusatznutzens sind. Es erfolgt keine Interpretation dieser Subgruppenergebnisse der Einzelstudien für die Fälle, in denen keine meta-analytisch signifikanten Interaktionen vorliegen.

Tabelle 4-111: p-Werte der Interaktionstests pro Endpunkt auf Einzelstudienebene

Endpunkt	**Studie**	**Interaktions-p-Werte**					
		Alter	**Geschlecht**	**Ethnizität**	**Region**	**Baseline CD4+ Zellzahl**	**Baseline Third Agent**
Viruslast < 50 Kopien/ml zu Woche 48	SWORD-1	p=0,262	p=0,928	p=0,740	p=0,868	p=0,444	p=0,444
	SWORD-2	p=0,107	p=0,669	p=0,056	p=0,207	p=0,417	p=0,590
EQ-5D zu Woche 48 Utility Score	SWORD-1	p=0,163	p=0,276	p=1,000	p=0,584	p=0,262	p=0,052
	SWORD-2	p=0,129	p=0,126	p=0,081	p=0,225	p=0,453	p=0,698
EQ-5D zu Woche 48 VAS	SWORD-1	p=0,303	p=0,628	p=0,471	p=0,956	p=0,131	p=0,231
	SWORD-2	p=0,433	p=0,558	p=0,937	p=0,204	p=0,151	p=0,741
Symptom Distress Module zu Woche 48 Symptom Count Score	SWORD-1	p=0,156	**p=0,036**	p=0,741	p=0,617	p=0,334	p=0,551
	SWORD-2	p=0,085	p=0,855	p=0,405	p=0,647	p=0,527	p=0,947
Symptom Distress Module zu Woche 48 Symptom Bother Score	SWORD-1	p=0,087	p=0,093	p=0,902	p=0,586	p=0,531	p=0,588
	SWORD-2	p=0,162	p=0,897	p=0,482	p=0,636	p=0,381	p=0,896
HIVTSQ Total Score zu Woche 48	SWORD-1	p=0,121	p=0,222	p=0,544	p=0,940	p=0,466	p=0,506
	SWORD-2	p=0,707	p=0,163	p=0,388	p=0,929	p=0,068	p=0,838
Unerwünschte Ereignisse bis Woche 48	SWORD-1	p=0,143	p=0,938	p=0,742	p=0,101	p=0,749	p=0,594
	SWORD-2	p=0,353	p=0,324	p=0,086	p=0,565	p=0,651	p=0,978
Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48	SWORD-1	p=0,991	p=0,145	p=0,676	p=0,144	**p=0,041**	p=0,250
	SWORD-2	p=0,575	p=0,442	p=0,466	p=0,378	**p=0,040**	p=0,385
Schwerwiegende unerwünschte Ereignisse bis Woche 48	SWORD-1	p=0,486	p=0,085	p=0,942	p=0,476	p=0,123	p=0,070
	SWORD-2	p=0,477	p=0,878	p=0,899	p=0,779	p=0,054	p=0,634
Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten	SWORD-1	p=0,690	p=0,373	p=1,000	p=0,681	p=0,524	p=0,711
	SWORD-2	p=0,844	p=0,286	p=0,185	p=0,884	p=0,112	p=0,966

Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen[a]	SWORD-1	-	-	-	-	-	-
	SWORD-2	-	-	-	-	-	-
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts	SWORD-1	p=0,485	p=0,526	p=0,757	p=0,513	p=0,420	p=0,584
	SWORD-2	p=0,373	p=0,705	p=0,454	p=0,173	p=0,168	p=0,664
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen[a]	SWORD-1	-	-	-	-	-	-
	SWORD-2	-	-	-	-	-	-
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität[a]	SWORD-1	-	-	-	-	-	-
	SWORD-2	-	-	-	-	-	-
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen	SWORD-1	p=0,364	p=0,326	p=0,450	p=1,000	p=0,910	p=0,136
	SWORD-2	p=0,655	p=0,390	p=0,963	p=0,993	p=0,967	p=0,654
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	SWORD-1	p=0,777	p=0,556	p=0,527	p=0,727	p=0,922	p=0,763
	SWORD-2	p=0,986	p=0,962	p=0,093	p=0,434	p=0,933	p=0,736
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems	SWORD-1	p=0,648	p=0,250	p=0,176	p=0,882	p=0,686	p=0,199
	SWORD-2	p=0,330	p=0,702	p=0,259	p=0,187	p=0,655	p=0,317
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen	SWORD-1	**p=0,040**	p=0,579	p=0,469	p=0,294	p=0,524	p=0,833
	SWORD-2	p=0,778	p=0,976	p=0,142	p=0,311	p=0,939	p=0,552
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege	SWORD-1	p=0,526	p=0,389	p=0,867	p=0,375	p=0,646	p=0,758
	SWORD-2	p=0,855	p=1,000	p=0,701	p=0,410	p=0,824	p=0,543
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes	SWORD-1	p=0,729	p=0,651	p=0,590	p=0,384	p=0,064	p=0,321
	SWORD-2	p=0,919	p=0,749	p=0,545	p=0,188	p=0,184	p=0,330
Knochenspezifischer Biomarker: BSAP zu Woche 48	SWORD-1	p=0,367	p=0,635	p=0,309	p=0,742	p=0,547	p=0,112
	SWORD-2	p=0,692	p=0,216	p=0,081	p=0,587	p=0,791	p=0,441
Knochenspezifischer Biomarker: P1NP zu Woche 48	SWORD-1	p=0,783	p=0,187	p=0,735	p=0,225	p=0,174	p=0,832
	SWORD-2	**p=0,026**	p=0,098	p=0,519	p=0,135	p=0,155	p=0,703

Knochenspezifischer Biomarker: CTx zu Woche 48	SWORD-1	p=0,319	p=0,114	p=0,405	p=0,463	p=0,135	**p=0,047**
	SWORD-2	p=0,608	p=0,229	p=0,589	p=0,604	**p=0,007**	p=0,553
Knochenspezifischer Biomarker: Osteocalcin zu Woche 48	SWORD-1	p=0,109	p=0,443	p=0,978	p=0,457	p=0,773	**p=0,001**
	SWORD-2	**p=0,023**	p=0,445	p=0,358	p=0,987	p=0,186	**p=0,044**

Quelle: eigene Berechnungen
[a] Aufgrund der Seltenheit der Ereignisse wurden für diesen Endpunkt keine Subgruppenanalysen durchgeführt.

4.3.1.3.25.1 Subgruppenanalysen: Viruslast < 50 Kopien/ml zu Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-112: p-Werte der meta-analytischen Interaktionstests für Viruslast < 50 Kopien/ml zu Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,057
Geschlecht	p=0,639
Ethnizität	p=0,256
Region	p=0,731
Baseline CD4+ Zellzahl	p=0,985
Baseline Third Agent	p=0,903

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.2 Subgruppenanalysen: EQ-5D zu Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-113: p-Werte der meta-analytischen Interaktionstests für EQ-5D zu Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**	
	Utility Score	**VAS**
Alter	p=0,955	p=0,896
Geschlecht	p=0,817	p=0,498
Ethnizität	p=0,360	p=0,631
Region	p=0,621	p=0,626
Baseline CD4+ Zellzahl	p=0,833	p=0,036
Baseline Third Agent	p=0,165	p=0,676

Quelle: eigene Berechnungen

Für den EQ-5D Utility Score werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen. Für EQ-5D VAS liegt bzgl. des Subgruppenfaktors der Baseline CD4+ Zellzahl ein Beleg für Effektmodifikation vor (p = 0,036). Diese Ergebnisse auf Subgruppenebene werden nachfolgend dargestellt.

Tabelle 4-114: Ergebnisse für EQ-5D VAS zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline CD4+ Zellzahl

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
Baseline CD4+ Zellzahl < 500							
SWORD-1 DTG/RPV CAR	 85 74	 87,69±11,42 87,77±13,83	 85 74	 92,34±11,44 89,64±11,45	 4,65±13,15 1,87±11,12	2,74 [-0,41; 5,89] p=0,088	0,27 [-0,04; 0,59] p=0,088
SWORD-2 DTG/RPV CAR	 80 74	 90,90±10,35 91,45±8,35	 79 73	 92,51±10,56 91,95±10,40	 1,66 ±7,82 0,48±10,70	0,95 [-1,83; 3,73] p=0,500	0,11 [-0,21; 0,43] p=0,499
Baseline CD4+ Zellzahl ≥ 500							
SWORD-1 DTG/RPV CAR	 167 181	 88,30±10,51 87,48±12,07	 167 181	 91,37±9,12 91,18±10,20	 3,07±10,47 3,70±10,73	-0,14 [-1,92; 1,64] p=0,878	-0,02 [-0,23; 0,20] p=0,878
SWORD-2 DTG/RPV CAR	 181 178	 89,44±10,15 85,83±11,26	 178 180	 90,52±11,75 90,78±9,09	 1,07±11,30 4,84±11,80	-1,62 [-3,63; 0,39] p=0,115	-0,17 [-0,38; 0,04] p=0,115

Quelle: Tabellen 6.7, 6.8, 6.11 und 6.12 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse, VAS = Visuelle Analogskala

ANCOVA: adjustiert für Geschlecht, Ethnizität, Alter, *Baseline Third Agent* und Baseline EQ-5D

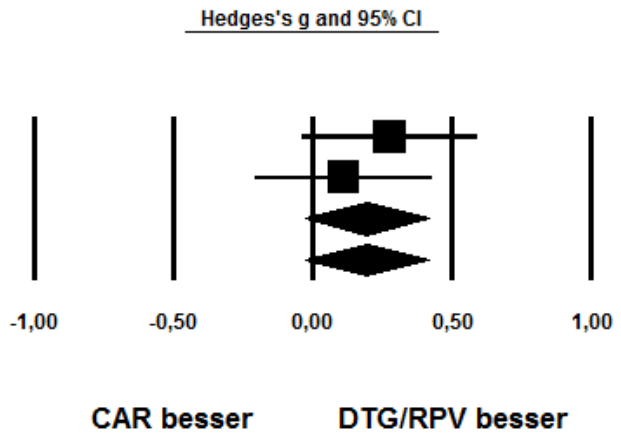
Während die Werte des mittleren Scores bei den Patienten mit niedriger Baseline CD4+ Zellzahl leichte Vorteile im DTG/RPV-Arm zeigen, liegt der Effekt bei den Patienten mit höherer Baseline CD4+ Zellzahl eher zugunsten des CAR-Arms. Diese numerischen Trends sind in beiden Studien ähnlich, es liegen jedoch keine statistischen Signifikanzen vor.

Tabelle 4-115: Meta-Analyse für EQ-5D VAS zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Baseline CD4+ Zellzahl

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
Baseline CD4+ Zellzahl < 500	
Heterogenität: I^2=0%	
Modell mit festen Effekten	0,193 [-0,030 – 0,416] p=0,090
Modell mit zufälligen Effekten	0,193 [-0,030 – 0,416] p=0,090
Baseline CD4+ Zellzahl ≥ 500	
Heterogenität: I^2=0,5%	
Modell mit festen Effekten	-0,093 [-0,241 – 0,054] p=0,216
Modell mit zufälligen Effekten	-0,093 [-0,241 – 0,055] p=0,217
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall, VAS = visuelle Analogskala	

Numerisch unterscheiden sich die Effektrichtungen zwischen den beiden Strata hinsichtlich der Baseline CD4-Zellzahl. Für beide Subgruppen liegt jeweils kein statistisch signifikanter Effekt in der Meta-Analyse vor.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	CD4+ < 500	0,274	-0,040	0,588	0,087
	SWORD-2	CD4+ < 500	0,110	-0,208	0,428	0,497
Fixed			0,193	-0,030	0,416	0,090
Random			0,193	-0,030	0,416	0,090

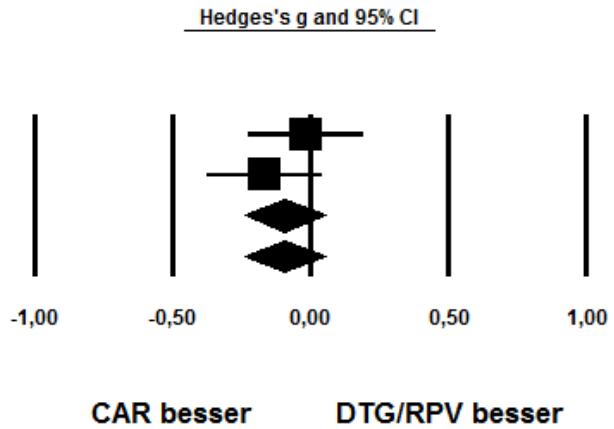


Heterogenität: I Quadrat = 0%

Abbildung 26: Meta-Analyse für EQ-5D VAS zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl < 500

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	CD4+ >= 500	-0,017	-0,227	0,193	0,874
	SWORD-2	CD4+ >= 500	-0,168	-0,376	0,040	0,113
Fixed			-0,093	-0,241	0,054	0,216
Random			-0,093	-0,241	0,055	0,217



Heterogenität: I Quadrat = 0,5%

Abbildung 27: Meta-Analyse für EQ-5D VAS zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl ≥ 500

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.25.3 Subgruppenanalysen: Symptom Distress Module zu Woche 48 - RCT

Für die Studie SWORD-1 liegt für den Subgruppenfaktor Geschlecht eine statistisch signifikante Interaktion vor (p=0,036). Die Ergebnisse werden in nachfolgender Tabelle dargestellt.

Tabelle 4-116: Ergebnisse für SDM Symptom Count Score zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Geschlecht

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
männlich							
SWORD-1 DTG/RPV CAR	 166 164	 5,3±4,99 5,5±4,84	 167 166	 4,7±4,93 6,0±5,33	 -0,6±5,41 0,4±4,32	-1,15 [-2,10; -0,20] p=0,018	-0,26 [-0,48; -0,05] p=0,018
weiblich							
SWORD-1 DTG/RPV CAR	 48 38	 5,8±4,88 7,4±6,59	 48 38	 6,1±6,13 5,8±5,68	 0,3±5,88 -1,6±4,86	1,23 [-0,93; 3,39] p=0,259	0,25 [-0,19; 0,68] p=0,257

Quelle: Tabelle 6.13 und 6.25 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse

ANCOVA: adjustiert für Baseline Symptom Count Score

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-117: p-Werte der meta-analytischen Interaktionstests für Symptom Distress Module zu Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**	
	Symptom Count Score	**Symptom Bother Score**
Alter	p=0,895	p=0,886
Geschlecht	p=0,294	p=0,305
Ethnizität	p=0,718	p=0,668
Region	p=0,996	p=0,983
Baseline CD4+ Zellzahl	p=0,839	p=0,281
Baseline Third Agent	p=0,793	p=0,909

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse der Meta-Analyse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.4 Subgruppenanalysen: HIVTSQ Total Score zu Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-118: p-Werte der meta-analytischen Interaktionstests für HIVTSQ Total Score zu Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,176
Geschlecht	p=0,064
Ethnizität	p=0,921
Region	p=0,961
Baseline CD4+ Zellzahl	p=0,581
Baseline Third Agent	p=0,463

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.5 Subgruppenanalysen: unerwünschte Ereignisse bis Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-119: p-Werte der meta-analytischen Interaktionstests für unerwünschte Ereignisse bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,731
Geschlecht	p=0,542
Ethnizität	p=0,456
Region	p=0,138
Baseline CD4+ Zellzahl	p=0,943
Baseline Third Agent	p=0,701

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.6 Subgruppenanalysen: unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 - RCT

Für beide Studien, SWORD-1 und SWORD-2, liegt für den Subgruppenfaktor Baseline CD4+ Zellzahl jeweils eine statistisch signifikante Interaktion vor (p=0,041 bzw. p=0,040). Die Ergebnisse werden in nachfolgender Tabelle dargestellt.

Tabelle 4-120: Ergebnisse für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline CD4+ Zellzahl

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
Baseline CD4+ Zellzahl < 500					
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	85	74	8 (9)	2 (3)	3,48 [0,76; 15,89], p=0,107 4,12 [0,83; 20,45], p=0,083 7,72% [0,28%; 15,16%], p=0,042
SWORD-2	80	75	2 (3)	3 (4)	0,69 [0,12; 3,94], p=0,673 0,68 [0,11; 4,24], p=0,677 -1,25% [-7,09%; 4,60%], p=0,676
Baseline CD4+ Zellzahl ≥ 500					
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	167	182	7 (4)	13 (7)	0,56 [0,23; 1,35], p=0,193 0,54 [0,21; 1,40], p=0,207 -2,85% [-7,64%; 1,95%], p=0,245
SWORD-2	181	180	15 (8)	2 (1)	7,46 [1,73; 32,15], p=0,007 8,35 [1,87; 37,31], p=0,005 7,19% [2,94%; 11,44%], p<0,001

Quellen: Tabellen 3.3 und 3.4 in den Zusatzanalysen sowie eigene Berechnungen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

Effektschätzer adjustiert für Alter und Baseline Third Agent (Ausnahme: RR für die Fälle, in denen das statistische Modell nicht konvergiert)

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-121: p-Werte der meta-analytischen Interaktionstests für unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,703
Geschlecht	p=0,338
Ethnizität	p=0,785
Region	p=0,283
Baseline CD4+ Zellzahl	p=0,922
Baseline Third Agent	p=0,104

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse für die Meta-Analyse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.7 Subgruppenanalysen: schwerwiegende unerwünschte Ereignisse bis Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-122: p-Werte der meta-analytischen Interaktionstests für schwerwiegende unerwünschte Ereignisse bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,450
Geschlecht	p=0,495
Ethnizität	p=0,953
Region	p=0,481
Baseline CD4+ Zellzahl	p=0,915
Baseline Third Agent	p=0,340

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.8 Subgruppenanalysen: Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-123: p-Werte der meta-analytischen Interaktionstests für Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,618
Geschlecht	p=0,194
Ethnizität	p=0,369
Region	p=0,941
Baseline CD4+ Zellzahl	p=0,622
Baseline Third Agent	p=0,748

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.9 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Herzerkrankungen bis Woche 48 - RCT

Aufgrund der Seltenheit dieses Ereignisses wurden keine Subgruppenanalysen für diesen Endpunkt durchgeführt

4.3.1.3.25.10 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Erkrankungen des Gastrointestinaltrakts bis Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-124: p-Werte der meta-analytischen Interaktionstests für Erkrankungen des Gastrointestinaltrakts bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,497
Geschlecht	p=0,715
Ethnizität	p=0,431
Region	p=0,511
Baseline CD4+ Zellzahl	p=0,294
Baseline Third Agent	p=0,430

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.11 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Leber- und Gallenerkrankungen bis Woche 48 - RCT

Aufgrund der Seltenheit dieses Ereignisses wurden keine Subgruppenanalysen für diesen Endpunkt durchgeführt

4.3.1.3.25.12 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Hypersensitivität Woche 48 - RCT

Aufgrund der Seltenheit dieses Ereignisses wurden keine Subgruppenanalysen für diesen Endpunkt durchgeführt.

4.3.1.3.25.13 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Stoffwechsel- und Ernährungsstörungen bis Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-125: p-Werte der meta-analytischen Interaktionstests für Stoffwechsel- und Ernährungsstörungen bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,946
Geschlecht	p=0,207
Ethnizität	p=0,687
Region	p=0,997
Baseline CD4+ Zellzahl	p=0,898
Baseline Third Agent	p=0,374

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.14 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen bis Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-126: p-Werte der meta-analytischen Interaktionstests für Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,886
Geschlecht	p=0,929
Ethnizität	p=0,368
Region	p=0,600
Baseline CD4+ Zellzahl	p=0,976
Baseline Third Agent	p=0,986

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.15 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Erkrankungen des Nervensystems bis Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-127: p-Werte der meta-analytischen Interaktionstests für Erkrankungen des Nervensystems bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,500
Geschlecht	p=0,419
Ethnizität	p=0,967
Region	p=0,537
Baseline CD4+ Zellzahl	p=0,973
Baseline Third Agent	p=0,345

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.16 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: psychiatrische Erkrankungen bis Woche 48 - RCT

Für die Studie SWORD-1 liegt für den Subgruppenfaktor Alter eine statistisch signifikante Interaktion vor (p=0,040). Die Ergebnisse werden in nachfolgender Tabelle dargestellt.

Tabelle 4-128: Ergebnisse für unerwünschte Ereignisse von besonderem Interesse: psychiatrische Erkrankungen bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Alter

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
Alter < 50 Jahre					
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	183	185	22 (12)	17 (9)	1,32 [0,73; 2,41], p=0,360 1,37 [0,70; 2,67], p=0,360 2,94% [-3,35%; 9,23%], p=0,359
Alter ≥ 50 Jahre					
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	69	71	12 (17)	1 (1)	11,91 [1,59; 89,20], p=0,016 14,26 [1,79; 113,42], p=0,012 15,55% [6,21%; 24,88%], p=0,001

Quellen: Tabellen 3.11 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

Effektschätzer adjustiert für Alter und Baseline Third Agent

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-129: p-Werte der meta-analytischen Interaktionstests für psychiatrische Erkrankungen bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,388
Geschlecht	p=0,727
Ethnizität	p=0,178
Region	p=0,291
Baseline CD4+ Zellzahl	p=0,583
Baseline Third Agent	p=0,670

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse für die Meta-Analyse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.17 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Erkrankungen der Niere und Harnwege bis Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-130: p-Werte der meta-analytischen Interaktionstests für Erkrankungen der Niere und Harnwege bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,925
Geschlecht	p=0,237
Ethnizität	p=0,632
Region	p=0,264
Baseline CD4+ Zellzahl	p=0,834
Baseline Third Agent	p=0,906

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.18 Subgruppenanalysen: unerwünschte Ereignisse von besonderem Interesse: Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-131: p-Werte der meta-analytischen Interaktionstests für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,785
Geschlecht	p=0,582
Ethnizität	p=0,972
Region	p=0,116
Baseline CD4+ Zellzahl	p=0,031
Baseline Third Agent	p=0,581

Quelle: eigene Berechnungen

Für diesen Endpunkt liegt bzgl. des Subgruppenfaktors der Baseline CD4+ Zellzahl ein Beleg für Effektmodifikation vor (p = 0,031). Diese Ergebnisse auf Subgruppenebene werden in diesem Abschnitt dargestellt.

Tabelle 4-132: Ergebnisse für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline CD4+ Zellzahl

Studie	**Anzahl der Patienten in der Analyse**		**Anzahl Patienten mit Ereignis, N (%)**		**RR [95%-KI], p-Wert OR]95%-KI], p-Wert RD [95%-KI], p-Wert**
Baseline CD4+ Zellzahl < 500					
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	85	74	7 (8)	9 (12)	0,66 [0,26; 1,68], p=0,388 0,63 [0,22; 1,81], p=0,394 -4,62% [-14,21%; 4,97%], p=0,345
SWORD-2	80	75	12 (15)	11 (15)	1,08 [0,51; 2,29], p=0,843 1,06 [0,43; 2,59], p=0,904 1,47% [-10,04%; 12,97%], p=0,803
Baseline CD4+ Zellzahl ≥ 500					
	DTG/RPV	CAR	DTG/RPV	CAR	DTG/RPV vs. CAR
SWORD-1	167	182	23 (14)	13 (7)	1,93 [1,01; 3,69], p=0,045 2,07 [1,01; 4,23], p=0,047 6,31% [-0,13%; 12,74%], p=0,055
SWORD-2	181	180	26 (14)	12 (7)	2,12 [1,10; 4,06], p=0,024 2,37 [1,15; 4,88], p=0,019 7,68% [1,38%; 13,97%], p=0,017

Quellen: Tabellen 3.11 und 3.12 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio, RD = Risikodifferenz

Effektschätzer adjustiert für Alter und Baseline Third Agent

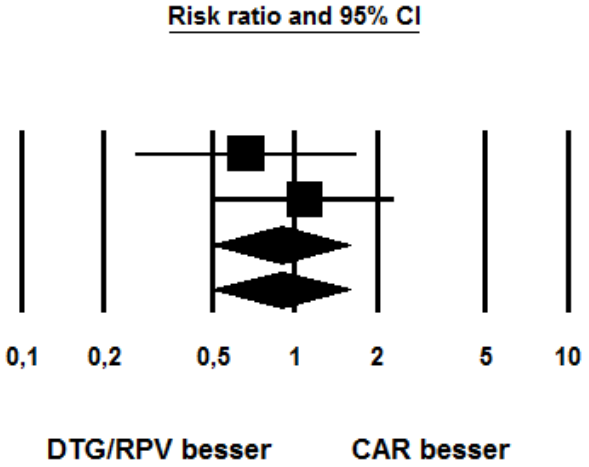
In der Subgruppe der Patienten mit niedrigerer Baseline CD4+ Zellzahl sind die Ereignisraten in SWORD-1 im CAR-Arm höher als im DTG/RPV-Arm (12 % vs. 8 %), während die Ereignisraten in SWORD-2 ausgeglichen sind (15 %). Es liegen jeweils keine statistisch signifikanten Unterschiede vor. Bei den Patienten mit höherer Baseline CD4+ Zellzahl sind die Ereignisraten im DTG/RPV-Arm in beiden Studien deutlich höher (jeweils 14 % vs. 7 %). In dieser Subgruppe sind die Unterschiede in beiden Studien statistisch signifikant (p = 0,045 bzw. p = 0,024).

Tabelle 4-133: Meta-Analyse für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Baseline CD4+ Zellzahl

	Relatives Risiko DTG/RPV vs. CAR [95% KI], p-Wert
Baseline CD4+ Zellzahl < 500	
Heterogenität: I^2=0%	
Modell mit festen Effekten	0,890 [0,496 – 1,597] p=0,696
Modell mit zufälligen Effekten	0,890 [0,496 – 1,597] p=0,696
Baseline CD4+ Zellzahl ≥ 500	
Heterogenität: I^2=0%	
Modell mit festen Effekten	2,022 [1,277 – 3,203] p=0,003
Modell mit zufälligen Effekten	2,022 [1,277 – 3,203] p=0,003
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall, VAS = visuelle Analogskala	

Für die Patienten mit niedriger Baseline CD4+ Zellzahl ist der beobachtete Effekt zugunsten von DTG/RPV nicht statistisch signifikant. In der Subgruppe der Patienten mit höherer Baseline CD4+ Zellzahl zeigt die Meta-Analyse, dass der beobachtete Effekt zuungunsten von DTG/RPV gegenüber CAR statistisch signifikant ist (p = 0,003).

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	CD4+ < 500	0,660	0,260	1,678	0,383
	SWORD2	CD4+ < 500	1,080	0,510	2,289	0,841
Fixed			0,890	0,496	1,597	0,696
Random			0,890	0,496	1,597	0,696

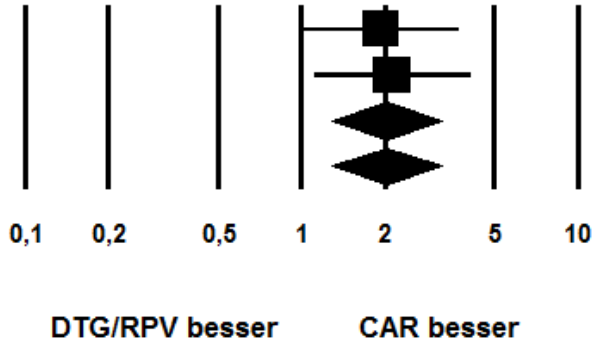


Heterogenität: I Quadrat = 0%

Abbildung 28: Meta-Analyse für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl < 500

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Model	Study name	Subgroup within study	Statistics for each study			
			Risk ratio	Lower limit	Upper limit	p-Value
	SWORD1	CD4+ >= 500	1,930	1,010	3,689	0,047
	SWORD2	CD4+ >= 500	2,120	1,103	4,073	0,024
Fixed			2,022	1,277	3,203	0,003
Random			2,022	1,277	3,203	0,003



Heterogenität: I Quadrat = 0%

Abbildung 29: Meta-Analyse für Erkrankungen der Haut und des Unterzellgewebes bis Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl ≥ 500

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.25.19 Subgruppenanalysen: alkalische Phosphatase (BSAP) zu Woche 48 - RCT

Auf Einzelstudienebene liegen für diesen Endpunkt keine statistisch signifikanten Interaktionstests vor. Es werden daher in diesem Modul keine Ergebnisse für die Einzelstudien SWORD-1 und SWORD-2 berichtet.

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-134: p-Werte der meta-analytischen Interaktionstests für BSAP zu Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,384
Geschlecht	p=0,636
Ethnizität	p=0,132
Region	p=0,958
Baseline CD4+ Zellzahl	p=0,537
Baseline Third Agent	p=0,234

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.20 Subgruppenanalysen: Prokollagen 1-N-terminales Propeptid (P1NP) zu Woche 48 - RCT

Für die Studie SWORD-2 liegt für den Subgruppenfaktor Alter eine statistisch signifikante Interaktion vor (p=0,026). Die Ergebnisse werden in nachfolgender Tabelle dargestellt.

Tabelle 4-135: Ergebnisse P1NP zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Alter

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
Alter < 50 Jahre							
SWORD-2						-4,13	-0,27
DTG/RPV	182	55,1±22,01	172	50,5±19,48	-4,2±19,94	[-7,44; -0,82]	[-0,48; -0,05]
CAR	180	56,1±20,63	172	55,2±18,70	-0,6±16,63	p=0,015	p=0,015
Alter ≥ 50 Jahre							
SWORD-2						-13,43	-0,72
DTG/RPV	78	52,2±24,80	74	42,8±21,07	-9,5±19,78	[-19,63; -7,23]	[-1,07; 0,38]
CAR	69	55,7±32,09	70	57,8±28,13	2,5±24,68	p<0,001	p<0,001

Quelle: Tabelle 3.16 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse

ANCOVA: adjustiert für Baseline P1NP

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-136: p-Werte der meta-analytischen Interaktionstests für P1NP zu Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,386
Geschlecht	p=0,072
Ethnizität	p=0,622
Region	p=0,431
Baseline CD4+ Zellzahl	p=0,071
Baseline Third Agent	p=0,636

Quelle: eigene Berechnungen

Es werden keine Subgruppenergebnisse für die Meta-Analyse dargestellt, da für keine der Subgruppen Belege für Effektmodifikation vorliegen.

4.3.1.3.25.21 Subgruppenanalysen: type 1 collagen cross-linked C-telopeptide (CTx) zu Woche 48 - RCT

Für die Studie SWORD-1 liegt für den Subgruppenfaktor Baseline Third Agent eine statistisch signifikante Interaktion vor (p=0,047). Die Ergebnisse werden in nachfolgender Tabelle dargestellt. Zudem liegt für SWORD-2 ein statistisch signifikanter Interaktions-p-Wert für Baseline CD4+ Zellzahl vor (p=0,007). Diese Ergebnisse werden im Anschluss mit den meta-analytischen Darstellungen präsentiert.

Tabelle 4-137: Ergebnisse CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline Third Agent

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
Baseline Third Agent = INI							
SWORD-1 DTG/RPV CAR	 46 48	 0,603±0,280 0,636±0,361	 45 46	 0,399±0,182 0,535±0,269	 -0,212±0,270 -0,099±0,263	-0,126 [-0,20; -0,05] p=0,002	-0,681 [-1,11; -0,25] p=0,002
Baseline Third Agent = NNRTI							
SWORD-1 DTG/RPV CAR	 131 133	 0,643±0,371 0,669±0,354	 124 130	 0,489±0,305 0,705±1,549	 -0,149±0,325 0,027±1,566	-0,199 [-0,48; 0,08] p=0,161	-0,178 [-0,43; 0,07] p=0,160
Baseline Third Agent = PI							
SWORD-1 DTG/RPV CAR	 75 73	 0,669±0,370 0,710±0,320	 65 67	 0,394±0,202 0,563±0,290	 -0,226±0,293 -0,145±0,262	-0,131 [-0,21; -0,06] p<0,001	-0,606 [-0,96; -0,25] p<0,001

Quelle: Tabelle 3.45 und 3.51 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse

ANCOVA: adjustiert für Baseline CTx

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-138: p-Werte der meta-analytischen Interaktionstests für CTx zu Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,283
Geschlecht	p=0,049
Ethnizität	p=0,800
Region	p=0,630
Baseline CD4+ Zellzahl	p=0,003
Baseline Third Agent	p=0,156

Quelle: eigene Berechnungen

Für diesen Endpunkt liegen jeweils bzgl. der Subgruppenfaktoren Geschlecht und Baseline CD4+ Zellzahl ein Beleg für Effektmodifikation vor (p = 0,049 bzw. p = 0,003). Diese Ergebnisse auf Subgruppenebene werden in diesem Abschnitt dargestellt.

Geschlecht:

Tabelle 4-139: Ergebnisse für CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Geschlecht

Studie	Deskriptive Analyse					Adjustierte Analyse: ANCOVA	
	Baseline		Woche 48				
	N	Mittelwert ± SD	N	Mittelwert ± SD	Mittlere Änderung zwischen Baseline und Woche 24 Mittelwert ± SD	Mittlere Änderung zu Woche 48 Differenz DTG/RPV – CAR [95%-KI] p-Wert	Mittlere Änderung zu Woche 48 Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert
weiblich							
SWORD-1						-0,152	-0,548
DTG/RPV	58	0,703±0,555	53	0,450±0,379	-0,228±0,436	[-0,26; -0,04]	[-0,96; -0,14]
CAR	51	0,668±0,296	46	0,597±0,305	-0,070±0,245	p=0,008	p=0,009
SWORD-2						-0,129	-0,500
DTG/RPV	61	0,626±0,335	56	0,478±0,328	-0,157±0,258	[-0,23; -0,03]	[-0,89; -0,11]
CAR	56	0,750±0,472	52	0,690±0,437	-0,067±0,315	p=0,012	p=0,013
männlich							
SWORD-1						-0,172	-0,185
DTG/RPV	194	0,626±0,268	181	0,444±0,218	-0,169 ±0,258	[-0,36; 0,02]	[-0,39; 0,02]
CAR	203	0,676±0,358	197	0,642±1,269	-0,038±1,285	p=0,074	p=0,074
SWORD-2						-0,065	-0,232
DTG/RPV	199	0,683±0,314	187	0,532±0,338	-0,146±0,328	[-0,12; -0,01]	[-0,44; -0,03]
CAR	194	0,697±0,359	190	0,609±0,294	-0,093±0,352	p=0,026	p=0,026

Quelle: Tabellen 3.17, 3.18, 3.25 und 3.26 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse, VAS = Visuelle Analogskala

ANCOVA: adjustiert für Baseline-CTx

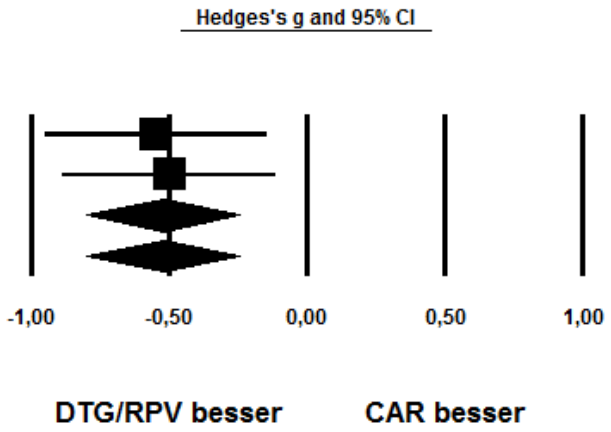
Sowohl bei den Frauen als auch bei den Männern liegt jeweils eine deutlichere Verbesserung dieses Biomarkers im DTG/RPV-Arm in beiden Studien vor. Bei den Frauen ist der Effekt deutlich stärker ausgeprägt als bei den Männern. Mit Ausnahme der SWORD-1 in der Subgruppe der Männer sind die Effekte statistisch signifikant.

Tabelle 4-140: Meta-Analyse für CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Geschlecht

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
weiblich	
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,523 [-0,801; -0,245] p<0,001
Modell mit zufälligen Effekten	-0,523 [-0,801; -0,245] p<0,001
männlich	
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,208 [-0,352; -0,065] p=0,004
Modell mit zufälligen Effekten	-0,208 [-0,352; -0,065] p=0,004
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall, VAS = visuelle Analogskala	

Sowohl für die Frauen als auch für die Männer liegen statistisch signifikante Effekte zugunsten von DTG/RPV vor (p < 0,001 bzw. p = 0,004), wobei der Effekt in der Subgruppe der Frauen deutlicher ausgeprägt ist.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	female	-0,548	-0,950	-0,146	0,008
	SWORD-2	female	-0,500	-0,886	-0,114	0,011
Fixed			-0,523	-0,801	-0,245	0,000
Random			-0,523	-0,801	-0,245	0,000

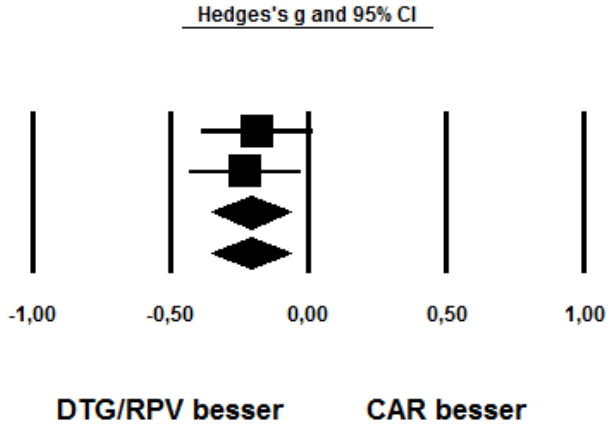


Heterogenität: I Quadrat 0%

Abbildung 30: Meta-Analyse für CTx zu Woche 48 aus RCT; DTG/RPV versus CAR – weiblich

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	male	-0,185	-0,387	0,017	0,072
	SWORD-2	male	-0,232	-0,436	-0,028	0,026
Fixed			-0,208	-0,352	-0,065	0,004
Random			-0,208	-0,352	-0,065	0,004



Heterogenität: I Quadrat 0%

Abbildung 31: Meta-Analyse für CTx zu Woche 48 aus RCT; DTG/RPV versus CAR – männlich

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Baseline CD4+ Zellzahl:

Tabelle 4-141: Ergebnisse für CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline CD4+ Zellzahl

Studie	Deskriptive Analyse					Adjustierte Analyse: ANCOVA	
	Baseline		Woche 48				
	N	Mittelwert ± SD	N	Mittelwert ± SD	Mittlere Änderung zwischen Baseline und Woche 24 Mittelwert ± SD	Mittlere Änderung zu Woche 48 Differenz DTG/RPV – CAR [95%-KI] p-Wert	Mittlere Änderung zu Woche 48 Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert
Baseline CD4+ Zellzahl < 500							
SWORD-1						-0,123	-0,492
DTG/RPV	85	0,711±0,437	76	0,462±0,337	-0,215±0,356	[-0,21; -0,04]	[-0,82; -0,16]
CAR	74	0,704±0,461	71	0,603±0,326	-0,110±0,305	p=0,004	p=0,004
SWORD-2						-0,174	-0,701
DTG/RPV	80	0,666±0,289	77	0,496±0,301	-0,173±0,260	[-0,26; -0,09]	[-1,04; -0,36]
CAR	74	0,689±0,306	68	0,688±0,320	-0,010±0,287	p<0,001	p<0,001
Baseline CD4+ Zellzahl ≥ 500							
SWORD-1						-0,189	-0,191
DTG/RPV	167	0,609±0,301	158	0,437±0,218	-0,166 ±0,281	[-0,40; 0,03]	[-0,41; 0,03]
CAR	180	0,663±0,286	172	0,646±1,351	-0,017±1,368	p=0,086	p=0,086
SWORD-2						-0,045	-0,156
DTG/RPV	180	0,672±0,332	166	0,531±0,351	-0,137±0,335	[-0,11; 0,02]	[-0,37; 0,06]
CAR	176	0,717±0,417	174	0,602±0,333	-0,118±0,360	p=0,154	p=0,154

Quelle: Tabellen 3.17, 3.18, 3.25 und 3.26 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse, VAS = Visuelle Analogskala

ANCOVA: adjustiert für Baseline-CTx

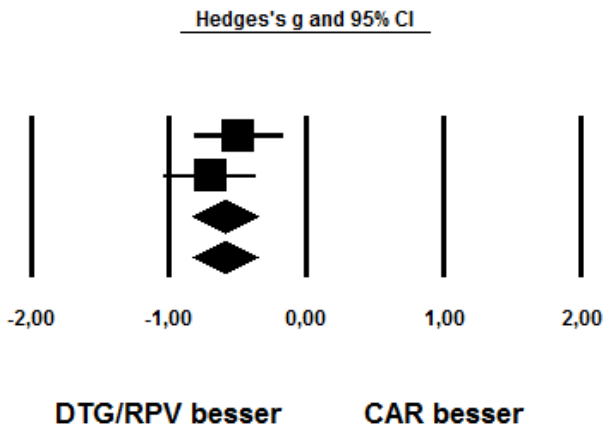
Sowohl bei den Patienten mit niedriger als auch bei den Patienten mit hoher Baseline CD4+ Zellzahl liegt jeweils in beiden Studien eine Verbesserung dieses Biomarkers vor. Bei den Patienten mit niedrigerer Baseline CD4+ Zellzahl ist der Effekt deutlich stärker ausgeprägt als bei denjenigen mit höherer CD4+Zellzahl. Für Erstere ist der Effekt zudem statistisch signifikant in beiden Einzelstudien (p = 0,004 bzw. p < 0,001).

Tabelle 4-142: Meta-Analyse für CTx zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Baseline CD4+ Zellzahl

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
Baseline CD4+ Zellzahl < 500	
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,595 [-0,830; -0,360] p<0,001
Modell mit zufälligen Effekten	-0,595 [-0,830; -0,360] p<0,001
Baseline CD4+ Zellzahl ≥ 500	
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,173 [-0,326; -0,021] p=0,026
Modell mit zufälligen Effekten	-0,173 [-0,326; -0,021] p=0,026
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall, VAS = visuelle Analogskala	

Sowohl für die Patienten mit niedriger als auch für die Patienten mit hoher Baseline CD4+ Zellzahl liegen statistisch signifikante Effekte zugunsten von DTG/RPV vor (p < 0,001 bzw. p = 0,026), wobei der Effekt in der Subgruppe der Patienten mit niedriger Baseline CD4+ Zellzahl deutlicher ausgeprägt ist.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	CD4+ < 500	-0,492	-0,821	-0,163	0,003
	SWORD-2	CD4+ < 500	-0,701	-1,036	-0,366	0,000
Fixed			-0,595	-0,830	-0,360	0,000
Random			-0,595	-0,830	-0,360	0,000

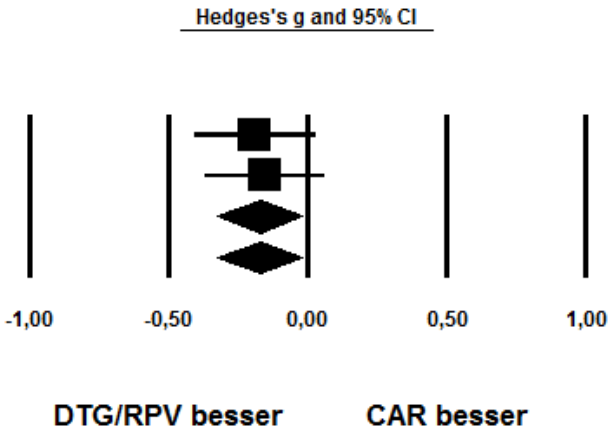


Heterogenität: I Quadrat 0%

Abbildung 32: Meta-Analyse für CTx zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl < 500

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	CD4+ >= 500	-0,191	-0,409	0,027	0,085
	SWORD-2	CD4+ >= 500	-0,156	-0,370	0,058	0,152
Fixed			-0,173	-0,326	-0,021	0,026
Random			-0,173	-0,326	-0,021	0,026



Heterogenität: I Quadrat 0%

Abbildung 33: Meta-Analyse für CTx zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline CD4+ Zellzahl ≥ 500

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.25.22 Subgruppenanalysen: Osteocalcin zu Woche 48 - RCT

Für die Studie SWORD-2 liegt für den Subgruppenfaktor Alter eine statistisch signifikante Interaktion vor (p=0,047). Die Ergebnisse werden in nachfolgender Tabelle dargestellt. Zudem liegt für beide Studien, SWORD-1 und SWORD-2, jeweils ein statistisch signifikanter Interaktions-p-Wert für Baseline Third Agent vor (p=0,001 bzw. p=0,044). Diese Ergebnisse werden im Anschluss mit den meta-analytischen Darstellungen präsentiert.

Tabelle 4-143: Ergebnisse Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Alter

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
Alter < 50 Jahre							
SWORD-2						-3,027	-0,582
DTG/RPV	183	24,45±8,726	171	19,78±6,249	-4,37±6,976	[-4,14; -1,915]	[-0,80; -0,36]
CAR	181	24,14±8,426	171	22,86±7,701	-1,32±6,057	p<0,001	p<0,001
Alter ≥ 50 Jahre							
SWORD-2						-6,176	-1,065
DTG/RPV	77	25,18±12,272	74	18,58±9,455	-6,84±7,882	[-8,13; -4,22]	[-1,42; -0,71]
CAR	67	25,48±12,025	70	24,48±10,690	-0,70±5,938	p<0,001	p<0,001

Quelle: Tabelle 3.28 und 3.30 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse

ANCOVA: adjustiert für Baseline Osteocalcin

Folgende Tabelle gibt einen Überblick über die p-Werte der meta-analytischen Interaktionstests für diesen Endpunkt.

Tabelle 4-144: p-Werte der meta-analytischen Interaktionstests für Osteocalcin zu Woche 48

Subgruppenfaktor	**Interaktions-p-Wert**
Alter	p=0,815
Geschlecht	p=0,278
Ethnizität	p=0,558
Region	p=0,710
Baseline CD4+ Zellzahl	p=0,469
Baseline Third Agent	p<0,001

Quelle: eigene Berechnungen

Für diesen Endpunkt liegen bzgl. des Subgruppenfaktors Baseline Third Agent ein Beleg für Effektmodifikation vor ($p < 0,001$). Die entsprechenden Ergebnisse auf Subgruppenebene werden in diesem Abschnitt dargestellt.

Tabelle 4-145: Ergebnisse für Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel – Subgruppenfaktor: Baseline Third Agent

Studie	**Deskriptive Analyse**					**Adjustierte Analyse: ANCOVA**	
	Baseline		**Woche 48**				
	N	**Mittelwert ± SD**	**N**	**Mittelwert ± SD**	**Mittlere Änderung zwischen Baseline und Woche 24** **Mittelwert ± SD**	**Mittlere Änderung zu Woche 48** **Differenz DTG/RPV – CAR [95%-KI] p-Wert**	**Mittlere Änderung zu Woche 48** **Hedges' g Differenz DTG/RPV – CAR [95%-KI] p-Wert**
Baseline Third Agent = INI							
SWORD-1						-3,335	-0,679
DTG/RPV	46	21,92±7,090	45	17,67±6,461	-4,42±5,460	[-5,41; -1,26]	[-1,11; -0,25]
CAR	47	22,65±8,159	46	21,05±6,970	-1,25±6,184	p=0,002	p=0,002
SWORD-2						-3,232	-0,545
DTG/RPV	59	24,77±9,754	53	18,88±7,789	-6,02±7,866	[-5,62; -0,84]	[-0,95; -0,14]
CAR	49	24,96±12,215	46	22,28±9,949	-2,91±7,297	p=0,009	p=0,009
Baseline Third Agent = NNRTI							
SWORD-1						-2,375	-0,453
DTG/RPV	131	21,93±9,570	123	18,63±7,674	-3,14±7,320	[-3,68; -1,07]	[-0,70; -0,20]
CAR	133	21,83±6,597	131	21,19±6,355	-0,87±5,300	p<0,001	p<0,001
SWORD-2						-3,088	-0,634
DTG/RPV	143	23,07±9,498	139	19,73±7,754	-3,36±5,733	[-4,26; -1,92]	[-0,88; -0,39]
CAR	139	23,29±8,920	134	22,90±8,557	-0,28±5,492	p<0,001	p<0,001
Baseline Third Agent = PI							
SWORD-1						-8,390	-1,319
DTG/RPV	74	25,11±10,515	66	17,29±5,178	-6,77±8,850	[-10,59; -6,19]	[-1,70; -0,94]
CAR	73	26,81±10,349	68	26,69±8,937	0,09±8,762	p<0,001	p<0,001
SWORD-2						-6,630	-1,182
DTG/RPV	58	28,50±10,078	53	19,13±5,810	-8,81±8,963	[-8,75; -4,51]	[-1,59; -0,78]
CAR	60	26,93±7,921	61	25,07±7,793	-1,67±5,768	p<0,001	p<0,001

Quelle: Tabellen 3.53, 3.54, 3.55 und 3.56 in den Zusatzanalysen

Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, N = Anzahl, KI = Konfidenzintervall, SD = Standardabweichung, ANCOVA = Kovarianzanalyse, INI = Integrase-Inhibitor, NNRTI = Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren, PI = Protease-Inhibitor

ANCOVA: adjustiert für Baseline-Osteocalcin

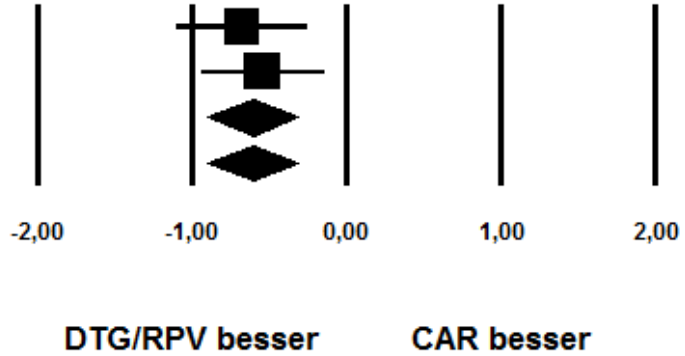
In allen drei Subgruppen und jeweils in beiden Studien ist die Osteocalcin-Senkung statistisch signifikant höher im DTG/RPV-Arm als im CAR-Arm. Am deutlichsten ausgeprägt ist der Unterschied bei den Patienten, die bei Baseline einen Protease-Inhibitor als Bestandteil ihrer ART bekamen (jeweils $p < 0,001$).

Tabelle 4-146: Meta-Analyse für Osteocalcin zu Woche 48 aus RCT mit dem zu bewertenden Arzneimittel - Subgruppenfaktor: Baseline Third Agent

	Hedges' g DTG/RPV vs. CAR [95% KI], p-Wert
Baseline Third Agent = INI	
Heterogenität: I^2=0%	
Modell mit festen Effekten	-0,608 [-0,900; -0,316] p<0,001
Modell mit zufälligen Effekten	-0,608 [-0,900; -0,316] p<0,001
Baseline Third Agent = NNRTI	
Heterogenität: I^2=3,1%	
Modell mit festen Effekten	-0,545 [-0,720; -0,370] p<0,001
Modell mit zufälligen Effekten	-0,545 [-0,722; -0,368] p<0,001
Baseline Third Agent = PI	
Heterogenität: I^2=0%	
Modell mit festen Effekten	-1,255 [-1,530; -0,980] p<0,001
Modell mit zufälligen Effekten	-1,255 [-1,530; -0,980] p<0,001
Abkürzungen: DTG/RPV = Dolutegravir/Rilpivirin, CAR = Current Antiretroviral Regimen, KI = Konfidenzintervall, INI = Integrase-Inhibitor, NNRTI = Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren, PI = Protease-Inhibitor	

Auch in der Meta-Analyse zeigt sich in allen drei Subgruppen jeweils ein statistisch signifikant besserer Effekt zugunsten von DTG/RPV gegenüber CAR (jeweils p<0,001). Am deutlichsten ausgeprägt ist der Unterschied bei den Patienten, die bei Baseline einen Protease-Inhibitor als Bestandteil ihrer ART bekamen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	INI	-0,679	-1,104	-0,254	0,002
	SWORD-2	INI	-0,545	-0,947	-0,143	0,008
Fixed			-0,608	-0,900	-0,316	0,000
Random			-0,608	-0,900	-0,316	0,000

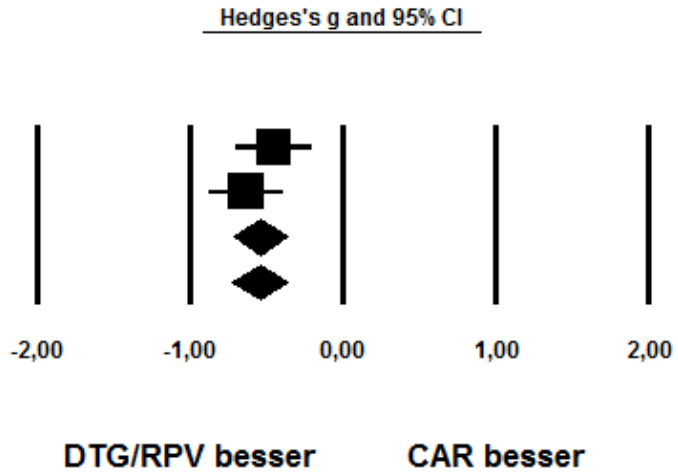


Heterogenität: I Quadrat 0%

Abbildung 34: Meta-Analyse für Osteocalcin zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline Third Agent = INI

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	NNRTI	-0,453	-0,702	-0,204	0,000
	SWORD-2	NNRTI	-0,634	-0,879	-0,389	0,000
Fixed			-0,545	-0,720	-0,370	0,000
Random			-0,545	-0,722	-0,368	0,000

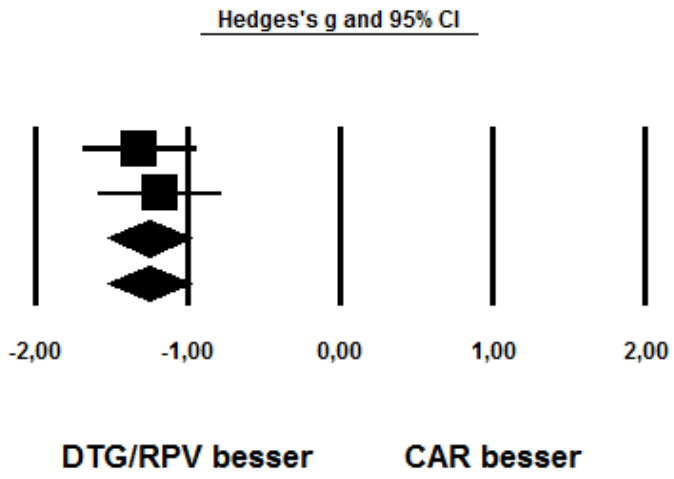


Heterogenität: I Quadrat 3,1%

Abbildung 35: Meta-Analyse für Osteocalcin zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline Third Agent = NNRTI

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

Model	Study name	Subgroup within study	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
	SWORD-1	PI	-1,319	-1,695	-0,943	0,000
	SWORD-2	PI	-1,182	-1,584	-0,780	0,000
Fixed			-1,255	-1,530	-0,980	0,000
Random			-1,255	-1,530	-0,980	0,000



Heterogenität: I Quadrat 0%

Abbildung 36: Meta-Analyse für Osteocalcin zu Woche 48 aus RCT; DTG/RPV versus CAR – Baseline Third Agent = PI

Etwaige Abweichungen der Einzelstudienergebnisse zu den weiter oben dargestellten, resultieren aus der Verwendung gerundeter Schätzwerte, die für die Berechnung der Meta-Analyse zur Verfügung standen.

4.3.1.3.26 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.

In der nachfolgenden Tabelle werden die Ergebnisse der einzelnen patientenrelevanten Endpunkte aus den obigen Abschnitten zu RCT – **Teilpopulation ohne Umstellindikation** - zusammengefasst. Pro Endpunkt ist der jeweilige Effektschätzer des meta-analytischen Modells entsprechend seiner Operationalisierung (relatives Risiko bzw. Hedges‘ g) inklusive des 95%-Konfidenzintervalls und des p-Werts angegeben. Zur Interpretation sind jeweils die Ergebnisse der Modelle mit festen Effekten dargestellt. Lediglich für Endpunkte, für die sich in der Meta-Analyse starke Heterogenität (I^2 > 70 %) ergibt, werden primär die Ergebnisse der Einzelstudien interpretiert.

Subgruppenergebnisse werden in der Tabelle jeweils aufgeführt, falls ein Beleg für Effektmodifikation vorliegt (d.h. Interaktions-p-Wert < 0,05), da in diesen Fällen i.A. eine Interpretation des Zusatznutzens auf Ebene der Subgruppen erforderlich ist. Auf diese Interpretationen wird in Abschnitt 4.4.2 eingegangen. Falls trotz eines Belegs auf Effektmodifikation die Ergebnisse in allen jeweiligen Subgruppen als gleichgerichtet interpretiert werden können, wird in diesen Fällen die Interpretation des Zusatznutzens für die gesamte Zielpopulation getroffen (und dieser Fall durch eine Fußnote kenntlich gemacht).

Für die **Teilpopulation mit Umstellindikation** lag keine RCT vor, demzufolge können in diesem Abschnitt dafür keine Ergebnisse aus RCT dargestellt werden.

Tabelle 4-147: Übersicht der Ergebnisse aus RCT für die Teilpopulation ohne Umstellindikation auf Endpunktebene

	Effektschätzer der Meta-Analyse (Modell mit festen Effekten) **Hedges' g oder mittlere Änderung oder RR [95% KI], p-Wert**		**Interaktions -test p-Wert**
Viruslast < 50 Kopien/ml zu Woche 48		RR: 1,00 [0,98; 1,03], p=0,970	
EQ-5D zu Woche 48 Utility Score		Hedges' g: 0,014 [-0,110; 0,137], p=0,830	
VAS		Hedges' g: -0,005 [-0,128; 0,118], p=0,958	p=0,036
	CD4+ Zahl < 500	Hedges' g: 0,193 [-0,030; 0,416], p=0,090	
	CD4+ Zahl ≥ 500	Hedges' g: -0,093 [-0,241; 0,054], p=0,216	
Symptom Distress Module zu Woche 48 Symptom Count Score		Hedges' g: -0,118 [-0,250; 0,015], p=0,081	
Symptom Bother Score		Hedges' g: -0,167 [-0,299; -0,034], p=0,014	
HIVTSQ Total Score zu Woche 48		Hedges' g: 0,168 [0,045; 0,291], p=0,007	
Mortalität bis Woche 48		wegen sehr weniger Ereignisse keine Meta-Analyse durchgeführt	
Unerwünschte Ereignisse bis Woche 48		RR: 1,08 [1,01; 1,16], p=0,037	
Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48[a]	SWORD-1	RR: 0,98 [0,49; 1,95], p=0,955	
	SWORD-2	RR: 3,28 [1,23; 8,76], p=0,018	
Schwerwiegende unerwünschte Ereignisse bis Woche 48		RR: 1,31 [0,73; 2,34], p=0,361	
Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten		RR: 6,33 [1,87; 21,38], p=0,003	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen		RR: 1,01 [0,14; 7,27], p=0,992	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts[a]	SWORD-1	RR: 1,28 [0,95; 1,74], p=0,104	
	SWORD-2	RR: 2,16 [1,41; 3,32], p<0,001	

Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen		RR: 2,52 [0,49; 12,97], p=0,268	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität		wegen sehr weniger Ereignisse keine Meta-Analyse durchgeführt	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen		RR: 1,05 [0,41; 2,69], p=0,926	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen[a]	SWORD-1 SWORD-2	RR: 0,71 [0,48; 1,05], p=0,088 RR: 1,33 [0,87; 2,03], p=0,184	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems[a]	SWORD-1 SWORD-2	RR: 1,36 [0,86; 2,14], p=0,193 RR: 2,78 [1,55; 4,98], p<0,001	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen		RR: 1,88 [1,25; 2,83], p=0,003	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege		RR: 1,61 [0,70; 3,72], p=0,263	
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes		RR: 1,49 [1,05; 2,13], p=0,027	
	CD4+ Zahl < 500	RR: 0,89 [0,50; 1,60], p=0,696	p=0,031
	CD4+ Zahl ≥ 500	RR: 2,02 [1,28; 3,20], p=0,003	
Knochenmineraldichte [g/cm^2][b]			
Hüfte		adj. mittlere Änderung: 1,29 [0,27; 2,31], p=0,014	
Lendenwirbel		adj. mittlere Änderung: 1,32 [0,07; 2,57], p=0,039	
Knochenspezifischer Biomarker: BSAP zu Woche 48		Hedges' g: -1,020 [-1,154; -0,886], p<0,001	
Knochenspezifischer Biomarker: P1NP zu Woche 48		Hedges' g: -0,529 [-0,658; -0,400], p<0,001	
Knochenspezifischer Biomarker: CTx zu Woche 48		Hedges' g: -0,252 [-0,380; -0,124], p<0,001	
	weiblich	Hedges' g: -0,523 [-0,801; -0,245], p<0,001	p=0,049
	männlich	Hedges' g: -0,208 [-0,352; -0,065], p=0,004	
	CD4+ Zahl < 500	Hedges' g: -0,595 [-0,830; -0,360], p<0,001	p=0,003
	CD4+ Zahl ≥ 500	Hedges' g: -0,173 [-0,326; -0,021], p=0,026	

Knochenspezifischer Biomarker: Osteocalcin zu Woche 48		Hedges' g: -0,734 [-0,866; -0,602], p<0,001	p<0,001
	Third Agent = INI	Hedges' g: -0,608 [-0,900; -0,316], p<0,001	
	Third Agent = NNRTI	Hedges' g: -0,545 [-0,720; -0,370], p<0,001	
	Third Agent = PI	Hedges' g: -1,255 [-1,530; -0,980], p<0,001	

[a] Darstellung der Ergebnisse der Studien SWORD-1 und SWORD-2, da in der Meta-Analyse eine hohe Heterogenität vorlag
[b] Ergebnis der Einzelstudie (DEXA Substudie)

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 (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

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

Es wurde kein indirekter Vergleich durchgeführt.

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 relevant.

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 4-148: 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 4-149: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	**Studie**	**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 relevant.

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

Tabelle 4-150: Operationalisierung von <Endpunkt xxx>

Studie	**Operationalisierung**
<Studie 1>	

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

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

Tabelle 4-151: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunktherheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**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 relevant.

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 4-152: 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 relevant.

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 der üblichen 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.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht relevant.

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

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.***

Nicht relevant.

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 (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

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

Es wurden keine nicht randomisierten vergleichenden Studien für den Nachweis eines Zusatznutzens herangezogen.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

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

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

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

Tabelle 4-153: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten F	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht relevant.

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 4-154: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-155: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht relevant.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht relevant.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt* 4.3.1.3.1.**

Nicht relevant.

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 (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

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

Es wurden keine weiteren Untersuchungen für den Nachweis eines Zusatznutzens herangezogen.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

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

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 relevant.

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 4-156: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte 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 relevant.

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. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht relevant.

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

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.***

Nicht relevant.

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 relevant.

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.

Teilpopulation ohne Umstellindikation:

Studienqualität:

Für die vorliegende Darstellung des Zusatznutzens wurden die Ergebnisse zweier randomisierter klinischer Studien (RCT) von hoher Qualität herangezogen (SWORD-1 und SWORD-2). Die Studien besitzen auf Studienebene jeweils niedriges Verzerrungspotenzial und ermöglichen den direkten Vergleich des zu bewertenden Arzneimittels Dolutegravir/Rilpivirin mit der zweckmäßigen Vergleichstherapie (CAR). Beide RCT sind für eine meta-analytische Zusammenfassung geeignet.

Zusätzlich wird die DEXA-Substudie dargestellt und herangezogen. Bei dieser Studie handelt es sich um einen randomisierten Vergleich einer Subgruppe von Patienten aus den beiden SWORD-Studien (Vergleich DTG/RPV vs. CAR wie in den SWORD-Studien) hinsichtlich der Veränderung der Knochenmineraldichte. Obwohl diese Studie daher selbst kein RCT ist, besitzt sie aufgrund des a priori geplanten Rekrutierungsprozesses als Subgruppe der beiden RCT hohe Qualität..

Validität der herangezogenen Endpunkte:

Alle Endpunkte, die für die Darstellung der Wirksamkeit und Sicherheit im Rahmen der Nutzenbewertung von Dolutegravir/Rilpiviren herangezogen wurden, sind patientenrelevant.

Die Instrumente, die zur Erhebung der Morbidität bzw. Lebensqualität verwendet wurden, sind allgemein anerkannte und im vorliegenden Indikationsgebiet des HIV validierte (SDM, HIVTSQ) bzw. generisch validierte Messinstrumente (EQ-5D).

Evidenzstufe:

Der Nachweis für den Zusatznutzen beruht auf den Ergebnissen einer Meta-Analyse von zwei randomisierten klinischen Studien (SWORD-1 und SWORD-2) und entspricht somit dem höchsten Evidenzgrad Ia. Die DEXA-Substudie als RCT von hoher Qualität entspricht dem Evidenzgrad Ib.

Teilpopulation mit Umstellindikation:

Für die **Teilpopulation mit Umstellindikation** lag keine RCT-Evidenz vor, somit kann die Aussagekraft nicht beurteilt werden.

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. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. 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.

Tabelle 4-157: Übersicht der Ergebnisse pro Endpunkt – Teilpopulation ohne Umstellindikation

	Effektschätzer der Meta-Analyse (Modell mit festen Effekten) **Hedges' g oder mittlere Änderung oder RR [95% KI], p-Wert**	**Wahrscheinlichkeit und Ausmaß des Zusatznutzens für DTG/RPV**
Viruslast < 50 Kopien/ml zu Woche 48	RR: 1,00 [0,98; 1,03], p=0,970	Zusatznutzen nicht belegt
EQ-5D zu Woche 48 Utility Score	Hedges' g: 0,014 [-0,110; 0,137], p=0,830	Zusatznutzen nicht belegt
VAS	Hedges' g: -0,005 [-0,128; 0,118], p=0,958 CD4+ Zahl < 500 Hedges' g: 0,193 [-0,030; 0,416], p=0,090 CD4+ Zahl ≥ 500 Hedges' g: -0,093 [-0,241; 0,054], p=0,216	- Zusatznutzen nicht belegt Zusatznutzen nicht belegt
Symptom Distress Module zu Woche 48 Symptom Count Score	Hedges' g: -0,118 [-0,250; 0,015], p=0,081	Zusatznutzen nicht belegt
Symptom Bother Score	Hedges' g: -0,167 [-0,299; -0,034], p=0,014	Beleg für geringen Zusatznutzen
HIVTSQ Total Score zu Woche 48	Hedges' g: 0,168 [0,045; 0,291], p=0,007	Beleg für geringen Zusatznutzen
Mortalität bis Woche 48	wegen sehr weniger Ereignisse keine Meta-Analyse durchgeführt	Zusatznutzen nicht belegt
Unerwünschte Ereignisse bis Woche 48	RR: 1,08 [1,01; 1,16], p=0,037	Zusatznutzen nicht belegt[b]
Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48[a]	SWORD-1 RR: 0,98 [0,49; 1,95], p=0,955 SWORD-2 RR: 3,28 [1,23; 8,76], p=0,018	Zusatznutzen nicht belegt[b]
Schwerwiegende unerwünschte Ereignisse bis Woche 48	RR: 1,31 [0,73; 2,34], p=0,361	Zusatznutzen nicht belegt
Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten	RR: 6,33 [1,87; 21,38], p=0,003	Zusatznutzen nicht belegt[b]
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen	RR: 1,01 [0,14; 7,27], p=0,992	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts[a]	SWORD-1 RR: 1,28 [0,95; 1,74], p=0,104 SWORD-2 RR: 2,16 [1,41; 3,32], p<0,001	Zusatznutzen nicht belegt[b]

Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen		RR: 2,52 [0,49; 12,97], p=0,268	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität		wegen sehr weniger Ereignisse keine Meta-Analyse durchgeführt	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen		RR: 1,05 [0,41; 2,69], p=0,926	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen[a]	SWORD-1 SWORD-2	RR: 0,71 [0,48; 1,05], p=0,088 RR: 1,33 [0,87; 2,03], p=0,184	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems[a]	SWORD-1 SWORD-2	RR: 1,36 [0,86; 2,14], p=0,193 RR: 2,78 [1,55; 4,98], p<0,001	Zusatznutzen nicht belegt[b]
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen		RR: 1,88 [1,25; 2,83], p=0,003	Zusatznutzen nicht belegt[b]
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege		RR: 1,61 [0,70; 3,72], p=0,263	Zusatznutzen nicht belegt
Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes	 CD4+ Zahl < 500 CD4+ Zahl ≥ 500	RR: 1,49 [1,05; 2,13], p=0,027 RR: 0,89 [0,50; 1,60], p=0,696 RR: 2,02 [1,28; 3,20], p=0,003	- Zusatznutzen nicht belegt Zusatznutzen nicht belegt[b]
Knochenmineraldichte [g/cm²][c] Hüfte		adj. mittlere Änderung: 1,29 [0,27; 2,31], p=0,014	Hinweis auf geringen Zusatznutzen
Lendenwirbel		adj. mittlere Änderung: 1,32 [0,07; 2,57], p=0,039	Hinweis auf geringen Zusatznutzen
Knochenspezifischer Biomarker: BSAP zu Woche 48		Hedges' g: -1,020 [-1,154; -0,886], p<0,001	Zusatznutzen nicht belegt[b]
Knochenspezifischer Biomarker: P1NP zu Woche 48		Hedges' g: -0,529 [-0,658; -0,400], p<0,001	Zusatznutzen nicht belegt[b]
Knochenspezifischer Biomarker: CTx zu Woche 48	 weiblich männlich CD4+ Zahl < 500 CD4+ Zahl ≥ 500	Hedges' g: -0,252 [-0,380; -0,124], p<0,001 Hedges' g: -0,523 [-0,801; -0,245], p<0,001 Hedges' g: -0,208 [-0,352; -0,065], p=0,004 Hedges' g: -0,595 [-0,830; -0,360], p<0,001 Hedges' g: -0,173 [-0,326; -0,021], p=0,026	Zusatznutzen nicht belegt[b,d] - - - -

Knochenspezifischer Biomarker: Osteocalcin zu Woche 48		Hedges' g: -0,734 [-0,866; -0,602], p<0,001	Zusatznutzen nicht belegt[b,d]
	Third Agent = INI	Hedges' g: -0,608 [-0,900; -0,316], p<0,001	-
	Third Agent = NNRTI	Hedges' g: -0,545 [-0,720; -0,370], p<0,001	-
	Third Agent = PI	Hedges' g: -1,255 [-1,530; -0,980], p<0,001	-

[a] Darstellung der Ergebnisse der Studien SWORD-1 und SWORD-2, da in der Meta-Analyse eine hohe Heterogenität vorlag
[b] Zur Interpretation siehe hierzu die Begründung im Text direkt im Anschluss an diese Tabelle
[c] Ergebnis der Einzelstudie (DEXA Substudie)
[d] Bewertung im Gesamtkollektiv möglich, da die Ergebnisse in allen relevanten Subgruppenanalysen (Geschlecht, CD4-Zellzahl, Third Agent) gleichgerichtet sind

Wie bereits in den Abschnitten 4.2.1 und 4.2.2 ausgeführt, wird das **Ausmaß und die Wahrscheinlichkeit des Zusatznutzens für die gesamte Zielpopulation, d.h. für vorbehandelte erwachsene Patienten,** jeweils auf der Grundlage der Ergebnisse aller patientenrelevanter Endpunkte über die längste verfügbare Behandlungsdauer von 48 Wochen bewertet.

Wie sich ebenfalls aus Abschnitt 4.2.1 ergibt, folgt aus der Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA, dass die **Zielpopulation** in zwei Teilpopulationen entsprechend des Grundes zur Therapieumstellung einzuteilen und separat zu betrachten und zu bewerten ist. Dies sind die **Teilpopulation mit Umstellindikation** (Umstellindikationen sind nach der Definition des G-BA virologisches Versagen, Nebenwirkungen oder Toxizitäten), sowie die **Teilpopulation ohne Umstellindikation**.

Diese beiden Teilpopulationen werden im Folgenden separat diskutiert.

<u>Patienten der Zielpopulation ohne Umstellindikation:</u>

Morbidität:

Ergebnisse zur Morbidität werden anhand des Endpunktes „Viruslast < 50 Kopien/ml zu Woche 48“ bewertet. In Konsequenz der Indikation und des Studien-Designs, das für diesen Endpunkt in beiden SWORD-Zulassungstudien eine Nicht-Unterlegenheits-Hypothese geprüft hat, war eine Überlegenheit für das zu bewertende Arzneimittel Dolutegravir/Rilpivirin (DTG/RPV) im Vergleich zur Fortführung der bisherigen ART (CAR) für diesen Endpunkt nicht zu erwarten. Die Patientenpopulation der beiden SWORD-Studien bestand entsprechend der Indikation aus Patienten, zu Studienbeginn die eine Viruslast < 50 Kopien/ml aufweisen. Dieses Kriterium war ein Einschlusskriterium für die SWORD-Studien, es gab einen sehr geringen Anteil von weniger als 2 % aller Patienten (für beide Studien SWORD-1 und SWORD-2), die bei Baseline dieses Kriterium nicht erfüllten, obwohl es auch für diese Patienten, also insgesamt für 100 % der Patienten, zum Screening-Zeitpunkt erfüllt war. Insofern waren Verbesserungen im Vergleich zu Baseline nicht zu erwarten und die Fragestellung bestand darin festzustellen, inwieweit sich ein Verlust des virologischen Erfolgs beider Therapien über 48 Wochen einstellen und ggfs. unterscheiden würde.

Im Ergebnis gab es keinen Unterschied bezüglich dieses Endpunktes zwischen DTG/RPV und CAR. Die Viruslast war auch am Ende beider Studien nach 48 Wochen Behandlungsdauer für mindestens 94 % der Patienten in beiden Behandlungsgruppen mit Werten < 50 Kopien/ml kontrolliert und die Nicht-Unterlegenheit wurde damit statistisch signifikant belegt. Die meta-analytische Auwertung zum Endpunkt „Viruslast < 50 Kopien/ml zu Woche 48“ konnte die nahezu perfekte Übereinstimmung zwischen DTG/RPV und der zweckmäßigen Vergleichstherapie CAR zeigen (RR = 1,0). Ein Zusatznutzen für DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie CAR wurde hinsichtlich dieses Endpunktes nicht belegt.

Ergänzend zu diesem Endpunkt sollte auch das Auftreten von AIDS-definierenden Ereignissen (CDC Klasse C) und/oder Todesfällen überprüft werden, um zu belegen, dass sich die beiden Behandlungsgruppen auch in diesem klinischen Endpunkt nicht unterscheiden. Wie im Patientenkollektiv ebenfalls zu erwarten war, traten unter beiden Behandlungen nur sehr wenige Ereignisse einer Progression zur Klasse CDC C im Studienverlauf auf und die relative Häufigkeit unterschied sich nicht relevant und auch nicht statistisch signifikant zwischen den Behandlungen.

Damit stützen diese Ergebnisse die zum Endpunkt „Viruslast < 50 Kopien/ml zu Woche 48“ getroffenen Aussagen sehr gut.

Mortalität:

Aufgrund des eingeschlossenen Patientenkollektivs war auch das Auftreten von Todesfällen im Studienverlauf beider SWORD-Zulassungsstudien nur äußerst selten zu erwarten. Es traten in beiden SWORD-Studien nur insgesamt zwei Todesfälle auf, einer je Behandlungsgruppe. Ein Zusatznutzen für DTG/RPV wurde hinsichtlich der Mortalität nicht belegt.

Gesundheitsbezogene Lebensqualität:

Zur Bewertung der gesundheitsbezogenen Lebensqualität wurden die Endpunkte „EQ-5D“, „*Symptom Distress Module*“ (SDM) und „*HIV Treatment Satisfaction Questionnaire*“ (HIVTSQ) in beiden SWORD-Zulassungsstudien erhoben und die Ergebnisse jeweils zu Woche 48 dargestellt.

Der EQ-5D wurde hinsichtlich des *Utitlity Scores* und der visuellen Analogskala (VAS) separat dargestellt. Es ergaben sich jeweils bei nur sehr geringen Änderungen keine statistisch signfikanten Unterschiede zwischen den Behandlungsgruppen DTG/RPV und CAR. Ein Zusatznutzen ist damit aus den Ergebnissen des EQ-5D nicht belegt.

Für das SDM wurden zwei Dimensionen ausgewertet und dargestellt: der *Symptom Count Score*, der die Häufigkeit aufgetretener Symptome angibt und der *Symptom Bother Score*, bei dem der Patient die Belastung durch aufgetretene Symptome der Erkrankung einschätzt. Verringerungen in diesen beiden Scores zeigen jeweils Verbesserungen an. In beiden Scores treten numerische Verbesserungen zugunsten von DTG/RPV im Vergleich zu CAR auf, die für den *Symptom Bother Score* auch statistisch signifikant sind. Dies bedeutet in der Gesamtschau für das SDM, dass die Patienten im DTG/RPV-Arm im Vergleich zu CAR weniger Symptome berichten und diese zusätzlich statistisch signifikant als weniger belastend eingeschätzt werden, was eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung der Lebensqualität darstellt. Da das Ergebnis durch die Meta-Analyse von zwei hochwertigen RCT nachgewiesen wurde, ergibt sich die Aussagesicherheit eines Belegs. Insgesamt wird aus den Ergebnissen zum SDM daher ein Beleg für einen geringen Zusatznutzen für DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie (CAR) abgeleitet.

Für den HIVTSQ ergibt sich zu Woche 48 eine statistisch signifikante Verbesserung der Therapiezufriedenheit für DTG/RPV im Vergleich zu CAR. Dies bedeutet eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung der Lebensqualität. Es ist dabei hervorzuheben, dass die Behandlungszufriedenheit bei Baseline mit Werten von 54 bis zu 55 Punkten bei einem Maximalscore von 60 (entspricht höchster Therapiezufriedenheit) bereits sehr hoch lag und von daher Verbesserungen im Patientenkollektiv nur schwer zu erreichen waren, weshalb deren Eintreten besonders hoch zu bewerten ist. Da das Ergebnis durch die Meta-Analyse von zwei hochwertigen RCT nachgewiesen wurde, ergibt sich die Aussagesicherheit eines Belegs. Insgesamt wird aus den Ergebnissen des HIVTSQ daher ein Beleg für einen geringen Zusatznutzen für DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie (CAR) abgeleitet.

Sicherheit/Verträglichkeit:

Aufgrund des in den beiden Studien bestehenden Verzerrungspotenzials, ergibt sich die Notwendigkeit, insbesondere die Ergebnisse zur Sicherheit und Verträglichkeit kritisch zu betrachten. In den beiden offenen RCT SWORD-1 und SWORD-2 besteht ein potenzieller Bias zur kritischeren Beurteilung insbesondere von Sicherheit und Verträglichkeit Die Patienten in der CAR-Gruppe setzten ihre seit mindestens 6 Monaten bestehende, effektive (Viruslast < 50 Kopien/ml) und stabil ohne Nebenwirkungen und Toxizitäten verlaufende ART fort, wobei die mediane Dauer der bestehenden Therapie sogar noch sehr viel größerwar, als die geforderten 6 Monate. Der randomisierte Vergleich in beiden Studien erfolgt jedoch gegenüber einer für 87 % der auf DTG/RPV randomisierten Patienten komplett neuen ART, d.h. 87 % der Patienten hatten bisher noch keinen der beiden Wirkstoffe jemals erhalten und für 100 % der Patienten auf DTG/RPV war die Wirkstoff-Kombination aus DTG/RPV neu. Diese Zuteilung erfolgte offen, so dass Prüfärzte und Patienten diese kannten. Es ist daher naheliegend, dass primär im Sinne der Patientensicherheit und Aufrechterhaltung einer effektiven ART das Potential besteht, die neue Therapie mit DTG/RPV, trotz der Bekanntheit ihrer Einzelkomponenten, im Studienkontext besonders kritisch zu bewerten, um im Zweifelsfall möglichst schnell die Studie für einen Patienten abzubrechen, um den Patienten auf eine andere Therapie umstellen zu können. Auch die Meldung und Bewertung von unerwünschten Ereignissen, für die die subjektive Einschätzung, ob diese als bereits melderelevant anzusehen sind oder nicht, entscheidend sein kann, ist das Verzerrungspotenzial im hier vorliegenden offenen Studiendesigns kritisch zu bewerten.

Vor diesem Hintergrund wurden die Endpunkte „unerwünschte Ereignisse bis Woche 48“, „unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten“ sowie alle dargestellten „unerwünschten Ereignisse von besonderem Interesse bis Woche 48“ als mit hohem Verzerrungspotenzial behaftet eingestuft. Die Kategorien „unerwünschte Ereignisse vom Grad 3 oder 4 bis Woche 48“ und „schwerwiegende unerwünschte Ereignisse bis Woche 48“ wurden wegen ihrer Schwere bzw. der aus den Ereignissen ggfs. entstandenen Konsequenzen oder der erforderlichen Gegenmaßnahmen als nur mit geringem Verzerrungspotenzial behaftet eingestuft. Der spezifische Charakter schwerer Ereignisse lässt

das oben beschriebene Verzerrungspotenzial somit eher in den Hintergrund treten. Diese Einschätzung zum Verzerrungspotenzial der Endpunkte zur Sicherheit/Verträglichkeit steht auch in Übereinstimmung mit der Bewertung, die das IQWiG kürzlich im Rahmen der Nutzenbewertung von Darunavir/Cobicistat/Emtricitabin/Tenofoviralafenamid in der gleichen Indikation und einem vom Studiendesign her weitgehend vergleichbaren Setting getroffen hat (6).

In ihrer Bewertung der SWORD-Studien hat das CHMP (Committee for Medicinal Products for Human Use) im Zusammenhang mit der Bewertung der Sicherheit/Verträglichkeit ebenfalls deutlich gemacht, dass unterschiedliche Häufigkeiten gemeldeter unerwünschter Ereignisse (allgemein), von Therapieabbrüchen wegen unerwünschter Ereignisse und zu psychiatrischen unerwünschten Ereignissen in diesem unverblindeten Switch-Studiendesign – was zu teilweise häufigeren Meldungen im DTG/RPV-Arm im Vergleich zum CAR-Arm führte – sehr plausibel durch Verzerrungen (Reporting-Bias) bedingt sein können. Insgesamt werden die Arten der berichteten Ereignisse von den Einzelsubstanzen her als generell bekannt und auch (in der hier geprüften Kombination von DTG und RPV) nicht häufiger auftretend, als bei den Einzelsubstanzen, eingeschätzt. Dies führt das CHMP zu der summarischen Bewertung, dass keine zusätzlichen Risiken oder Sicherheitsprobleme identifiziert wurden. (7)

Vor diesem Hintergrund werden im Folgenden die Aspekte des Zusatznutzens hinsichtlich der Sicherheit/Verträglichkeit diskutiert. Zu den Endpunkten „unerwünschte Ereignisse bis Woche 48“ und „Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten“ liegen Ergebnisse vor, die eine höhere relative Häufigkeit unter DTG/RPV im Vergleich zu CAR beschreiben. Die Ergebnisse sind für beide Endpunkte statistisch signifikant.

Der meta-analytische Effektschätzer zu allen „unerwünschten Ereignissen bis Woche 48“ beschreibt trotz der statistischen Signifikanz einen sehr kleinen Effekt. Bei Invertieren der Effektrichtung im Vergleich zur Darstellung in Tabelle 4-3 ergibt sich eine obere Konfidenzintervallgrenze für das relative Risiko von 0,99. Damit liegt nach Einschätzung des IQWiG (wie kürzlich in der Nutzenbewertung für Darunavir/Cobicistat/Emtricitabin/Tenofoviralafenamid getroffen (6)), ein für diese Kategorie der nicht schwerwiegenden bzw. nicht schweren unerwünschten Ereignisse, ein nicht mehr als geringfügiges Schadenspotenzial für DTG/RPV vor. Unter zusätzlicher Berücksichtigung des zuvor beschriebenden hohen Verzerrungspotenzials dieses Endpunktes kann aus dem Ergebnis zu diesem Endpunkt kein Aspekt eines geringeren Nutzens des zu bewertenden Kombinationsarzneimittels DTG/RPV abgeleitet werden. Zur Häufigkeit aller unerwünschter Ereignisse konstatiert das CHMP ebenfalls, dass der zuungunsten von DTG/RPV verzerrende Einfluss des Switch-Designs eine plausible Erklärung für die berichteten unterschiedlichen Häufigkeiten ist (7).

Bei dem Endpunkt „Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten“ ist das Ausmaß des meta-analytischen Effektschätzers größer, jedoch liegt für diesen Endpunkt ebenfalls ein hohes Verzerrungspotenzial vor. Es liegt die Vermutung nahe, dass es sich um einen Trend handelt, der zu erheblichen Anteilen durch einen Reporting-Bias verursacht wird. Gerade offene RCT sowie die hier vorliegenden Studien (SWORD-1 und

SWORD-2) jeweils zusätzlich durch Vorhandensein des Switch-Designs, sind anfällig für diese Art von Bias. Da auch die Ergebnisse zur Lebensqualität, zum SDM sowie zu schwerwiegenden unerwünschten Ereignissen und zu unerwünschten Ereignissen vom Grad 3 oder 4 das Ergebnis dieses Endpunktes „Unerwünschten Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten“ nicht widerspiegeln, wird auch aus diesem Ergebnis kein Aspekt eines geringeren Nutzens des zu bewertenden Kombinationsarzneimittels DTG/RPV abgeleitet. Auch das CHMP bewertet die häufigeren Meldungen von unerwünschten Ereignissen, die zum Studienabbruch führten, als primär durch die Gegebenheiten des Switch-Designs verursacht und kommt zu dem Schluss, dass sich daraus kein Sicherheitsproblem ableiten lässt (7).

Die Analyse der Endpunkte „Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48“ sowie der beiden Kategorien unerwünschter Ereignisse von besonderem Interesse bis Woche 48 „Erkrankungen des Gastrointestinaltrakts“ und „Erkrankungen des Nervensystems“ ergaben, dass die Meta-Analysen über beide SWORD-Studien eine hohe Heterogenität aufwiesen. Wegen der begründeten Auswertung im Modell mit festen Effekten ist die Meta-Analyse zur Interpretation der Gesamtaussage daher nicht geeignet und es werden daher bei der Diskussion dieser Endpunkte sowie in der Gesamtschau die Ergebnisse der Einzelstudien herangezogen. Es zeigt sich für alle drei Endpunkte, dass ein statistisch signifikanter Effekt im Sinne einer höheren relativen Auftretenshäufigkeit unter DTG/RPV im Vergleich zu CAR in jeweils nur einer Studie (SWORD-2) auftritt. Die jeweils andere Studie (SWORD-1) zeigte ein statistisch nicht signifikantes Ergebnis, wobei ein Trend zuungunsten von DTG/RPV entweder nur sehr schwach (unerwünschte Ereignisse von besonderem Interesse „Erkrankungen des Gastrointestinaltrakts“ und „Erkrankungen des Nervensystems“) oder gar nicht („unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48“) vorlag. Es ist auch auffällig, dass für alle drei Endpunkte in der Studie SWORD-2 ein hauptsächlicher Unterschied in den Ergebnissen in einer im Vergleich zu SWORD-1 deutlich geringeren relativen Häufigkeit der jeweiligen Ereignisse im CAR-Arm lag und nicht etwa in einer Erhöhung im DTG/RPV-Arm. Die Ergebnisse im DTG/RPV-Arm waren hingegen für diese drei hier diskutierten Endpunkte über beide Studien sehr gut vergleichbar. Ohne, dass dafür eine plausible Erklärung identifiziert werden konnte, ist das Ergebnis der hohen Heterogenität der Meta-Analysen im Wesentlichen auf diesen Unterschied im CAR-Arm zwischen den beiden SWORD-Studien zurückzuführen. Folglich kann für diese drei Endpunkte jeweils nicht von einer klaren Evidenzsituation über die vorhandene Gesamtevidenz ausgegangen werden. Daher wird festgestellt, dass für diese Endpunkte weder ein Zusatznutzen noch ein geringerer Nutzen für DTG/RPV als belegt angesehen werden kann. Bei der Interpretation ist ebenfalls das hohe Verzerrungspotenzial bezüglich der beiden Kategorien der unerwünschten Ereignisse von besonderem Interesse zu berücksichtigen.

Für die beiden Kategorien unerwünschter Ereignisse von besonderem Interesse bis Woche 48 „Psychiatrische Erkrankungen“ und „Erkrankungen der Haut und des Unterhautzellgewebes“ ergaben sich bei hohem Verzerrungspotenzial in der Meta-Analyse jeweils statistisch signifikante Ergebnisse einer höheren relativen Häufigkeit dieser Ereignisse unter DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie CAR.

Bei dem Endpunkt „Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes“ lag zusätzlich ein Beleg auf Effektmodifikation hinsichtlich der CD4-Zellzahl bei Baseline vor. Die Effekte in den beiden Subgruppen waren nicht gleichgerichtet, so dass eine Interpretation je Subgruppe erfolgt. Nur in der Subgruppe der Patienten, die bei Baseline eine CD4-Zellzahl von mindestens 500 Zellen/mm^3 hatten, lag ein statistisch signifikantes Ergebnis für eine höhere relative Häufigkeit unter DTG/RPV vor. Im alternativen Stratum gab es einen leichten numerischen Trend in die entgegengesetzte Richtung (weniger unerwünschte Ereignisse der Kategorie unter DTG/RPV im Vergleich zu CAR), der jedoch statistisch nicht signifikant war. Für diesen Endpunkt waren die aufgetretenen Ereignisse für alle Patienten mit Ereignissen maximal vom Grad 1 oder 2 (bis auf einen Patienten im CAR-Arm).

Für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Ereignisse“ waren nur für insgesamt 5 Patienten in der Gruppe DTG/RPV vom Grad 3 oder 4 (das waren 8,2 % aller unerwünschten psychiatrischen Ereignisse unter DTG/RPV) im Vergleich zu 2 Patienten in der CAR-Gruppe (6,3 % aller unerwünschten psychiatrischen Ereignisse unter CAR). Damit waren in beiden Behandlungsgruppen deutlich über 90 % aller unerwünschten psychiatrischen Ereignisse vom Grad 1 oder 2 und der jeweilige Anteil der Ereignisse vom Grad 3 oder 4 unterscheidet sich auch nur geringfügig zwischen DTG/RPV und CAR.

Da es sich somit in der großen Mehrheit der Ereignisse in beiden hier diskutierten Endpunkten um leichte oder maximal mäßig schwere Ereignisse handelte wird – auch vor dem Hintergrund des hohen Verzerrungspotenzials der beiden Endpunkte – insgesamt daraus kein geringerer Nutzen für DTG/RPV als belegt angesehen.

Da generell keine neuen Arten von unerwünschten psychiatrischen Ereignissen aufgetreten sind als solche, die bereits von den Einzelsubstanzen DTG und RPV bekannt sind, und da auch die Häufigkeit dem Erwartungswert dafür entspricht, kommt das CHMP zum Schluss, dass es keinen Hinweis darauf gibt, dass DTG/RPV zu einem additiven oder supraadditiven Risiko für psychiatrische unerwünschte Ereignisse beiträgt (7).

Zu allen weiteren dargestellten Endpunkten der Sicherheit/Verträglichkeit (wie insbesondere schwerwiegenden unerwünschten Ereignissen und den weiteren dargestellten unerwünschten Ereigniskategorien von besonderem Interesse) gab es keine weiteren Aspekte, die auf eine zwischen DTG/RPV und der zweckmäßigen Vergleichstherapie CAR unterschiedliche Sicherheit/Verträglichkeit hinweisen würden.

Bezüglich des Endpunktes „Knochenmineraldichte“ zeigte sich an beiden Messlokalisationen (Hüfte und Lendenwirbel) in der DEXA-Substudie jeweils ein statistisch signifikanter Anstieg über 48 Wochen in der Gruppe DTG/RPV im Vergleich zu CAR. Die ebenfalls dargestellten knochenspezifischen Biomarker (BSAP, P1NP, CTx und Osteocalcin jeweils zu Woche 48) unterstützen dieses Ergebnis zur Knochenmineraldichte durch für jeden dieser Endpunkte in den Meta-Analysen gezeigte statistisch signifikante Verbesserungen unter DTG/RPV im

Vergleich zu CAR. Diese Endpunkte der knochenspezifischen Endpunkte werden jedoch nicht zur separaten Ableitung eines Zusatznutzens herangezogen.

Der Endpunkt Knochenmineraldichte besitzt prädiktive Wertigkeit für Knochenfrakturen. Die Auswertung erfolgte durch verblindete Bewerter, so dass für diesen Endpunkt kein hohes Verzerrungspotenzial vorliegt. Die Ergebnisse zu im Studienverlauf über 48 Wochen aufgetretenen Frakturen zeigen einen leichten positiven Trend zugunsten von DTG/RPV, der aber keine statistische Signifikanz aufweist. Die Aussagesicherheit des Endpunktes „Knochenmineraldichte“ wird (auf der Subpopulation zweier RCT beruhend) als Hinweis eingeschätzt. Insgesamt ergibt sich aus den Ergebnissen zum Endpunkt „Knochenmineraldichte“ damit ein Hinweis für einen geringen Zusatznutzen für DTG/RPV im Vergleich zur zweckmäßigen Vergleichstherapie.

Zusammenfassung für die Teilpopulation ohne Umstellindikation:

In der **Teilpopulation ohne Umstellindikation** ergibt sich zusammenfassend auf der Grundlage einer vollkommen vergleichbaren Wirksamkeit, einem in der Gesamtschau nicht belegten höheren oder geringeren Nutzen hinsichtlich der dargestellten Endpunkte zu Nebenwirkungen, einem Hinweis auf einen geringen Zusatznutzen hinsichtlich der Knochenmineraldichte und einem Beleg für einen geringen Zusatznutzen hinsichtlich der Lebensqualität in der Gesamtschau ein **Hinweis auf einen geringen Zusatznutzen** für das zu bewertende Kombinationsarzneimittel Dolutegravir/Rilpivirin im Vergleich zur zweckmäßigen Vergleichstherapie „Fortführung der bestehenden ART“.

Die Bewertung zu den diskutierten, teilweise häufigeren, Meldungen unerwünschter Ereignisse berücksichtigt auch die Einschätzung des CHMP, das zu dem Schluss kommt, dass diese plausibel als Folge des diesbezüglich zuungunsten von DTG/RPV verzerrenden Einflusses des (unverblindeten) Switch-Designs sind und daraus keine zusätzlichen Risiken oder Verträglichkeitsprobleme abzuleiten sind (7).

Trotzdem erfolgt die vorgenommene Herabstufung der Aussagesicherheit von einem möglichen Beleg auf einen Hinweis unter Berücksichtigung der Aspekte hinsichtlich der Heterogenität in drei Endpunkten der Sicherheit/Verträglichkeit zwischen den beiden Zulassungsstudien SWORD-1 und SWORD-2 und dem mit hohem Verzerrungspotenzial behafteten Ergebnissen zu dem Endpunkt „unerwünschte Ereignisse bis Woche 48, die zum Studienabbruch geführt hatten“.

Patienten der Zielpopulation mit Umstellindikation:

Für die **Teilpopulation mit Umstellindikation** lag keine RCT-Evidenz vor, somit kann für diese Teilpopulation kein Zusatznutzen begründet werden. Für die Teilpopulation mit Umstellindikation ist damit kein Zusatznutzen belegt.

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 4-158: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	**Ausmaß des Zusatznutzens**
Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen - **antiretroviral vorbehandelte Patienten ohne Umstellindikation.**	Hinweis auf geringen Zusatznutzen
Erwachsene, die mit dem humanen Immundefizienz-Virus Typ 1 (HIV-1) infiziert sind und die virologisch supprimiert sind (HIV-1 RNA < 50 Kopien/ml) und seit mindestens sechs Monaten auf einem stabilen antiretroviralen Regime eingestellt sind, kein virologisches Versagen in der Vergangenheit und keine bekannten oder vermuteten Resistenzen gegen Nicht-Nukleosidale Reverse-Transkriptase-Inhibitoren oder Integrase-Inhibitoren aufweisen - **antiretroviral vorbehandelte Patienten mit Umstellindikation.**	Kein Zusatznutzen belegt

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Es wurden keine indirekten Vergleiche vorgelegt.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Es wurden keine nicht randomisierten, vergleichenden Studien bzw. weitere Untersuchungen vorgelegt.

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.

Es wurden die Ergebnisse zweier randomisierter kontrollierter Studien vorgelegt, welche valide Daten zu patientenrelevanten Endpunkten umfassen.

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[13], Molenberghs 2010[14]). 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[15]) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der

[13] *Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.*
[14] *Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.*
[15] *Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.*

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[16]) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

– Patientenpopulation
– Intervention
– Kontrolle
– Datenherkunft
– verwendete Methodik
– entsprechende Ergebnisse (abhängig von der Methode)
– Untersuchungen zur Robustheit
– ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Der Endpunkt zum virologischen Ansprechen (Viruslast < 50 Kopien/ml) gilt im vorliegenden Anwendungsgebiet als anerkannter Surrogatparameter für das Auftreten von AIDS-definierenden Ereignissen oder Tod (8, 14, 15).

Wie schon in vorangegangenen Nutzenbewertungen im Indikationsgebiet HIV belegt, gilt das virologische Ansprechen (Absinken der Viruslast unter 50 Kopien/ml) als patientenrelevanter Endpunkt. Dies wurde vom G-BA in den Beschlüssen zu den Nutzenbewertungen insbesondere der Einzelsubstanz Dolutegravir (10) sowie der antiretroviralen Substanz Rilpivirin (11) sowie weiteren Bewertungen im vorliegenden Anwendungsgebiet anerkannt (z.B. kürzlich durchgeführte Bewertungen zu Emtricitabin/Tenofoviralafenamid oder Elvitegravir/Cobicistat/Emtricitabin/Tenofoviralafenamid) (12, 13).

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

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 Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Tabelle 4-159: Studienpool – Studien mit dem zu bewertenden Arzneimittel (Teilpopulation ohne Umstellindikation)

Studie	**verfügbare Quellen**			
	Studienberichte	**Registereinträge**	**Publikationen**	**Zusatzanalysen**
201636 SWORD-1	(40)	(29-33)	(34)	(46)
201637 SWORD-2	(41)	(35-39)	(34)	(47)
DEXA Substudie	(42)	(43, 44)	(45)	-

Für die **Teilpopulation mit Umstellindikation** konnte keine RCT identifiziert und in die Bewertung eingeschlossen werden.

4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. Gilead Sciences Limited International. Fachinformation zu Odefsey® 2017.
2. Gilead Sciences Limited International. Fachinformation zu Eviplera® 2017.
3. ViiV Healthcare. Fachinformation zu Triumeq. 2017.
4. Gemeinsamer Bundesausschuss G-BA. Niederschrift (finale Fassung) zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV. Beratungsanforderung 2017-B-128. 2017.
5. Gemeinsamer Bundesausschuss G-BA. Anschreiben zur Niederschrift (finale Fassung) zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV. Beratungsanforderung 2017-B-128. 2017.
6. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen IQWiG. Darunavir/Cobicistat/Emtricitabin/Tenofoviralafenamid - Nutzenbewertung gemäß § 35a SGB V. Dossierbewertung. Auftrag: A17-48, Version 1.0 vom 22.12.2017. 2017.
7. Committee for Medicinal Products for Human Use (CHMP). CHMP assessment report Juluca. 2018.
8. EMA, EMA. 2016. Guideline on the Clinical Development of Medicinal Products for the Treatment of HIV Infection. EMEA/CPMP/EWP/633/02 Rev. 3. 20. April 2016; Abrufbar unter: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/07/WC500209918.pdf [Zugriff am: 11.04.2018].
9. Wong, SS, Wilczynski, NL, Haynes, RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc. 2006;94(4):451-5.
10. Gemeinsamer Bundesausschuss G-BA. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschuss über eine Änderung Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Dolutegravir vom 7. August 2014. 2014.
11. Gemeinsamer Bundesausschuss G-BA. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschuss über eine Änderung Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Rilpivirin vom 5. Juli 2012. 2012.
12. Gemeinsamer Bundesausschuss G-BA. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschuss über eine Änderung Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Elvitegravir/Cobicistat/Emtricitabin/Tenofoviralafenamid vom 16. Juni 2016. 2016.
13. Gemeinsamer Bundesausschuss G-BA. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschuss über eine Änderung Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Emtricitabin/Tenofoviralafenamid vom 3. November 2016. 2016.
14. Deutsche AIDS-Gesellschaft DAIG. 2017. Deutsch-Österreichische Leitlinien zur antiretroviralen Therapie der HIV-Infektion. AWMF-Register-Nr.: 055-001. Version 7 vom 29.11.2017; Abrufbar unter: http://www.daignet.de/site-content/hiv-therapie/leitlinien-

1/Deutsch%20Osterreichische%20Leitlinien%20zur%20antiretroviralen%20Therapie%20der%20HIV%201%20Infektion.pdf [Zugriff am: 11.04.2018].
15. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen IQWiG. Rilpivirin - Nutzenbewertung gemäß § 35a SGB V. Dossierbewertung. Auftrag: A12-04, Version 1.0 vom 12.04.2012. 2012.
16. Mielck, A, Vogelmann, M, Schweikert, B, et al. Health status of adults in Germany: results from a representative survey using the EuroQol 5D (EQ-5D). Gesundheitswesen. 2010;72:476-86.
17. Herdman, M, Gudex, C, Lloyd, A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
18. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen IQWiG. Dolutegravir/Abacavir/Lamivudin - Nutzenbewertung gemäß § 35a SGB V. Dossierbewertung. Auftrag: A14-34, Version 1.0 vom 18.12.2014. 2014.
19. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen IQWiG. Dolutegravir - Nutzenbewertung gemäß § 35a SGB V. Dossierbewertung. Auftrag: A14-08, Version 1.0 vom 12.05.2014. 2014.
20. Justice, AC, Holmes, W, Gifford, AL, et al. Development and validation of a self-completed HIV symptom index. J Clin Epidemiol. 2001;54 Suppl 1:S77-90.
21. Marc, LG, Wang, MM, Testa, MA. Psychometric evaluation of the HIV symptom distress scale. AIDS Care. 2012;24:1432-41.
22. Woodcock, A, Bradley, C. Validation of the revised 10-item HIV Treatment Satisfaction Questionnaire status version and new change version. Value Health. 2006;9(5):320-33.
23. Battalora, L, Buchacz, K, Armon, C, et al. Low bone mineral density and risk of incident fracture in HIV-infected adults. Antivir Ther. 2016;21(1):45-54.
24. DerSimonian, R, Laird, N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-88.
25. Higgins, JP, Thompson, SG, Deeks, JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557-60.
26. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen IQWiG. 2017. Allgemeine Methoden Version 5.0 vom 10.07.2017; Abrufbar unter: https://www.iqwig.de/de/methoden/methodenpapier.3020.html [Zugriff am: 11.04.2018].
27. Knapp, G, Hartung, J. Improved tests for a random effects meta-regression with a single covariate. Stat Med. 2003;22(17):2693-710.
28. Higgins, JP, White, IR, Wood, AM. Imputation methods for missing outcome data in meta-analysis of clinical trials. Clin Trials. 2008;5(3):225-39.
29. Clinicaltrials.gov. Regimen Switch to Dolutegravir + Rilpivirine From Current Antiretroviral Regimen in Human Immunodeficiency Virus Type 1 Infected and Virologically Suppressed Adults (SWORD-1); Abrufbar unter: https://ClinicalTrials.gov/show/NCT02429791 [Zugriff am: 03.04.2018].
30. ICTRP. 2015. A Phase III, randomized, multicenter, parallel-group, noninferiority study evaluating the efficacy, safety, and tolerability of switching to dolutegravir plus rilpivirine from current INI-, NNRTI-, or PI-based antiretroviral regimen in HIV-1-infected adults who are virologically suppressed. SWORD-1; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005147-40 [Zugriff am: 03.04.2018].
31. ICTRP. 2015. Regimen Switch to Dolutegravir + Rilpivirine From Current Antiretroviral Regimen in Human Immunodeficiency Virus Type 1 Infected and Virologically

Suppressed Adults (SWORD-1); Abrufbar unter: https://clinicaltrials.gov/show/NCT02429791 [Zugriff am: 03.04.2018].
32. EUCTR. 2015. A Phase III, randomized, multicenter, parallel-group, noninferiority study evaluating the efficacy, safety, and tolerability of switching to dolutegravir plus rilpivirine from current INI-, NNRTI. SWORD-1; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005147-40 [Zugriff am: 03.04.2018].
33. PharmNet.Bund. A Phase III, randomized, multicenter, parallel-group, noninferiority study evaluating the efficacy, safety, and tolerability of switching to dolutegravir plus rilpivirine from current INI-, NNRTI-, or PI-based antiretroviral regimen in HIV-1-infected adults who are virologically suppressed. SWORD-1; [Zugriff am: 03.04.2018].
34. Llibre, JM, Hung, CC, Brinson, C, et al. Efficacy, safety, and tolerability of dolutegravir-rilpivirine for the maintenance of virological suppression in adults with HIV-1: phase 3, randomised, non-inferiority SWORD-1 and SWORD-2 studies. Lancet. 2018.
35. Clinicaltrials.gov. Regimen Switch to Dolutegravir + Rilpivirine From Current Antiretroviral Regimen in Human Immunodeficiency Virus Type 1 Infected and Virologically Suppressed Adults (SWORD-2); Abrufbar unter: https://ClinicalTrials.gov/show/NCT02422797 [Zugriff am: 03.04.2018].
36. ICTRP. 2015. A Phase III, randomized, multicenter, parallel-group, noninferiority study evaluating the efficacy, safety, and tolerability of switching to dolutegravir plus rilpivirine from current INI-, NNRTI-, or PI-based antiretroviral regimen in HIV-1-infected adults who are virologically suppressed. SWORD-2; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005148-16 [Zugriff am: 03.04.2018].
37. ICTRP. 2015. Regimen Switch to Dolutegravir + Rilpivirine From Current Antiretroviral Regimen in Human Immunodeficiency Virus Type 1 Infected and Virologically Suppressed Adults (SWORD-2); Abrufbar unter: https://clinicaltrials.gov/show/NCT02422797 [Zugriff am: 03.04.2018].
38. EUCTR. 2015. A Phase III, randomized, multicenter, parallel-group, noninferiority study evaluating the efficacy, safety, and tolerability of switching to dolutegravir plus rilpivirine from current INI-, NNRTI. SWORD-2; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005148-16 [Zugriff am: 03.04.2018].
39. PharmNet.Bund. A Phase III, randomized, multicenter, parallel-group, noninferiority study evaluating the efficacy, safety, and tolerability of switching to dolutegravir plus rilpivirine from current INI-, NNRTI-, or PI-based antiretroviral regimen in HIV-1-infected adults who are virologically suppressed. SWORD-2; [Zugriff am: 03.04.2018].
40. ViiV Healthcare. Clinical Study Report of the SWORD-1 Study (201636). A Phase III, Randomized, Multicenter, Parallel-Group, Non-Inferiority Study Evaluating the Efficacy, Safety, and Tolerability of Switching to Dolutegravir plus Rilpivirine from Current INI-, NNRTI-, or PI-Based Antiretroviral Regimen in HIV-1-Infected Adults who are Virologically Suppressed. 03.05.2017. 2017.
41. ViiV Healthcare. Clinical Study Report of the SWORD-2 Study (201637). A Phase III, Randomized, Multicenter, Parallel-Group, Non-Inferiority Study Evaluating the Efficacy, Safety, and Tolerability of Switching to Dolutegravir plus Rilpivirine from Current INI-, NNRTI-, or PI-Based Antiretroviral Regimen in HIV-1-Infected Adults who are Virologically Suppressed. 03.05.2017. 2017.

42. ViiV Healthcare. Clinical Study Report of the DEXA Substudy. An Evaluation of Bone Mineral Density in HIV-1-Infected Adult Subjects Switching from a Tenofovir-Containing Antiretroviral Therapy Regimen to a Dolutegravir plus Rilpivirine Regimen. 03.05.2017. 2017.
43. Clinicaltrials.gov. Bone Mineral Density in Human Immunodeficiency Virus Type 1 (HIV-1)-Infected Adult Subjects Switching From a Tenofovir Regimen to a Dolutegravir Plus Rilpivirine Regimen; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02478632 [Zugriff am: 03.04.2018].
44. ICTRP. 2018. Bone Mineral Density in Human Immunodeficiency Virus Type 1 (HIV-1)-Infected Adult Subjects Switching From a Tenofovir Regimen to a Dolutegravir Plus Rilpivirine Regimen; Abrufbar unter: https://clinicaltrials.gov/show/NCT02478632 [Zugriff am: 03.04.2018].
45. McComsey, GA, Lupo, S, Parks, D, et al. Switch from tenofovir disoproxil fumarate combination to dolutegravir plus rilpivirine improves parameters of bone health. AIDS. 2017.
46. ViiV Healthcare. Zusatzanalysen der Studie SWORD-1 zur Dossiererstellung. 2017.
47. ViiV Healthcare. Zusatzanalysen der Studie SWORD-2 zur Dossiererstellung. 2017.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. 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. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[17]] – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

[17] Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Datenbankname	MEDLINE
Suchoberfläche	ProQuest Dialog
Datum der Suche	03.04.2018
Zeitsegment	Keine zeitliche Beschränkung
Suchfilter	Filter RCT nach: Wong SS et al. 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: 451-5. Table 1: „Strategy minimizing difference between sensitivity and specificity“. Angepasst an die Kommandosprache von ProQuest.

#	Suchbegriffe	Ergebnis
S1	dolutegravir*	581
S2	(*GSK1349572* OR "*GSK 1349572*")	45
S3	(TIVICAY OR DTG)	1406
S4	(S1 OR S2 OR S3)	1828
S5	MESH.EXACT("Rilpivirine")	263
S6	rilpivirin*	528
S7	(*TMC278* OR "*TMC278*" OR R278474 OR "R 278474")	65
S8	(edurant OR RPV)	679
S9	(S5 OR S6 OR S7 OR S8)	1073
S10	(S4 AND S9)	75
S11	juluca	0
S12	(S10 OR S11)	75
S13	DTYPE("randomized controlled trial")	456753
S14	randomised OR randomized	761285
S15	placebo	207368
S16	(S13 OR S14 OR S15)	826329
S17	(MESH.EXACT.EXPLODE("Animals") NOT MESH.EXACT.EXPLODE("HUMANS"))	4438989
S18	S16 NOT S17	788887
S19	S12 AND S18	12

Datenbankname	EMBASE
Suchoberfläche	ProQuest Dialog
Datum der Suche	03.04.2018
Zeitsegment	Keine zeitliche Beschränkung
Suchfilter	Filter RCT nach: Wong SS et al. 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: 451-5. Table 1: „Strategy minimizing difference between sensitivity and specificity“. Angepasst an die Kommandosprache von ProQuest.

#	Suchbegriffe	Ergebnis
S20	EMB.EXACT("dolutegravir")	1419
S21	dolutegravir*	1517
S22	*GSK1349572* OR "*GSK 1349572*"	122
S23	tivicay OR DTG	2007
S24	S20 OR S21 OR S22 OR S23	3153
S25	EMB.EXACT("rilpivirine")	1743
S26	rilpivirin*	1958
S27	*TMC278* OR "*TMC278*"	292
S28	edurant OR RPV	1720
S29	S25 OR S26 OR S27 OR S28	3299
S30	S24 AND S29	387
S31	EMB.EXACT("dolutegravir plus rilpivirine")	20
S32	juluca	1
S33	S30 OR S31 OR S32	387
S34	pub(random*) OR ab(random*) OR ti(random*)	1290970
S35	pub(double NEAR/1 blind*) OR ab(double NEAR/1 blind*) OR ti(double NEAR/1 blind*)	190957
S36	placebo*	418678
S37	EMB.EXACT.EXPLODE("placebo")	353951
S38	S34 OR S35 OR S36 OR S37	1541972
S39	S33 AND S38	52
S40	S19 OR S39	64
S41	S19 OR S39 – mit Duplikat-Entfernung	56[a]

[a] Dublikate zwischen MEDLINE und EMBASE werden in diesem Schritt entfernt. Insgesamt resultieren 7 Treffer aus MEDLINE und 49 Treffer aus EMBASE.

Datenbankname	Cochrane Central Register of Controlled Trials	
Suchoberfläche	Cochrane Library	
Datum der Suche	03.04.2018	
Zeitsegment	Keine zeitliche Beschränkung	
Suchfilter	Keine Suchfilter verwendet	
#	**Suchbegriffe**	**Ergebnis**
1	dolutegravir or DTG or tivicay or GSK1349572 or "GSK 1349572":ti,ab,kw (Word variations have been searched)	176
2	Rilpivirine or RPV or edurant or TMC278 or "TMC 278":ti,ab,kw (Word variations have been searched)	180
3	#1 and #2	16
4	juluca:ti,ab,kw (Word variations have been searched)	0
5	#3 or #4	16

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht relevant, da keine indirekten Vergleiche durchgeführt wurden.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant, da keine nicht randomisierten vergleichenden Studien in die Ergebnisdarstellung dieses Dossiers einbezogen wurden.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht relevant, da keine weiteren Untersuchungen dargestellt wurden.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. 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. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer	23

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Studienregister	Treffer insgesamt	Anzahl Duplikate	Treffer identifiziert
clinicaltrials.gov	26	4	22
EU Clinical Trials Register	13	0	13
WHO – ICTRP-Portal	14	0	14
PharmNet.Bund	4	0	4

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	03.04.2018
Suchstrategie	(dolutegravir OR DTG OR tivicay) AND (Rilpivirine OR RPV OR edurant) [other terms]
Treffer	22

Studienregister	EU Clinical Trials Register
Internetadresse	www.clinicaltrialsregister.eu
Datum der Suche	03.04.2018
Suchstrategie	(dolutegravir OR DTG OR tivicay) AND (Rilpivirine OR RPV OR edurant)
Treffer	13

Studienregister	WHO – ICTRP-Portal
Internetadresse	http://apps.who.int/trialsearch/Default.aspx
Datum der Suche	03.04.2018
Suchstrategie	Rilpivirine AND dolutegravir
Treffer	14

Studienregister	PharmnetBund
Internetadresse	http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm
Datum der Suche	03.04.2018
Suchstrategie	Dolutegravir? [active substance] AND rilpivirin? [active substance]
Treffer	4

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht relevant, da keine indirekten Vergleiche durchgeführt wurden.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant, da keine nicht randomisierten vergleichenden Studien in die Ergebnisdarstellung dieses Dossiers einbezogen wurden.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht relevant, da keine weiteren Untersuchungen dargestellt wurden.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der / den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nr.	Publikation	Auschlussgrund
1.	Boswell, R, Foisy, MM, Hughes, CA. Dolutegravir Dual Therapy as Maintenance Treatment in HIV-Infected Patients: A Review. Annals of pharmacotherapy. 2018.	A5
2.	Capetti, AF, Astuti, N, Cattaneo, D, Rizzardini, G. Pharmacokinetic drug evaluation of dolutegravir plus rilpivirine for the treatment of HIV. Expert Opinion on Drug Metabolism and Toxicology. 2017;13(11):1183-92.	A5
3.	Capetti, AF, Cossu, MV, Paladini, L, Rizzardini, G. Dolutegravir plus rilpivirine dual therapy in treating HIV-1 infection. Expert Opinion on Pharmacotherapy. 2017:1-13.	A5
4.	Gianotti, N, Poli, A, Nozza, S, Galli, L, Galizzi, N, Ripa, M, et al. Durability of switch regimens based on rilpivirine or on integrase inhibitors, both in association with tenofovir and emtricitabine, in HIV-infected, virologically suppressed patients. BMC Infectious Diseases. 2017;17(1).	A2
5.	Llibre, JM, Hung, C-C, Brinson, C, Castelli, F, Girard, P-M, Kahl, L, et al. Phase III sword 1&2: Switch to DTG+RPV maintains virologic suppression through 48 wks. Topics in Antiviral Medicine. 2017;25(1):19s-20s.	A6
6.	McComsey, GA, Lupo, S, Parks, D, Poggio, MC, De Wet, J, Kahl, LP, et al. Switch from tenofovir disoproxil fumarate combination to dolutegravir with rilpivirine improves parameters of bone health. AIDS. 2018;32(4):477-85.	A5
7.	Sun, H-Y, Chang, S-Y, Hung, C-C. Dolutegravir-rilpivirine coformulation. Current Opinion in HIV and AIDS. 2018.	A5
8.	Sebaaly, JC, Kelley, D. Single-Tablet Regimens for the Treatment of HIV-1 Infection. Annals of pharmacotherapy. 2017;51:332-44.	A2
9.	Palacios, R, Mayorga, M, Gonzalez-Domenech, CM, Hidalgo, TC, Galvez, C, Munoz, ML, et al. Safety and efficacy of dolutegravir plus rilpivirine (DTG/RPV) in treatment-experienced HIV-infected patients: preliminary results at 24 weeks of the DORIVIR study. Journal of the international AIDS societyConference: international congress of drug therapy in HIV infection 2016United kingdom. 2016;19:54.	A5

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht relevant, da keine indirekten Vergleiche durchgeführt wurden.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant, da keine nicht randomisierten vergleichenden Studien in die Ergebnisdarstellung dieses Dossiers einbezogen wurden.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht relevant, da keine weiteren Untersuchungen dargestellt wurden.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nr.	Registertreffer	Ausschlussgrund
1.	Clinicaltrials.gov. Bioequivalence Study of a Fixed-dose Combination (FDC) of Dolutegravir (DTG) and Rilpivirine (RPV); Abrufbar unter: https://ClinicalTrials.gov/show/NCT02741557 [Zugriff am: 03.04.2018].	A4
2.	Clinicaltrials.gov. Relative Oral Bioavailability Study of Different Fixed Dose Combinations of Dolutegravir and Rilpivirine in Healthy Subjects; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02373930 [Zugriff am: 03.04.2018].	A4
3.	Clinicaltrials.gov. A Study to Evaluate the Pharmacokinetics and Safety of GSK1265744 and Rilpivirine and Dolutegravir and Rilpivirine in Healthy Adult Subjects; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01467531 [Zugriff am: 03.04.2018].	A5
4.	Clinicaltrials.gov. Dolutegravir + Rilpivirine Switch Study (DORISS); Abrufbar unter: https://ClinicalTrials.gov/show/NCT02069834 [Zugriff am: 03.04.2018].	A7
5.	Clinicaltrials.gov. Bone Mineral Density in Human Immunodeficiency Virus Type 1 (HIV-1)-Infected Adult Subjects Switching From a Tenofovir Regimen to a Dolutegravir Plus Rilpivirine Regimen; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02478632 [Zugriff am: 03.04.2018].	A5
6.	Clinicaltrials.gov. Evaluation of Neurologic and Psychiatric Adverse Events of Several Antiretroviral Drugs in Real Life Setting; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02882230 [Zugriff am: 03.04.2018].	A2
7.	Clinicaltrials.gov. Study to Evaluate the Efficacy, Safety, and Tolerability of Long-acting Intramuscular Cabotegravir and Rilpivirine for Maintenance of Virologic Suppression Following Switch From an Integrase Inhibitor in HIV-1 Infected Therapy Naive Participants; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02938520 [Zugriff am: 03.04.2018].	A2
8.	Clinicaltrials.gov. Population Pharmacokinetics of Antiretroviral in Children; Abrufbar unter: https://ClinicalTrials.gov/show/NCT03194165 [Zugriff am: 03.04.2018].	A5
9.	Clinicaltrials.gov. Characterization of Acute and Recent HIV-1 Infections in Zurich: a Long-term Observational Study; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00537966 [Zugriff am: 03.04.2018].	A5
10.	Clinicaltrials.gov. Study Evaluating the Efficacy, Safety, and Tolerability of Switching to Long-acting Cabotegravir Plus Long-acting Rilpivirine From Current Antiretroviral Regimen in Virologically Suppressed HIV-1-infected Adults; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02951052 [Zugriff am: 03.04.2018].	A2

11.	Clinicaltrials.gov. Antiretroviral Treatment Taken 4 Days Per Week Versus Continuous Therapy 7/7 Days Per Week in HIV-1 Infected Patients; Abrufbar unter: https://ClinicalTrials.gov/show/NCT03256422 [Zugriff am: 03.04.2018].	A7
12.	Clinicaltrials.gov. Safety and Efficacy of Switching From Regimens of ABC/3TC + a 3rd Agent to E/C/F/TAF Fixed-Dose Combination (FDC) in Virologically-Suppressed HIV 1 Infected Adults; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02605954 [Zugriff am: 03.04.2018].	A2
13.	Clinicaltrials.gov. Safety & Efficacy of Dual Therapy With Raltegravir/Lamivudine; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02284035 [Zugriff am: 03.04.2018].	A2
14.	Clinicaltrials.gov. Safety and Biological Activity of Vesatolimod in HIV-1 Infected, Virologically Suppressed Adults; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02858401 [Zugriff am: 03.04.2018].	A2
15.	Clinicaltrials.gov. Trial to Evaluate the Interest of a Reductive Anti Retroviral Strategy Using Dual Therapy Inspite of Triple Therapy; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02302547 [Zugriff am: 03.04.2018].	A7
16.	Clinicaltrials.gov. Switch Study to Evaluate F/TAF in HIV-1 Positive Participants Who Are Virologically Suppressed on Regimens Containing FTC/TDF; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02121795 [Zugriff am: 03.04.2018].	A2
17.	Clinicaltrials.gov. Switch Study to Evaluate F/TAF in HIV-1 Infected Adults Who Are Virologically Suppressed on Regimens Containing ABC/3TC; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02469246 [Zugriff am: 03.04.2018].	A7
18.	Clinicaltrials.gov. Pharmacokinetics, Safety, and Efficacy of Cobicistat-boosted Atazanavir or Cobicistat-boosted Darunavir in HIV-1 Infected, Treatment-Experienced, Virologically Suppressed Pediatric Subjects; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02016924 [Zugriff am: 03.04.2018].	A7
19.	Clinicaltrials.gov. Switching From a Tenofovir Disoproxil Fumarate (TDF) Containing Regimen to Elvitegravir/Cobicistat/Emtricitabine/ Tenofovir Alafenamide (E/C/F/TAF) Fixed-Dose Combination (FDC) in Virologically-Suppressed, HIV-1 Infected Adults Aged ƒ%¾ 60 Years; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02616783 [Zugriff am: 03.04.2018].	A2
20.	Clinicaltrials.gov. Interaction With HIV Antiretroviral Agents; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02556268 [Zugriff am: 03.04.2018].	A5
21.	Euctr. 2009. Study on Pharmacokinetics of newly developed ANtiretroviral agents in HIV-infected pregNAnt women (PANNA); Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-006158-16 [Zugriff am: 03.04.2018].	A7
22.	Euctr. 2009. Strategic Timing of AntiRetroviral Treatment(START); Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-006439-12 [Zugriff am: 03.04.2018].	A7
23.	Euctr. 2012. TAILoR ƒ_" (TelmisArtan and InsuLin Resistance in HIV): A Dose-Ranging Phase II Randomised Open-Labelled Trial of Telmisartan as a strategy for the Reduction of Insulin Resistance in HIV-Positive Ind; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-000935-18 [Zugriff am: 03.04.2018].	A7

24.	Euctr. 2014. A Phase II, Randomized, Open-Label Clinical Trial to Study the Efficacy and Safety of the Combination Regimen of MK-5172 and MK-3682 with Either MK-8742 or MK-8408 in Subjects with Chronic HCV GT3; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-003347-35 [Zugriff am: 03.04.2018].	A4
25.	Euctr. 2015. An open-label, randomized, controlled clinical trial to assess the safety, tolerability and efficacy of two dolutegravir-based simplification strategies in HIV-infected patients with prolonged viro; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-000274-35 [Zugriff am: 03.04.2018].	A7
26.	Euctr. 2015. A Phase II, Randomized, Open-Label Clinical Trial to Study the Efficacy and Safety of the Combination Regimen of MK-5172 and MK-3682 with Either MK-8742 or MK-8408 in Subjects with Chronic HCV GT1; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-003304-73 [Zugriff am: 03.04.2018].	A2
27.	Euctr. 2016. A Phase III, Randomized, Multicenter, Parallel-group, Open- Label Study Evaluating the Efficacy, Safety, and Tolerability of Long-Acting Intramuscular Cabotegravir and Rilpivirine for Maintenance o; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-001646-25 [Zugriff am: 03.04.2018].	A2
28.	Euctr. 2016. A randomised study of interferon-free treatment for recently acquired hepatitis C in people who inject drugs and people with HIV coinfection; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-004243-39 [Zugriff am: 03.04.2018].	A2
29.	Euctr. 2017. Randomized, open-label and multicentric trial evaluating the non-inferiority of antiretroviral treatment taken 4 consecutive days per week versus continuous therapy 7/7 days per week in HIV-1 infec; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-000040-17 [Zugriff am: 03.04.2018].	A7
30.	Euctr. 2017. A Phase 3 Clinical Study to Evaluate the Efficacy and Safety of the Combination Regimen of MK-3682B (Grazoprevir/Ruzasvir/Uprifosbuvir) in Participants with Chronic Hepatitis C Virus Genotype 3 Inf; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-001463-21 [Zugriff am: 03.04.2018].	A2
31.	Euctr. 2017. A PHASE III, OPEN-LABEL, MULTICENTRIC CLINICAL TRIAL OF A SINGLE ARM OF 16 LENGTHS OF TIME TO EVALUATE RETENTION WITH ELBASVIR / GRAZOPREVIR PLUS SOFOSBUVIR AND RIBAVIRIN IN PATIENTS WITH HEPATITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-000432-34 [Zugriff am: 03.04.2018].	A2
32.	ICTRP. 2007. Characterization of Acute and Recent HIV-1 Infections in Zurich: a Long-term Observational Study; Abrufbar unter: https://clinicaltrials.gov/show/NCT00537966 [Zugriff am: 03.04.2018].	A5
33.	ICTRP. 2011. A Study to Evaluate the Pharmacokinetics and Safety of GSK1265744 and Rilpivirine and Dolutegravir and Rilpivirine in Healthy Adult Subjects; Abrufbar unter: http://clinicaltrials.gov/show/NCT01467531 [Zugriff am: 03.04.2018].	A5
34.	ICTRP. 2014. Dolutegravir + Rilpivirine Switch Study (DORISS); Abrufbar unter: https://clinicaltrials.gov/show/NCT02069834 [Zugriff am: 03.04.2018].	A4

35.	ICTRP. 2015. Relative Oral Bioavailability Study of Different Fixed Dose Combinations of Dolutegravir and Rilpivirine in Healthy Subjects; Abrufbar unter: https://clinicaltrials.gov/show/NCT02373930 [Zugriff am: 03.04.2018].	A4
36.	ICTRP. 2015. Bone Mineral Density in Human Immunodeficiency Virus Type 1 (HIV-1)-Infected Adult Subjects Switching From a Tenofovir Regimen to a Dolutegravir Plus Rilpivirine Regimen; Abrufbar unter: https://clinicaltrials.gov/show/NCT02478632 [Zugriff am: 03.04.2018].	A5
37.	ICTRP. 2016. 201584 or *f*_oFLAIR*f*__ is a study to compare two drugs (called Cabotegravir and Rilpivirine, as oral tablets followed by long-acting injections) to a single tablet regimen containing 3 HIV drugs (called Triumeq); Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-001646-25 [Zugriff am: 03.04.2018].	A2
38.	ICTRP. 2016. Bioequivalence Study of a Fixed-dose Combination (FDC) of Dolutegravir (DTG) and Rilpivirine (RPV); Abrufbar unter: https://clinicaltrials.gov/show/NCT02741557 [Zugriff am: 03.04.2018].	A4
39.	ICTRP. 2016. Evaluation of Neurologic and Psychiatric Adverse Events of Several Antiretroviral Drugs in Real Life Setting; Abrufbar unter: https://clinicaltrials.gov/show/NCT02882230 [Zugriff am: 03.04.2018].	A5
40.	ICTRP. 2017. Antiretroviral regimen 4 days a week; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-000040-17 [Zugriff am: 03.04.2018].	A7
41.	ICTRP. 2017. Population Pharmacokinetics of Antiretroviral in Children; Abrufbar unter: https://clinicaltrials.gov/show/NCT03194165 [Zugriff am: 03.04.2018].	A5
42.	PharmNet.Bund. A Phase III, Randomized, Multicenter, Parallel-group, Open-Label Study Evaluating the Efficacy, Safety, and Tolerability of Long-Acting Intramuscular Cabotegravir and Rilpivirine for Maintenance of Virologic Suppression Following Switch from an Integrase Inhibitor Single Tablet Regimen in HIV-1 Infected Antiretroviral Therapy Naive Adult Participants; [Zugriff am: 03.04.2018].	A2
43.	PharmNet.Bund. Study on Pharmacokinetics of newly developed ANtiretroviral agents in HIV-infected pregNAnt women (PANNA); [Zugriff am: 03.04.2018].	A1

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht relevant, da keine indirekten Vergleiche durchgeführt wurden.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant, da keine nicht randomisierten vergleichenden Studien in die Ergebnisdarstellung dieses Dossiers einbezogen wurden.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht relevant, da keine weiteren Untersuchungen dargestellt wurden.

Anhang 4-E: 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 4-160 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 4-160 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-160 (Anhang): Studiendesign und -methodik für Studie 201636 (SWORD-1)

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primäres Ziel: Zeigen der Nichtunterlegenheit hinsichtlich der antiretroviralen Aktivität bei Wechsel zu DTG + RPV einmal täglich im Vergleich zum Fortführen der aktuellen antiretroviralen Therapie über 48 Wochen bei HIV-1 infizierten antiretroviral vorbehandelten Patienten.
Methoden		
3	Studiendesign	

3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multizentrische, randomisierte, offene, aktiv-kontrollierte Nichtunterlegenheitsstudie der Phase III; Randomisierung im Verhältnis 1:1 stratifiziert nach Baseline *Third Agent* (INI, NNRTI, PI), Alter (< 50 Jahre, ≥ 50 Jahre) und geplanter Teilnahme an der DEXA-Substudie. Der primäre Endpunkt wurde nach der randomisierten Phase (*Early Switch Phase*) nach 48 Wochen erhoben. Anschließend wurden auch die Patienten im Kontrollarm auf DTG+RPV umgestellt, wenn sie eine Viruslast < 50c/ml hatten. Diese *Late Switch Phase* dauert bis zu Woche 148.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	**Amendment 1** (26.02.2015) [vor Einschluss des ersten Patienten]: Hinzufügen zusätzlicher pharmakokinetischer Studienbesuche; Hinzufügen des Stratifikationsfaktors aufgrund der geplanten Teilnahme an der DEXA-Substudie; kleinere Änderungen der Ein- und Ausschlusskriterien; Änderung der Beschreibung des Monitorings des unabhängigen Kommittees bzgl. der Kriterien für einen Studienabbruch aufgrund von virologischen Gründen; Anpassungen bzgl. nicht zugelassener Begleitmedikation; Anpassung der Definition des Studienendes; Anpassungen des suizidalen Risikomonitorings; Überarbeitung der Tabellen bzgl. Zeiten und Ereignisse. **Amendment 2** (08.06.2015): Angleichen der Gründe für einen Wechsel der PI-Klasse an andere retrovirale Klassen; Verdeutlichung des Ausschlusskriteriums 10 bzgl. einer Hepatitis B Infektion; Überarbeitung der stratifizierten Auswertemethodik des primären Endpunkts; Überarbeitung der Kriterien für einen vorzeitigen Studienabbruch aufgrund von virologischen Gründen; Überarbeitung der Strategie zur Risikoverringerung bzgl. Hautausschlag vom Grad 3 oder 4 und dadurch bedingtem Studienabbruch.
4	Probanden / Patienten	

4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**Einschlusskriterien:** 1) HIV-1 infizierte Männer oder Frauen ≥ 18 Jahre. 2) Aktuelle antiretrovirale Kombinationstherapie, die für mind. 6 Monate stabil war (erste oder zweite ART des Patienten); Frühere Wechsel der ART sollten aufgrund von Verträglichkeit/Sicherheit oder Zugang zu Medikamenten oder Vereinfachung der Therapie geschehen sein. Folgende Kombinationstherapien vor dem Screeningzeitpunkt waren zulässig: 2 NRTI plus: INI oder NNRTI (jeweils als erste oder zweite ART) bzw. geboosterter PI (oder Atazanavir ohne Booster) als erste ART. 3) Mind. zwei dokumentierte Messungen der HIV-1 Viruslast < 50 c/ml im Plasma innerhalb der letzten 12 Monate vor Screening; eine innerhalb Monat 6 bis 12 und die andere innerhalb der 6 Monate vor Screening. 4) HIV-1 Viruslast < 50 c/ml zum Screeningzeitpunkt. 5) Frauen im gebärfähigen Alter müssen jeweils einen negativen Schwangerschaftstest zum Screening und an Tag 1 aufweisen und den protokollgemäßen Kontrazeptionsvorgaben zustimmen. 6) Vorhandensein von Wille und Fähigkeit, die Anforderungen der Studienteilnahme zu verstehen und Vorliegen einer unterschriebenen und datierten Einverständniserklärung vor Screening. 7) Patienten in Frankreich mussten entweder Angehöriger oder Leistungsempfänger einer Sozialversicherung sein. **Ausschlusskriterien:** 1) Jedwede Messung der HIV-1 Viruslast ≥ 50 c/ml innerhalb 6 Monate vor Screening und nach bestätigter Suppression auf < 50 c/ml unter der aktuellen ART. 2) Jedwede Messung der HIV-1 Viruslast > 200 c/ml innerhalb 6 bis 12 Monaten vor Screening und nach bestätigter Suppression auf < 50 c/ml unter der aktuellen ART. 3) Zwei oder mehr Messungen der HIV-1 Viruslast ≥ 50 c/ml innerhalb 6 bis 12

		Monaten vor Screening und nach bestätigter Suppression auf < 50 c/ml unter der aktuellen ART. 4) Vorliegen einer Pause von mehr als einem Monat, in der alle antiretroviralen Medikamente abgesetzt wurden (wegen Verträglichkeit/Sicherheit) im Zeitraum bis 6 Monate vor Screening. 5) Jedweder Wechsel zu einem Zweitlinienregime aufgrund virologischem Therapieversagen (Viruslast ≥ 400 c/ml nach anfänglicher Suppression unter 50 c/ml während der Erstlinientherapie). 6) Schwangere oder stillende Frauen, oder solche, die eine Schwangerschaft während der Studienlaufzeit planen. 7) Vorliegen von Anzeichen einer aktiven CDC Kategorie C Krankheit. Ausnahmen hierbei sind: kutanes Kaposi's Sarkom ohne notwendiger systemischer Therapie und einer früheren CD4+ Zellzahl von < 200 pro mm^3. 8) Patienten mit jedweder hepatischer Beeinträchtigung. 9) Patienten mit positivem Hepatitis B Test oder mit einer absehbaren Therapie gegen Hepatitis C während der Studienlaufzeit. 10) Zurückliegende oder aktuelle Allergie gegen die Studienmedikation oder eine derer Komponenten in der jeweiligen Klasse. 11) Bestehende Malignität außer kutanes Kaposi's Sarkom, basales Zellkarzinom, entferntes nicht-invasives kutanes Plattenepithelkarzinom, intraepitheliale Neoplasie des Gebärmutterhalses; andere Malignitäten mussten durch den Prüfarzt und den Monitor beurteilt werden. 12) Patienten mit erhöhtem Suizidrisiko (nach Einschätzung durch den Prüfarzt). 13) Jedweder existierender körperlicher oder mentaler Zustand, der sich auf die Teilnahme an der Studie (Protokollvorgaben, Compliance) negativ auswirkt bzw. die Sicherheit des Patienten beeinträchtigt (nach Beurteilung des Prüfarztes). 14) Jedweder Zustand, der die Absorption, Distribution, Metabolisierung oder das Ausscheiden der Studienmedikation beeinflusst oder der es für den Patienten

		unmöglich macht, die orale Medikation einzunehmen (nach Beurteilung des Prüfarztes). 15) Einnahme von Medikation gegen Torsades de Pointes. 16) Vakzinierung mit einer HIV-1 Immuntherapie innerhalb von 90 Tagen vor Screening. 17) Behandlung mit einer der folgenden Therapien innerhalb von 28 Tagen vor Screening: Strahlentherapie, zytotoxische Chemotherapeutika, jedwede Immunmodulatoren, die die Immunantwort beeinflussen. 18) Exposition einer experimentellen Medikation oder eines Impfstoffs innerhalb des längsten Zeitintervalls vor Studienbeginn, dessen Länge definiert ist: entweder 28 Tage oder die fünffache Halbwertszeit der Testsubstanz oder die zweifache Dauer des biologischen Effekts. 19) Teilnahme an einer anderen interventionellen Studie nach Randomisierung mit Ausnahme der DEXA-Substudie. 20) Frühere Behandlung mit einem Therapieregime, das entweder aus nur einem NNRTI (auch wenn nur pari-partum) oder aus nur einem oder zwei NRTI bestand. 21) Frühere oder aktuelle Behandlung mit Etravirin. 22) Frühere Behandlung mit Tipranavir/Ritonavir oder Fosamprenavir/Ritonavir. 23) Patienten, die eine laut Studienprotokoll nicht zulässige Medikation erhalten und diese nicht ersetzten können oder wollen. 24) Nachweis des Vorliegens von resistenz-assoziierten Mutationen gegenüber PI, INI, NRTI oder NNRTI oder Substitution R263K gegenüber INI. 25) Jedwede Laborabweichung von Grad 4 (außer Grad 4 Lipidwerte). Eine einfache Testwiederholung während der Screeningphase war zulässig als Nachweis. 26) Jedwede Laborabweichung zum Screeningzeitpunkt, die eine Studienteilnahme ausschließt (nach Beurteilung des Prüfarztes).

		27) ALT ≥ 5×ULN oder ALT ≥ 5×ULN und Bilirubin ≥ 1,5×ULN. 28) QTc Intervall (Bazett) > 450 msek oder QTc (Bazett) > 480 msek für Patienten mit Schenkelblock.
4b	Studienorganisation und Ort der Studiendurchführung	66 Zentren in 13 Ländern weltweit: Argentinien, Australien, Belgien, Kanada, Frankreich, Deutschland, Italien, Niederlande, Russland, Spanien, Taiwan, UK und USA.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	DTG/RPV-Arm: • DTG: 1 × 50 mg Tablette einmal täglich, oral einzunehmen • RPV: 1 × 25 mg Tablette einmal täglich, oral einzunehmen mit einer Mahlzeit CAR-Arm: • 2 NRTI + INI oder • 2 NRTI + NNRTI oder • 2 NRTI + PI gemäß den jeweiligen Fachinformationen einzunehmen.
6	Zielkriterien	

6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	primärer Endpunkt: • Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 48 sekundäre Endpunkte: • Änderung der CD4+ Zellzahl zu Woche 24 und zu Woche 48 jeweils gegenüber Baseline • Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 24 • Anteil der Patienten zu Woche 48, die nach dem Snapshot-Algorithmus als virologische Versager klassifiziert wurden Für Patienten, die bei Baseline in den DTG/RPV-Arm randomisiert wurden: • Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 100 und zu Woche 148 • Änderung der CD4+ Zellzahl zu Woche 100 und zu Woche 148 jeweils gegenüber Baseline • Inzidenz der Krankheitsprogression (HIV-assoziierte Erkrankungen, AIDS und Tod) bis Woche 148 Für Patienten, die nach der randomisierten Phase zu DTG/RPV wechselten: • Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 100 und zu Woche 148 • Änderung der CD4+ Zellzahl zu Woche 100 und zu Woche 148 jeweils gegenüber Baseline • Inzidenz der Krankheitsprogression (HIV-assoziierte Erkrankungen, AIDS und Tod) bis Woche 148 • Sicherheit/Verträglichkeit • Resistenzen • Pharmakokinetik • *Patient Reported Outcomes* (SDM, HIV TSQ, EQ-5D-5L, *Willingness to Switch*)
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Der Endpunkt des Anteils der Patienten zu Woche 48, die nach dem Snapshot-Algorithmus als virologische Versager klassifiziert wurden, wurde nachträglich hinzugefügt.

7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Bei einem tatsächlichen Anteil von 87% für den primären Endpunkt in beiden Studienarmen, einer Nichtunterlegenheitsschwelle von -10% und einem einseitigen Signifikanzniveau von 2,5% sind 238 Patienten pro Studienarm nötig, um 90% Power zu erreichen. Die Studienrekrutierung wurde etwas höher als die berechneten 476 Patienten angesetzt, um die nötige Zahl an Teilnehmern an der DEXA-Substudie zu ermöglichen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es war eine Zwischenanalyse zur Beurteilung einer etwaigen Zwecklosigkeit (*futility*) der Studie geplant. Diese sollte ca. 9 Monate nach Einschluss des ersten Patienten stattfindne, wenn ca. 50% der Patienten 24 Wochen eingeschlossen waren. Die Regel für einen daraus resultierenden Studienabbruch folgte bayesianischer Methodik. Zu Woche 36 wurde zudem von einem unabhängigen Kommittee die Inzidenz von Studienabbrüchen aufgrund virologischen Versagens untersucht, um individuell zu untersuchen, ob die Behandlung der Patienten ggf. nicht ausreichend optimiert ist.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Zentrale Randomisierung mittels der eigenen validierten Software von GlaxoSmithKline (RANDALL).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung im Verhältnis 1:1 stratifiziert nach Baseline *Third Agent* (INI, NNRTI, PI), Alter (< 50 Jahre, ≥ 50 Jahre) und geplanter Teilnahme an der DEXA-Substudie.
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	Zentrale Randomisierung mittels der eigenen validierten Software von GlaxoSmithKline (RANDALL). Weitergabe und Zuteilung der Patienten zu den Studienarmen erfolgte zentral durch ein interaktives Responsesystem (RAMOS NG). Die Geheimhaltung war gewährleistet.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Rekrutierung der Patienten lokal in den Studienzentren; Randomisierung/Zuteilung zentral.
11	Verblindung	

11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Es handelt sich um eine offene Studie, d.h. weder die Patienten noch die Behandler waren verblindet. Das Studienteam des Sponsors hatte bis zu Woche 48 keinerlei Kenntnis über die Randomisierung, um eine etwaige Verzerrung zu minimieren.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht relevant
12	Statistische Methoden	

12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	• Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 24 bzw. Woche 48: unadjustierte und adjustierte Risikodifferenzen (Cochran-Mantel Haenszel); Nichtunterlegenheit gilt als erreicht, falls die untere Grenze des 95%-Konfidenzintervalls > -10% ist (nur zu Woche 48). • Änderung der CD4+ Zellzahl zu Woche 24 und zu Woche 48 jeweils gegenüber Baseline: Übersichtstabellen mit den gängigen statistischen Kennzahlen stetiger Variablen für die Änderungen. • Anteil der Patienten zu Woche 48, die nach dem Snapshot-Algorithmus als virologische Versager klassifiziert wurden: unadjustierte und adjustierte Risikodifferenzen (Cochran-Mantel Haenszel). • Inzidenz der Krankheitsprogression (HIV-assoziierte Erkrankungen, AIDS und Tod): Deskriptive Darstellungen der Häufigkeiten; Listing. Analysen für die Langzeitbeobachtungen zu Woche 100 und 148 sind analog zu den Analysen zu Woche 48 geplant. • *Patient Reported Outcomes*: Deskriptive Darstellungen sowie Kovarianzanalysen für die Änderungen der jeweiligen Scores gegenüber Baseline. • Virologie: Inzidenz genotypischer und phänotypischer Resistenzen: deskriptive Darstellung der Häufigkeiten; Listing. • Verträglichkeit/Sicherheit: Deskriptive Darstellungen der Häufigkeiten nach MedDRA kodiert.

12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	A priori waren Analysen des primären Endpunkts für folgende Subgruppen geplant: • Alter • Rasse • Geschlecht • Land • CD4+ Zellzahl bei Baseline • CDC Klassifikation • Baseline *Third Agent*
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	siehe auch Flussdiagramm im Anschluss an diese Tabelle DTG+RPV / CAR a) n=252 / n=256 b) ITT-E/Safety: n=252 / ITT-E/Safety: n=256 c) ITT-E/Safety n=252 Per-Protocol n=226 / ITT-E/Safety n=256 Per-Protocol n=226
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	siehe Flussdiagramm im Anschluss an diese Tabelle
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studiendauer: Beginn: 14.04.2015 Letzter Patientenbesuch für Woche 48: 16.09.2016 Datenschluss für Woche 48: 22.11.2016 Studie läuft noch.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie läuft noch planmäßig.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

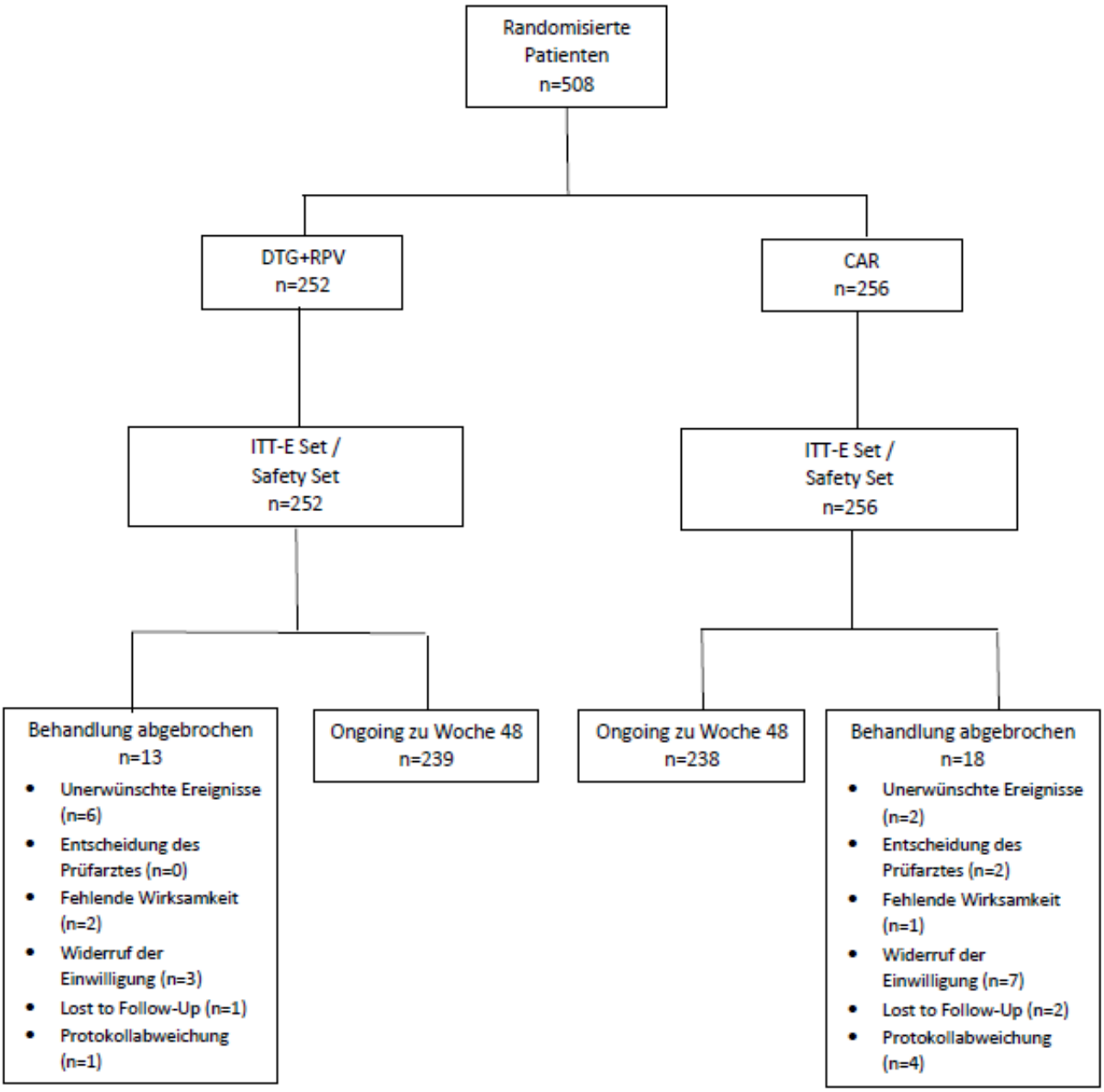


Abbildung 37: Flowchart für die Studie 201636 (SWORD-1)

Tabelle 4-161 (Anhang): Studiendesign und -methodik für Studie 201637 (SWORD-2)

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primäres Ziel: Zeigen der Nichtunterlegenheit hinsichtlich der antiretroviralen Aktivität bei Wechsel zu DTG + RPV einmal täglich im Vergleich zum Fortführen der aktuellen antiretroviralen Therapie über 48 Wochen bei HIV-1 infizierten antiretroviral vorbehandelten Patienten.
Methoden		
3	Studiendesign	

3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multizentrische, randomisierte, offene, aktiv-kontrollierte Nichtunterlegenheitsstudie der Phase III; Randomisierung im Verhältnis 1:1 stratifiziert nach Baseline *Third Agent* (INI, NNRTI, PI), Alter (< 50 Jahre, ≥ 50 Jahre) und geplanter Teilnahme an der DEXA-Substudie. Der primäre Endpunkt wurde nach der randomisierten Phase (*Early Switch Phase*) nach 48 Wochen erhoben. Anschließend wurden auch die Patienten im Kontrollarm auf DTG+RPV umgestellt, wenn sie eine Viruslast < 50c/ml hatten. Diese *Late Switch Phase* dauert bis zu Woche 148.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	**Amendment 1** (26.02.2015) [vor Einschluss des ersten Patienten]: Hinzufügen zusätzlicher pharmakokinetischer Studienbesuche; Hinzufügen des Stratifikationsfaktors aufgrund der geplanten Teilnahme an der DEXA-Substudie; kleinere Änderungen der Ein- und Ausschlusskriterien; Änderung der Beschreibung des Monitorings des unabhängigen Kommittees bzgl. der Kriterien für einen Studienabbruch aufgrund von virologischen Gründen; Anpassungen bzgl. nicht zugelassener Begleitmedikation; Anpassung der Definition des Studienendes; Anpassungen des suizidalen Risikomonitorings; Überarbeitung der Tabellen bzgl. Zeiten und Ereignisse. **Amendment 2** (08.06.2015): Angleichen der Gründe für einen Wechsel der PI-Klasse an andere retrovirale Klassen; Verdeutlichung des Ausschlusskriteriums 10 bzgl. einer Hepatitis B Infektion; Überarbeitung der stratifizierten Auswertemethodik des primären Endpunkts; Überarbeitung der Kriterien für einen vorzeitigen Studienabbruch aufgrund von virologischen Gründen; Überarbeitung der Strategie zur Risikoverringerung bzgl. Hautausschlag vom Grad 3 oder 4 und dadurch bedingtem Studienabbruch.
4	Probanden / Patienten	

4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**Einschlusskriterien:** 1) HIV-1 infizierte Männer oder Frauen ≥ 18 Jahre. 2) Aktuelle antiretrovirale Kombinationstherapie, die für mind. 6 Monate stabil war (erste oder zweite ART des Patienten); Frühere Wechsel der ART sollten aufgrund von Verträglichkeit/Sicherheit oder Zugang zu Medikamenten oder Vereinfachung der Therapie geschehen sein. Folgende Kombinationstherapien vor dem Screeningzeitpunkt waren zulässig: 2 NRTI plus: INI oder NNRTI (jeweils als erste oder zweite ART) bzw. geboosterter PI (oder Atazanavir ohne Booster) als erste ART. 3) Mind. zwei dokumentierte Messungen der HIV-1 Viruslast < 50 c/ml im Plasma innerhalb der letzten 12 Monate vor Screening; eine innerhalb Monat 6 bis 12 und die andere innerhalb der 6 Monate vor Screening. 4) HIV-1 Viruslast < 50 c/ml zum Screeningzeitpunkt. 5) Frauen im gebärfähigen Alter müssen jeweils einen negativen Schwangerschaftstest zum Screening und an Tag 1 aufweisen und den protokollgemäßen Kontrazeptionsvorgaben zustimmen. 6) Vorhandensein von Wille und Fähigkeit, die Anforderungen der Studienteilnahme zu verstehen und Vorliegen einer unterschriebenen und datierten Einverständniserklärung vor Screening. 7) Patienten in Frankreich mussten entweder Angehöriger oder Leistungsempfänger einer Sozialversicherung sein. **Ausschlusskriterien:** 1) Jedwede Messung der HIV-1 Viruslast ≥ 50 c/ml innerhalb 6 Monate vor Screening und nach bestätigter Suppression auf < 50 c/ml unter der aktuellen ART. 2) Jedwede Messung der HIV-1 Viruslast > 200 c/ml innerhalb 6 bis 12 Monaten vor Screening und nach bestätigter Suppression auf < 50 c/ml unter der aktuellen ART. 3) Zwei oder mehr Messungen der HIV-1 Viruslast ≥ 50 c/ml innerhalb 6 bis 12

		Monaten vor Screening und nach bestätigter Suppression auf < 50 c/ml unter der aktuellen ART. 4) Vorliegen einer Pause von mehr als einem Monat, in der alle antiretroviralen Medikamente abgesetzt wurden (wegen Verträglichkeit/Sicherheit) im Zeitraum bis 6 Monate vor Screening. 5) Jedweder Wechsel zu einem Zweitlinienregime aufgrund virologischem Therapieversagen (Viruslast ≥ 400 c/ml nach anfänglicher Suppression unter 50 c/ml während der Erstlinientherapie). 6) Schwangere oder stillende Frauen, oder solche, die eine Schwangerschaft während der Studienlaufzeit planen. 7) Vorliegen von Anzeichen einer aktiven CDC Kategorie C Krankheit. Ausnahmen hierbei sind: kutanes Kaposi's Sarkom ohne notwendiger systemischer Therapie und einer früheren CD4+ Zellzahl von < 200 pro mm^3. 8) Patienten mit jedweder hepatischer Beeinträchtigung. 9) Patienten mit positivem Hepatitis B Test oder mit einer absehbaren Therapie gegen Hepatitis C während der Studienlaufzeit. 10) Zurückliegende oder aktuelle Allergie gegen die Studienmedikation oder eine derer Komponenten in der jeweiligen Klasse. 11) Bestehende Malignität außer kutanes Kaposi's Sarkom, basales Zellkarzinom, entferntes nicht-invasives kutanes Plattenepithelkarzinom, intraepitheliale Neoplasie des Gebärmutterhalses; andere Malignitäten mussten durch den Prüfarzt und den Monitor beurteilt werden. 12) Patienten mit erhöhtem Suizidrisiko (nach Einschätzung durch den Prüfarzt). 13) Jedweder existierender körperlicher oder mentaler Zustand, der sich auf die Teilnahme an der Studie (Protokollvorgaben, Compliance) negativ auswirkt bzw. die Sicherheit des Patienten beeinträchtigt (nach Beurteilung des Prüfarztes). 14) Jedweder Zustand, der die Absorption, Distribution, Metabolisierung oder das Ausscheiden der Studienmedikation beeinflusst oder der es für den Patienten

		unmöglich macht, die orale Medikation einzunehmen (nach Beurteilung des Prüfarztes). 15) Einnahme von Medikation gegen Torsades de Pointes. 16) Vakzinierung mit einer HIV-1 Immuntherapie innerhalb von 90 Tagen vor Screening. 17) Behandlung mit einer der folgenden Therapien innerhalb von 28 Tagen vor Screening: Strahlentherapie, zytotoxische Chemotherapeutika, jedwede Immunmodulatoren, die die Immunantwort beeinflussen. 18) Exposition einer experimentellen Medikation oder eines Impfstoffs innerhalb des längsten Zeitintervalls vor Studienbeginn, dessen Länge definiert ist: entweder 28 Tage oder die fünffache Halbwertszeit der Testsubstanz oder die zweifache Dauer des biologischen Effekts. 19) Teilnahme an einer anderen interventionellen Studie nach Randomisierung mit Ausnahme der DEXA-Substudie. 20) Frühere Behandlung mit einem Therapieregime, das entweder aus nur einem NNRTI (auch wenn nur pari-partum) oder aus nur einem oder zwei NRTI bestand. 21) Frühere oder aktuelle Behandlung mit Etravirin. 22) Frühere Behandlung mit Tipranavir/Ritonavir oder Fosamprenavir/Ritonavir. 23) Patienten, die eine laut Studienprotokoll nicht zulässige Medikation erhalten und diese nicht ersetzten können oder wollen. 24) Nachweis des Vorliegens von resistenz-assoziierten Mutationen gegenüber PI, INI, NRTI oder NNRTI oder Substitution R263K gegenüber INI. 25) Jedwede Laborabweichung von Grad 4 (außer Grad 4 Lipidwerte). Eine einfache Testwiederholung während der Screeningphase war zulässig als Nachweis. 26) Jedwede Laborabweichung zum Screeningzeitpunkt, die eine Studienteilnahme ausschließt (nach Beurteilung des Prüfarztes).

		27) ALT ≥ 5×ULN oder ALT ≥ 5×ULN und Bilirubin ≥ 1,5×ULN. 28) QTc Intervall (Bazett) > 450 msek oder QTc (Bazett) > 480 msek für Patienten mit Schenkelblock.
4b	Studienorganisation und Ort der Studiendurchführung	60 Zentren in 11 Ländern weltweit: Spanien (15), USA (13), Kanada (7), Frankreich (6), Deutschland (5), Russland (4), Taiwan (4), Argentinien (3), Australien (3), Italien (2), UK (1).
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	DTG/RPV-Arm: • DTG: 1 × 50 mg Tablette einmal täglich, oral einzunehmen • RPV: 1 × 25 mg Tablette einmal täglich, oral einzunehmen mit einer Mahlzeit CAR-Arm: • 2 NRTI + INI oder • 2 NRTI + NNRTI oder • 2 NRTI + PI gemäß den jeweiligen Fachinformationen einzunehmen.
6	Zielkriterien	

6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	primärer Endpunkt: • Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 48 sekundäre Endpunkte: • Änderung der CD4+ Zellzahl zu Woche 24 und zu Woche 48 jeweils gegenüber Baseline • Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 24 • Anteil der Patienten zu Woche 48, die nach dem Snapshot-Algorithmus als virologische Versager klassifiziert wurden Für Patienten, die bei Baseline in den DTG/RPV-Arm randomisiert wurden: • Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 100 und zu Woche 148 • Änderung der CD4+ Zellzahl zu Woche 100 und zu Woche 148 jeweils gegenüber Baseline • Inzidenz der Krankheitsprogression (HIV-assoziierte Erkrankungen, AIDS und Tod) bis Woche 148 Für Patienten, die nach der randomisierten Phase zu DTG/RPV wechselten: • Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 100 und zu Woche 148 • Änderung der CD4+ Zellzahl zu Woche 100 und zu Woche 148 jeweils gegenüber Baseline • Inzidenz der Krankheitsprogression (HIV-assoziierte Erkrankungen, AIDS und Tod) bis Woche 148 • Sicherheit/Verträglichkeit • Resistenzen • Pharmakokinetik • *Patient Reported Outcomes* (SDM, HIV TSQ, EQ-5D-5L, *Willingness to Switch*)
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Der Endpunkt des Anteils der Patienten zu Woche 48, die nach dem Snapshot-Algorithmus als virologische Versager klassifiziert wurden, wurde nachträglich hinzugefügt.

7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Bei einem tatsächlichen Anteil von 87% für den primären Endpunkt in beiden Studienarmen, einer Nichtunterlegenheitsschwelle von -10% und einem einseitigen Signifikanzniveau von 2,5% sind 238 Patienten pro Studienarm nötig, um 90% Power zu erreichen. Die Studienrekrutierung wurde etwas höher als die berechneten 476 Patienten angesetzt, um die nötige Zahl an Teilnehmern an der DEXA-Substudie zu ermöglichen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es war eine Zwischenanalyse zur Beurteilung einer etwaigen Zwecklosigkeit (*futility*) der Studie geplant. Diese sollte ca. 9 Monate nach Einschluss des ersten Patienten stattfindne, wenn ca. 50% der Patienten 24 Wochen eingeschlossen waren. Die Regel für einen daraus resultierenden Studienabbruch folgte bayesianischer Methodik. Zu Woche 36 wurde zudem von einem unabhängigen Kommittee die Inzidenz von Studienabbrüchen aufgrund virologischen Versagens untersucht, um individuell zu untersuchen, ob die Behandlung der Patienten ggf. nicht ausreichend optimiert ist.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Zentrale Randomisierung mittels der eigenen validierten Software von GlaxoSmithKline (RANDALL).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung im Verhältnis 1:1 stratifiziert nach Baseline *Third Agent* (INI, NNRTI, PI), Alter (< 50 Jahre, ≥ 50 Jahre) und geplanter Teilnahme an der DEXA-Substudie.
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	Zentrale Randomisierung mittels der eigenen validierten Software von GlaxoSmithKline (RANDALL). Weitergabe und Zuteilung der Patienten zu den Studienarmen erfolgte zentral durch ein interaktives Responsesystem (RAMOS NG). Die Geheimhaltung war gewährleistet.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Rekrutierung der Patienten lokal in den Studienzentren; Randomisierung/Zuteilung zentral.
11	Verblindung	

11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Es handelt sich um eine offene Studie, d.h. weder die Patienten noch die Behandler waren verblindet. Das Studienteam des Sponsors hatte bis zu Woche 48 keinerlei Kenntnis über die Randomisierung, um eine etwaige Verzerrung zu minimieren.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht relevant
12	Statistische Methoden	

12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	• Anteil der Patienten mit einer Viruslast < 50 c/ml zu Woche 24 bzw. Woche 48: unadjustierte und adjustierte Risikodifferenzen (Cochran-Mantel Haenszel); Nichtunterlegenheit gilt als erreicht, falls die untere Grenze des 95%-Konfidenzintervalls > -10% ist (nur zu Woche 48). • Änderung der CD4+ Zellzahl zu Woche 24 und zu Woche 48 jeweils gegenüber Baseline: Übersichtstabellen mit den gängigen statistischen Kennzahlen stetiger Variablen für die Änderungen. • Anteil der Patienten zu Woche 48, die nach dem Snapshot-Algorithmus als virologische Versager klassifiziert wurden: unadjustierte und adjustierte Risikodifferenzen (Cochran-Mantel Haenszel). • Inzidenz der Krankheitsprogression (HIV-assoziierte Erkrankungen, AIDS und Tod): Deskriptive Darstellungen der Häufigkeiten; Listing. Analysen für die Langzeitbeobachtungen zu Woche 100 und 148 sind analog zu den Analysen zu Woche 48 geplant. • *Patient Reported Outcomes*: Deskriptive Darstellungen sowie Kovarianzanalysen für die Änderungen der jeweiligen Scores gegenüber Baseline. • Virologie: Inzidenz genotypischer und phänotypischer Resistenzen: deskriptive Darstellung der Häufigkeiten; Listing. • Verträglichkeit/Sicherheit: Deskriptive Darstellungen der Häufigkeiten nach MedDRA kodiert.

12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	A priori waren Analysen des primären Endpunkts für folgende Subgruppen geplant: • Alter • Rasse • Geschlecht • Land • CD4+ Zellzahl bei Baseline • CDC Klassifikation • Baseline *Third Agent*
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	siehe auch Flussdiagramm im Anschluss an diese Tabelle [table 1]
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	siehe Flussdiagramm im Anschluss an diese Tabelle
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studiendauer: Beginn: 21.04.2015 Letzter Patientenbesuch für Woche 48: 16.09.2016 Datenschluss für Woche 48: 22.11.2016 Studie läuft noch.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie läuft noch planmäßig.
a: nach CONSORT 2010.		

	DTG+RPV	CAR
a)	n=261	n=255
b)	ITT-E/Safety: n=261	ITT-E/Safety: n=255
c)	ITT-E/Safety n=261 Per-Protocol n=231	ITT-E/Safety n=255 Per-Protocol n=227

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

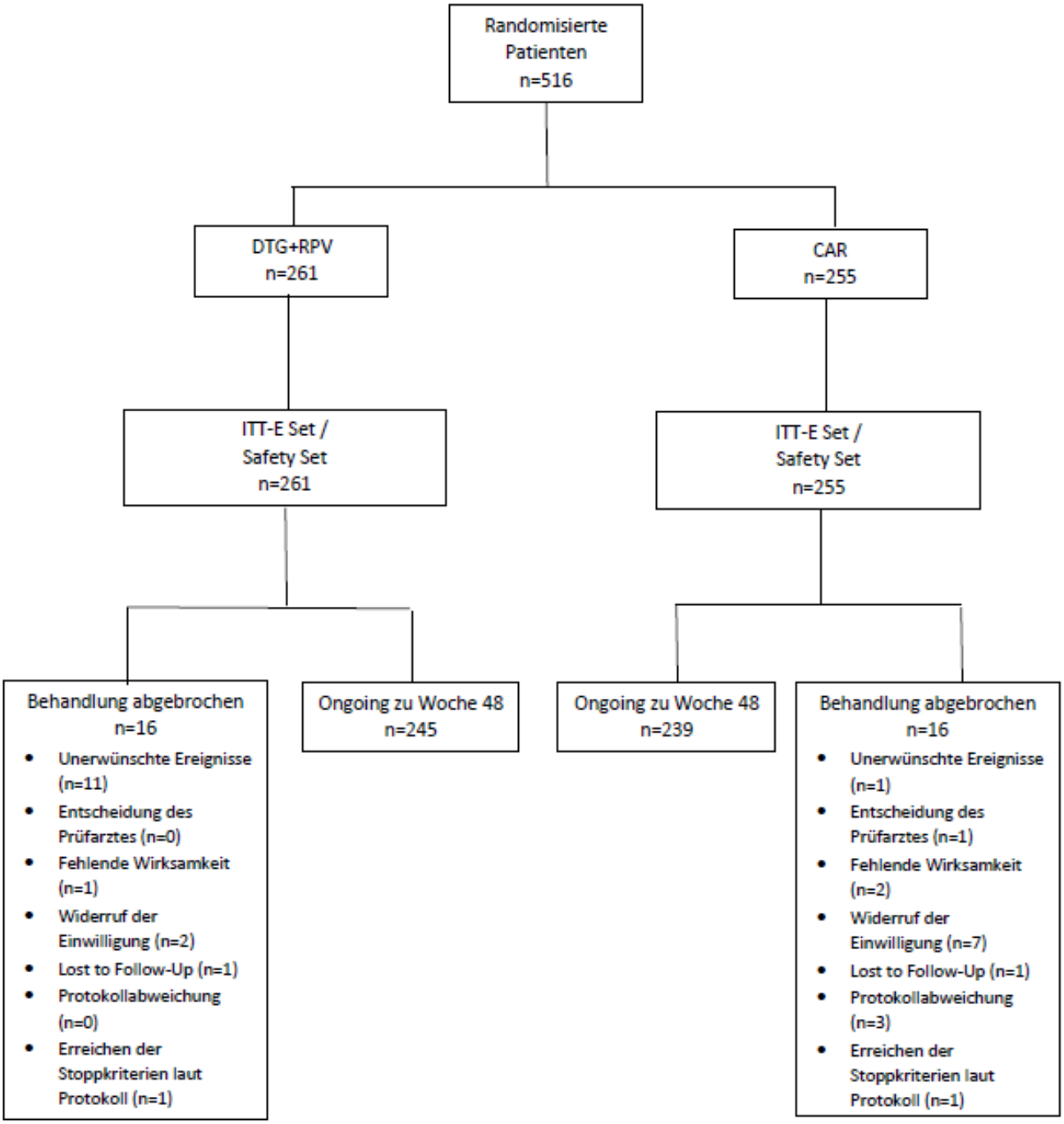


Abbildung 38: Flowchart für die Studie 201637 (SWORD-2)

Tabelle 4-162 (Anhang): Studiendesign und -methodik für Studie 202094 (DEXA-Substudie)

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primäres Ziel: Untersuchung der Änderung der Knochenmineraldichte bei Wechsel von einer TDF-haltigen ART zu DTG + RPV einmal täglich im Vergleich zum Fortführen der aktuellen ART über 48 Wochen bei HIV-1 infizierten antiretroviral vorbehandelten Patienten.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Offene, multizentrische, aktiv-kontrollierte Substudie der beiden weiter oben beschriebenen Zulassungsstudien SWORD-1 und SWORD-2.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	**Amendment 1** (17.03.2016): Begründung für das zusätzliche Erfassen des Geburtsmonats; geringere administrative Anpassungen.
4	Probanden / Patienten	

4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**Einschlusskriterien:** 1) Gescreente und geeignete Patienten der Studien SWORD-1 und SWORD-2, die noch nicht randomisiert waren. 2) Aktuelle antiretrovirale Kombinationstherapie, welche TDF als Bestandteil hatte. 3) Frauen im gebärfähigen Alter müssen jeweils einen negativen Schwangerschaftstest zum Screening und an Tag 1 aufweisen und den protokollgemäßen Kontrazeptionsvorgaben zustimmen. 4) Vorhandensein von Wille und Fähigkeit, die Anforderungen der Studienteilnahme zu verstehen und Vorliegen einer unterschriebenen und datierten Einverständniserklärung vor Screening. 5) Patienten in Frankreich mussten entweder Angehöriger oder Leistungsempfänger einer Sozialversicherung sein. **Ausschlusskriterien:** 1) Vorliegen von weniger als drei Wirbeln im Bereich L1–L4, die für eine DEXA-Messung geeignet sind. 2) Bilaterale künstliche Hüfte 3) Unkontrollierte Erkrankung der Schilddrüse 4) Hypogonadismus bei Männern 5) Stoffwechselerkankung (inkl. Cushing Syndrom, Diabetes Mellitus) 6) Fragilitätsfrakturen in der Vorgeschichte 7) Frühere Therapie gegen Osteoporose mit Bisphosphonaten, Östrogenrezeptor-Modulatoren oder anderen Therapien und/oder Vorliegen einer schweren Osteoporose. 8) BMI < 18 kg/m^2 oder ≥ 40 kg/m^2 9) Vitamin D Mangel: 25 Hydroxy Vitamin D < 15 ng/mL 10) Jedweder existierender körperlicher oder mentaler Zustand, der sich auf die Teilnahme an der Studie (Protokollvorgaben, Compliance) negativ auswirkt bzw. die Sicherheit des Patienten beeinträchtigt (nach Beurteilung des Prüfarztes). 11) Aktuelle oder geplante Behandlung mit Tamoxifen, knochenbezogene Therapien (z.B. Bisphosphonate, Osteoporose-Medikamente),

Item[a]	Charakteristikum	Studieninformation
		Wachstumshormone, anabolische Steroide (mit Ausnahme der im Folgenden spezifizierten Testosterone) 12) Folgende Therapien sind davon ausgeschlossen, sofern sie mind. 6 Monate vor Tag 1 verabreicht wurden und während der Studiendauer weitergeführt werden: • Antikonvulsivum • Hormontherapie (inkl. Hormonersatztherapie bei Frauen oder Testosteron als Ersatz- oder Ergänzungstherapie) 13) Schwangere oder stillende Frauen, oder solche, die eine Schwangerschaft während der Studienlaufzeit planen 14) Teilnahme an einer anderen interventionellen Studie nach mit Ausnahme der SWORD-Studien
4b	Studienorganisation und Ort der Studiendurchführung	32 Zentren in 6 Ländern weltweit: Argentinien, Belgien, Kanada, Spanien, UK und USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Für diese Substudie wurde keine zusätzliche Medikation verabreicht als derjenigen, die im Rahmen der SWORD-Studien wie folgt gegeben wurde: DTG/RPV-Arm: • DTG: 1 × 50 mg Tablette einmal täglich, oral einzunehmen • RPV: 1 × 25 mg Tablette einmal täglich, oral einzunehmen mit einer Mahlzeit CAR-Arm: • 2 NRTI inkl. TDF + INI oder PI oder NNRTI gemäß den jeweiligen Fachinformationen einzunehmen.
6	Zielkriterien	

Item[a]	Charakteristikum	Studieninformation
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	primärer Endpunkt: • Prozentuale Änderung der Knochenmineraldichte der Hüfte zu Woche 48 sekundäre Endpunkte: • Prozentuale Änderung der Knochenmineraldichte der Lendenwirbelsäule zu Woche 48 • Prozentuale Änderung der Knochenmineraldichte der Hüfte und der Lendenwirbelsäule zu Woche 148 im DTG+RPV Arm • Prozentuale Änderung der Knochenmineraldichte der Hüfte und der Lendenwirbelsäule zu Woche 148 (96 Wochen nach Switch) im CAR Arm • Änderung der T- und Z-Scores der Hüfte und der Lendenwirbelsäule zu Woche 48 • Änderung der T- und Z-Scores der Hüfte und der Lendenwirbelsäule zu Woche 148 im DTG+RPV Arm • Änderung der T- und Z-Scores der Hüfte und der Lendenwirbelsäule zu Woche 148 (96 Wochen nach Switch) im CAR Arm
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	keine
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die ursprüngliche Zielsetzung war der Einschluss von 150 Patienten. Bei Annahme einer SD von 3,5% wäre der kleinste beobachtete Effekt, der statistische Signifikanz (p<0,05) erreicht, 0,6%. Basierend auf einem tatsächlichen Effekt von 1,9% (SD=3,5%) führt das zu einer Power von 90%. Mit einer Stichprobe von 100 Patienten läge die Power bei 77%.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es waren keine Zwischenanalysen explizit geplant. Falls es der Zulassungsprozess erforderte oder Publikationen angestrebt wurden, konnten Zwischenauswertungen durchgeführt werden. Es wurden keine Analysen vor der primären 48-Wochen Analyse durchgeführt.
8	Randomisierung, Erzeugung der Behandlungsfolge	

Item[a]	Charakteristikum	Studieninformation
8a	Methode zur Generierung der zufälligen Zuteilung	Zentrale Randomisierung mittels der eigenen validierten Software von GlaxoSmithKline (RANDALL) im Rahmen der SWORD-Studien.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung im Rahmen der SWORD-Studien im Verhältnis 1:1 stratifiziert nach Baseline *Third Agent* (INI, NNRTI, PI), Alter (< 50 Jahre, ≥ 50 Jahre) und geplanter Teilnahme an der DEXA-Substudie.
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	Zentrale Randomisierung im Rahmen der SWORD-Studien mittels der eigenen validierten Software von GlaxoSmithKline (RANDALL). Weitergabe und Zuteilung der Patienten zu den Studienarmen erfolgte zentral durch ein interaktives Responsesystem (RAMOS NG). Die Geheimhaltung war gewährleistet.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Rekrutierung der Patienten lokal in den Studienzentren; Randomisierung/Zuteilung zentral.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Es handelt sich um eine offene Studie, d.h. weder die Patienten noch die Behandler waren verblindet. Das Studienteam des Sponsors hatte bis zu Woche 48 keinerlei Kenntnis über die Randomisierung, um eine etwaige Verzerrung zu minimieren.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	nicht relevant
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Deskriptive Darstellungen sowie Kovarianzanalysen für die Änderungen der jeweiligen Parameter zu Woche 48 gegenüber Baseline (adjustiert für Baseline Knochenmineraldichte, Alter und Baseline BMI)
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	A priori waren Analysen für folgende Subgruppen geplant: • Baseline *Third Agent*
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	

<table>
<tr><th>Item[a]</th><th>Charakteristikum</th><th>Studieninformation</th></tr>
<tr><td>13a</td><td>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</td><td>siehe auch Flussdiagramm im Anschluss an diese Tabelle
<table>
<tr><th></th><th>DTG+RPV</th><th>CAR</th></tr>
<tr><td>a)</td><td>n=53</td><td>n=49</td></tr>
<tr><td>b)</td><td>n=53</td><td>n=49</td></tr>
<tr><td>c)</td><td>n=53</td><td>n=49</td></tr>
</table></td></tr>
<tr><td>13b</td><td>Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen</td><td>siehe Flussdiagramm im Anschluss an diese Tabelle</td></tr>
<tr><td>14</td><td>Aufnahme / Rekrutierung</td><td></td></tr>
<tr><td>14a</td><td>Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung</td><td>Studiendauer:
Beginn: 12.06.2015
Letzter Patientenbesuch für Woche 48: 16.09.2016
Datenschluss für Woche 48: 22.11.2016
Studie läuft noch.</td></tr>
<tr><td>14b</td><td>Informationen, warum die Studie endete oder beendet wurde</td><td>Die Studie läuft noch planmäßig.</td></tr>
<tr><td colspan="3">a: nach CONSORT 2010.</td></tr>
</table>

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

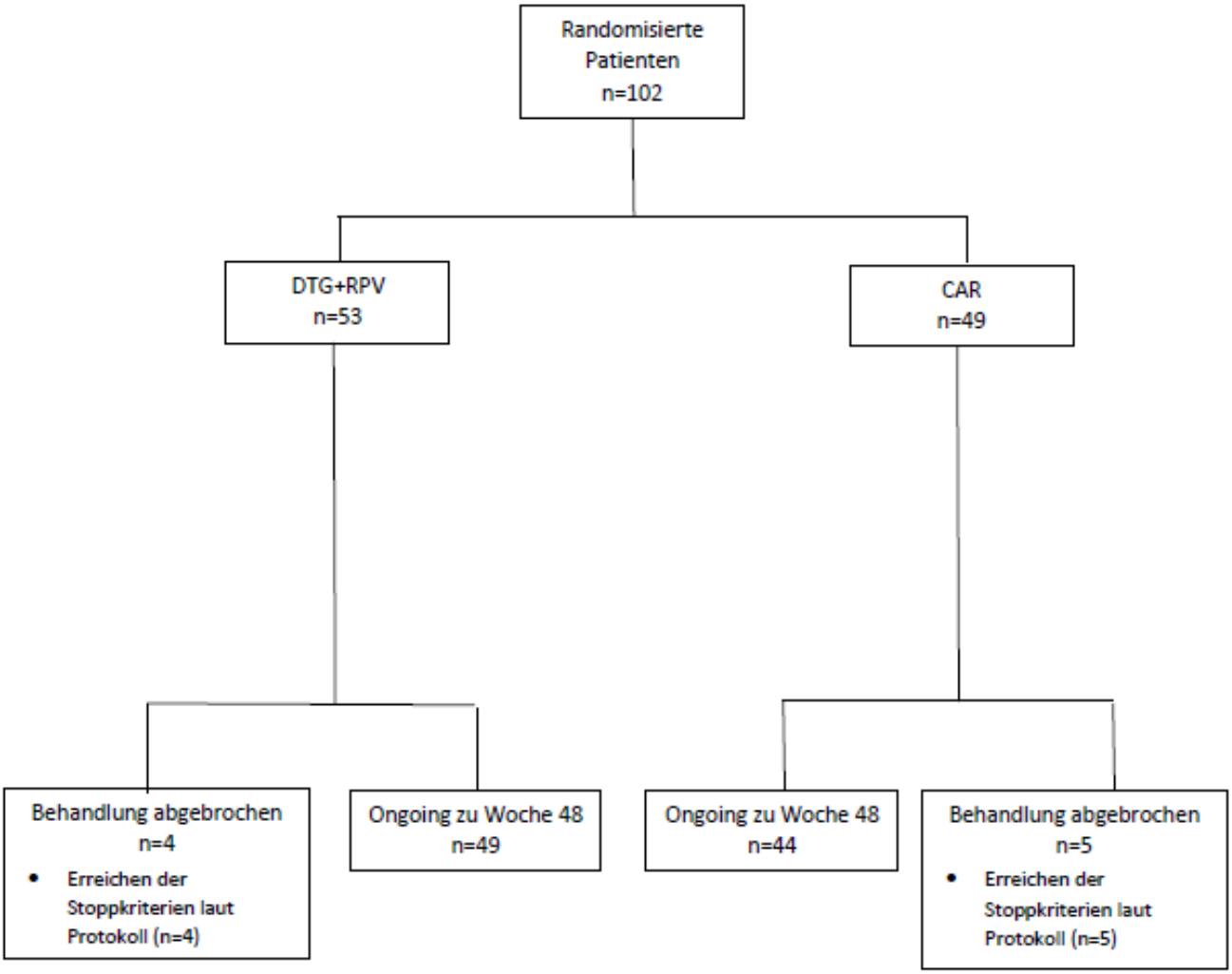


Abbildung 39: Flowchart für die Studie 202094 (DEXA-Substudie)

Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Der nachfolgend dargestellte 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:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.

- Verzerrungsaspekte auf Endpunktebene. 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 sollen die Bögen studienbezogen ausgefüllt werden. 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.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

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 4-163 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SWORD-1

Studie: SWORD-1 (201636)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht SWORD-1	A

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☒ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein** → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.
<u>für randomisierte Studien:</u> Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

2.
<u>für randomisierte Studien:</u> Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie war nicht verblindet, sowohl der Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation.

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie war nicht verblindet, sowohl der Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Bei der Studie handelt es sich um eine randomisierte, kontrollierte Studie. Die Zuordnung der Patienten zu den Behandlungsgruppen erfolgte jeweils mittels eines zuvor generierten zufälligen Randomisierungscodes. Die Studie war nicht verblindet, sowohl der behandelnde Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation. Des Weiteren existieren keine

Hinweise auf eine ergebnisabhängige Berichterstattung sowie keine sonstigen Aspekte, die für ein zusätzliches hohes Verzerrungspotenzial ursächlich wären.

Die offene Studiendurchführung allein führt nicht dazu, dass die Studie generell als mit hohem Verzerrungspotenzial behaftet eingestuft werden sollte. Sofern sich dies auf Endpunktebene anders verhält, wird dies individuell bei den entsprechenden Endpunkten diskutiert.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt:

- **Viruslast < 50 Kopien/ml zu Woche 48**
- **Knochenspezifischer Biomarker: BSAP zu Woche 48**
- **Knochenspezifischer Biomarker: P1NP zu Woche 48**
- **Knochenspezifischer Biomarker: CTx zu Woche 48**
- **Knochenspezifischer Biomarker: Osteocalcin zu Woche 48**

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Endpunkt wurde verblindet in einem zentralen Labor analysiert.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Diese Endpunkte wurden verblindet in einem zentralen Labor analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten.

Endpunkt:

- **EQ-5D zu Woche 48**
- **Symptom Distress Module zu Woche 48**
- **HIVTSQ Total Score zu Woche 48**

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Endpunkt wurde von den Patienten selbst erhoben, die unverblindet bzgl. der Behandlungsgruppe waren. Es kann davon ausgegangen werden, dass trotz des offenen Studiendesigns der Patient in seiner individuellen Bewertung nicht von den Bewertungen anderer Patienten oder insbesondere anderer Patienten der alternativen Behandlungsgruppe beeinflusst wurde.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Diese Endpunkte wurden von den Patienten selbst erhoben, die unverblindet bzgl. der Behandlungsgruppe waren. Es kann davon ausgegangen

werden, dass trotz des offenen Studiendesigns der Patient in seiner individuellen Bewertung nicht von den Bewertungen anderer Patienten oder insbesondere anderer Patienten der alternativen Behandlungsgruppe beeinflusst wurde. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten.

Endpunkt:

- **Mortalität bis Woche 48**

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Dieser Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Diese Tatsache hat jedoch bei diesem objektiv erhobenen Endpunkt keine Auswirkungen auf ein mögliches erhöhtes Verzerrungspotenzial.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Dieser Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Diese Tatsache hat jedoch bei diesem objektiv erhobenen Endpunkt keine Auswirkungen auf ein mögliches erhöhtes Verzerrungspotenzial.Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige

Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten.

Endpunkt:

- **Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48**
- **Schwerwiegende unerwünschte Ereignisse bis Woche 48**

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkte wurden vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Trotz der a priori bestehenden Ungleichheit zwischen beiden Behandlungsgruppen hinsichtlich der Beurteilung der Endpunkte bei dem offenem Studiendesign wird bei diesen Endpunkten nicht von einem erhöhten Verzerrungspotenzial ausgegangen, da hier eine im Vergleich zu weniger schweren unerwünschten Ereignissen noch genauere kausale Berücksichtigung der Umstände bei der Beurteilung durch den Prüfarzt sowie das Einbeziehen von Plausibilitätsargumenten vorausgesetzt werden kann.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Die Endpunkte wurden vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten. Trotz der bereits erläuterten a priori bestehenden Ungleichheit zwischen beiden Behandlungsgruppen hinsichtlich der Beurteilung der Endpunkte bei dem offenem Studiendesign wird bei diesen Endpunkten nicht von einem erhöhten Verzerrungspotenzial ausgegangen, da hier eine im Vergleich zu weniger schweren unerwünschten Ereignissen noch genauere kausale Berücksichtigung der Umstände bei der Beurteilung durch den Prüfarzt sowie das Einbeziehen von Plausibilitätsargumenten vorausgesetzt werden kann.

Endpunkt:

- **Unerwünschte Ereignisse bis Woche 48**
- **Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes**

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Diese Endpunkte wurden vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Diese Endpunkte wurden vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Es existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten.

Bei diesen Endpunkten werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt potentiell dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diese Endpunkte gesehen.

Tabelle 4-164 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SWORD-2

Studie: SWORD-2 (201637)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht SWORD-2	A

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☒ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein** → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.
<u>für randomisierte Studien:</u> Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

2.
<u>für randomisierte Studien:</u> Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie war nicht verblindet, sowohl der Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation.

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie war nicht verblindet, sowohl der Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Bei der Studie handelt es sich um eine randomisierte, kontrollierte Studie. Die Zuordnung der Patienten zu den Behandlungsgruppen erfolgte jeweils mittels eines zuvor generierten zufälligen Randomisierungscodes. Die Studie war nicht verblindet, sowohl der behandelnde Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation. Des Weiteren existieren keine

Hinweise auf eine ergebnisabhängige Berichterstattung sowie keine sonstigen Aspekte, die für ein zusätzliches hohes Verzerrungspotenzial ursächlich wären.

Die offene Studiendurchführung allein führt nicht dazu, dass die Studie generell als mit hohem Verzerrungspotenzial behaftet eingestuft werden sollte. Sofern sich dies auf Endpunktebene anders verhält, wird dies individuell bei den entsprechenden Endpunkten diskutiert.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt:

- **Viruslast < 50 Kopien/ml zu Woche 48**
- **Knochenspezifischer Biomarker: BSAP zu Woche 48**
- **Knochenspezifischer Biomarker: P1NP zu Woche 48**
- **Knochenspezifischer Biomarker: CTx zu Woche 48**
- **Knochenspezifischer Biomarker: Osteocalcin zu Woche 48**

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Endpunkt wurde verblindet in einem zentralen Labor analysiert.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Diese Endpunkte wurden verblindet in einem zentralen Labor analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten.

Endpunkt:

- **EQ-5D zu Woche 48**
- **Symptom Distress Module zu Woche 48**
- **HIVTSQ Total Score zu Woche 48**

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Endpunkt wurde von den Patienten selbst erhoben, die unverblindet bzgl. der Behandlungsgruppe waren. Es kann davon ausgegangen werden, dass trotz des offenen Studiendesigns der Patient in seiner individuellen Bewertung nicht von den Bewertungen anderer Patienten oder insbesondere anderer Patienten der alternativen Behandlungsgruppe beeinflusst wurde.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Diese Endpunkte wurden von den Patienten selbst erhoben, die unverblindet bzgl. der Behandlungsgruppe waren. Es kann davon ausgegangen

werden, dass trotz des offenen Studiendesigns der Patient in seiner individuellen Bewertung nicht von den Bewertungen anderer Patienten oder insbesondere anderer Patienten der alternativen Behandlungsgruppe beeinflusst wurde. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten.

Endpunkt:

- **Mortalität bis Woche 48**

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Dieser Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Diese Tatsache hat jedoch bei diesem objektiv erhobenen Endpunkt keine Auswirkungen auf ein mögliches erhöhtes Verzerrungspotenzial.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Dieser Endpunkt wurde vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Diese Tatsache hat jedoch bei diesem objektiv erhobenen Endpunkt keine Auswirkungen auf ein mögliches erhöhtes Verzerrungspotenzial.Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige

Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten.

Endpunkt:

- **Unerwünschte Ereignisse von Grad 3 oder 4 bis Woche 48**
- **Schwerwiegende unerwünschte Ereignisse bis Woche 48**

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkte wurden vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Trotz der a priori bestehenden Ungleichheit zwischen beiden Behandlungsgruppen hinsichtlich der Beurteilung der Endpunkte bei dem offenem Studiendesign wird bei diesen Endpunkten nicht von einem erhöhten Verzerrungspotenzial ausgegangen, da hier eine im Vergleich zu weniger schweren unerwünschten Ereignissen noch genauere kausale Berücksichtigung der Umstände bei der Beurteilung durch den Prüfarzt sowie das Einbeziehen von Plausibilitätsargumenten vorausgesetzt werden kann.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Die Endpunkte wurden vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten. Trotz der bereits erläuterten a priori bestehenden Ungleichheit zwischen beiden Behandlungsgruppen hinsichtlich der Beurteilung der Endpunkte bei dem offenem Studiendesign wird bei diesen Endpunkten nicht von einem erhöhten Verzerrungspotenzial ausgegangen, da hier eine im Vergleich zu weniger schweren unerwünschten Ereignissen noch genauere kausale Berücksichtigung der Umstände bei der Beurteilung durch den Prüfarzt sowie das Einbeziehen von Plausibilitätsargumenten vorausgesetzt werden kann.

Endpunkt:

- **Unerwünschte Ereignisse bis Woche 48**
- **Unerwünschte Ereignisse bis Woche 48, die zum Therapieabbruch geführt hatten**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Herzerkrankungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Gastrointestinaltrakts**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Leber- und Gallenerkrankungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Hypersensitivität**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Stoffwechsel- und Ernährungsstörungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen des Nervensystems**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Psychiatrische Erkrankungen**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Nieren und Harnwege**
- **Unerwünschte Ereignisse von besonderem Interesse bis Woche 48: Erkrankungen der Haut und des Unterhautzellgewebes**

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Diese Endpunkte wurden vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Studie besitzt auf Studienebene niedriges Verzerrungspotenzial. Diese Endpunkte wurden vom Prüfarzt dokumentiert, der unverblindet bzgl. der Behandlungsgruppe war. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es liegt keine ergebnisabhängige Berichterstattung vor. Es existieren keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diese Endpunkte hindeuten.

Bei diesen Endpunkten werden alle Ereignisse unabhängig von ihrem Schweregrad berichtet und gezählt. Bei dem Charakter der Studie (in einem Behandlungsarm bleibt die Behandlung unverändert, die der Patient bereits mindestens 6 Monate stabil erhalten hatte, im anderen Arm wird die ART komplett verändert) und der offenen Durchführung der Studie liegt es nahe, dass hier ein Reporting-Bias vorliegen könnte und aus den genannten Gründen unerwünschte Ereignisse potentiell häufiger im DTG/RPV-Arm berichtet werden. Auch der Aspekt der Patientensicherheit könnte den Prüfarzt potentiell dazu bewegen, bei Beobachtungen im DTG/RPV-Arm schneller und/oder häufiger ein unerwünschtes Ereignis zu registrieren als bei der bereits bewährten Therapie (CAR). Daher wird ein erhöhtes Verzerrungspotenzial für diese Endpunkte gesehen.

Tabelle 4-165 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie DEXA-Substudie

Studie: DEXA-Substudie (202094)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht DEXA-Substudie	A

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☒ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein** → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.
<u>für randomisierte Studien:</u> Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

2.
<u>für randomisierte Studien:</u> Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie war nicht verblindet, sowohl der Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation.

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie war nicht verblindet, sowohl der Arzt als auch der Patient hatten Kenntnis über die tatsächlich verabreichte Medikation.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Substudie besitzt basierend auf der Einstufung der beiden Hauptstudien SWORD-1 und SWORD-2 auf Studienebene ebenfalls niedriges Verzerrungspotenzial. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial hindeuten.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt:

- **Knochenmineraldichte zu Woche 48**

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Endpunkt wurde verblindet von einem zentralen Begutachter analysiert.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Substudie besitzt basierend auf der Einstufung der beiden Hauptstudien SWORD-1 und SWORD-2 auf Studienebene ebenfalls niedriges Verzerrungspotenzial. Der Endpunkt wurde verblindet von einem zentralen Begutachter analysiert. Das ITT-Prinzip wurde nach den anerkannten Standards umgesetzt, es

liegt keine ergebnisabhängige Berichterstattung vor. Des Weiteren existieren auch keine sonstigen Aspekte, die auf ein erhöhtes Verzerrungspotenzial für diesen Endpunkt hindeuten.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A Verzerrungsaspekte 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 Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.
→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

☐ **ja**: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

☐ **unklar**: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

☐ **nein**: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

☐ **ja**: Die Gruppen wurden zeitlich parallel verfolgt.

☐ **unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

☐ **nein**: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

☐ **ja**: Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

☐ **unklar**: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

☐ **nein**: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja**: Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen 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 behandelnden Personen

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:

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. 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 behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:
- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter
- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:
- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
  Zulässige Gründe sind:
  - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
  - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
  - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*

- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung 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 dies nicht zur Verzerrung anderer Endpunkte führt.

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen 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. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

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.

Begründung für die Einstufung:

_____

_____

B Verzerrungsaspekte auf Endpunktebene 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: _____

1. **Verblindung der Endpunkterheber**

   Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.
   In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

   ☐ **ja:** Der Endpunkt wurde verblindet erhoben.

   ☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

   ☐ **nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

   Angaben zum Kriterium; obligate Begründung für die Einstufung:

   _____

   _____

2. **Adäquate Umsetzung des ITT-Prinzips**

   Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set"). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

   ☐ **ja:** Eines der folgenden Merkmale trifft zu:
   - Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und

es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

☐ **unklar**: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

☐ **nein**: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

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.

Begründung für die Einstufung: