Dokumentvorlage, Version vom 16.12.2021

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Upadacitinib (RINVOQ®)

AbbVie Deutschland GmbH & Co.KG

Modul 4 A

Behandlung der mittelschweren bis schweren aktiven Colitis ulcerosa bei erwachsenen Patienten, die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 24.08.2022

Inhaltsverzeichnis

Seite

Tabellenverzeichnis ... 4
Abbildungsverzeichnis ... 12
Abkürzungsverzeichnis ... 14
4 Modul 4 – allgemeine Informationen ... 18
4.1 Zusammenfassung der Inhalte von Modul 4 ... 19
4.2 Methodik ... 33
4.2.1 Fragestellung ... 33
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung ... 35
4.2.3 Informationsbeschaffung ... 40
4.2.3.1 Studien des pharmazeutischen Unternehmers ... 40
4.2.3.2 Bibliografische Literaturrecherche ... 41
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken ... 42
4.2.3.4 Suche auf der Internetseite des G-BA ... 43
4.2.3.5 Selektion relevanter Studien ... 44
4.2.4 Bewertung der Aussagekraft der Nachweise ... 45
4.2.5 Informationssynthese und -analyse ... 47
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien ... 47
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien ... 48
4.2.5.3 Meta-Analysen ... 67
4.2.5.4 Sensitivitätsanalysen ... 69
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren ... 70
4.2.5.6 Indirekte Vergleiche ... 71
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen ... 75
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel ... 75
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel ... 75
4.3.1.1.1 Studien des pharmazeutischen Unternehmers ... 75
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche ... 77
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken ... 79
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA ... 80
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel ... 81
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel ... 83
4.3.1.2.1 Studiendesign und Studienpopulationen ... 83
4.3.1.2.2 Verzerrungspotenzial auf Studienebene ... 85
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien ... 86
4.3.1.3.1 <Endpunkt xxx> – RCT ... 87
4.3.1.3.2 Subgruppenanalysen – RCT ... 91
4.3.1.4 Liste der eingeschlossenen Studien – RCT ... 94

4.3.2 Weitere Unterlagen 94
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien 94
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche 94
4.3.2.1.1.1 Studien des pharmazeutischen Unternehmers 95
4.3.2.1.1.2 Studien aus der bibliografischen Literaturrecherche 96
4.3.2.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken 100
4.3.2.1.1.4 Studien aus der Suche auf der Internetseite des G-BA 101
4.3.2.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel 102
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche 104
4.3.2.1.2.1 Studiendesign und Studienpopulationen 104
4.3.2.1.2.2 Verzerrungspotenzial auf Studienebene 141
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen 145
4.3.2.1.3.1 Mortalität – indirekte Vergleiche aus RCT 145
4.3.2.1.3.2 Klinische Remission – indirekte Vergleiche aus RCT 149
4.3.2.1.3.3 Klinisches Ansprechen – indirekte Vergleiche aus RCT 163
4.3.2.1.3.4 Mukosaheilung – indirekte Vergleiche aus RCT 172
4.3.2.1.3.5 Stuhlfrequenz – indirekte Vergleiche aus RCT 180
4.3.2.1.3.6 Rektale Blutungen – indirekte Vergleiche aus RCT 189
4.3.2.1.3.7 EQ-5D VAS – indirekte Vergleiche aus RCT 198
4.3.2.1.3.8 Gesundheitsbezogene Lebensqualität: IBDQ – indirekte Vergleiche aus RCT 207
4.3.2.1.3.9 Gesundheitsbezogene Lebensqualität: SF-36 – indirekte Vergleiche aus RCT 216
4.3.2.1.3.10 Verträglichkeit – indirekte Vergleiche aus RCT 230
4.3.2.1.4 Subgruppenanalysen – indirekte Vergleiche aus RCT 246
4.3.2.1.4.1 Klinische Remission – Subgruppen 246
4.3.2.1.4.2 Mukosaheilung - Subgruppen 247
4.3.2.1.5 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT 247
4.3.2.2 Nicht randomisierte vergleichende Studien 248
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien 248
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien 248
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien 249
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien 250
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien 251
4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien 251
4.3.2.3 Weitere Untersuchungen 251
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen 251
4.3.2.3.2 Charakteristika der weiteren Untersuchungen 252
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen 252
4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen 252
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen 253
4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen 253

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens 254
4.4.1 Beurteilung der Aussagekraft der Nachweise 254
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß 258
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht 271
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte 272
4.5.1 Begründung für die Vorlage indirekter Vergleiche 272
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen 273
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen 273
4.5.4 Verwendung von Surrogatendpunkten 273
4.6 Referenzliste 274
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche 282
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken 292
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche) 295
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken) 297
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT 307
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 350

Tabellenverzeichnis

Seite

Tabelle 4-1: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien mit dem zu bewertenden Arzneimittel – RCT 21

Tabelle 4-2: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für den indirekten Vergleich (RCT) mit dem zu bewertenden Arzneimittel 22

Tabelle 4-3: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für den indirekten Vergleich (RCT) mit der ZVT (Ustekinumab)..... 24

Tabelle 4-4: Zusammenfassung der Ergebnisse des adjustierten indirekten Vergleichs der Studien U-ACHIEVE, U-ACCOMPLISH und UNIFI und daraus resultierender Zusatznutzen..... 27

Tabelle 4-5: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien mit dem zu bewertenden Arzneimittel – RCT 35

Tabelle 4-6: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für den indirekten Vergleich (RCT) mit dem zu bewertenden Arzneimittel 37

Tabelle 4-7: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für den indirekten Vergleich (RCT) mit der ZVT (Ustekinumab)..... 39

Tabelle 4-8: Nutzendimensionen und patientenrelevante Endpunkte zur Bewertung des Ausmaßes des medizinischen Zusatznutzens von Upadacitinib 50

Tabelle 4-9: Domänen und Summenskalen des SF-36-Fragebogens..... 59

Tabelle 4-10: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel 76

Tabelle 4-11: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel..... 77

Tabelle 4-12: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern/ Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel..... 80

Tabelle 4-13: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel..... 81

Tabelle 4-14: Studienpool – RCT mit dem zu bewertenden Arzneimittel..... 82

Tabelle 4-15: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 84

Tabelle 4-16: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel..... 85

Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel..... 85

Tabelle 4-18: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel..... 86

Tabelle 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel 86

Tabelle 4-20: Operationalisierung von <Endpunkt xxx> ... 90
Tabelle 4-21: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT mit dem zu bewertenden Arzneimittel ... 90
Tabelle 4-22: Ergebnisse für <Endpunkt xxx> aus RCT mit dem zu bewertenden Arzneimittel ... 91
Tabelle 4-23 Matrix der durchgeführten Subgruppenanalysen ... 93
Tabelle 4-24: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für <Studie> und <Effektmodifikator> ... 93
Tabelle 4-25: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel ... 95
Tabelle 4-26: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel ... 96
Tabelle 4-27: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern/ Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel ... 100
Tabelle 4-28: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel ... 102
Tabelle 4-29: Studienpool – RCT mit dem zu bewertenden Arzneimittel ... 103
Tabelle 4-30: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel und der ZVT für indirekte Vergleiche ... 105
Tabelle 4-31: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel und der ZVT für indirekte Vergleiche ... 109
Tabelle 4-32: Dosierung von Ustekinumab nach Körpergewicht ... 112
Tabelle 4-33: Charakterisierung der Studienpopulationen – Demografische Charakteristika zu Baseline nach Behandlungsarm in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel ... 114
Tabelle 4-34: Charakterisierung der Studienpopulationen – Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel ... 115
Tabelle 4-35: Charakterisierung der Studienpopulationen – Medizinische Vorgeschichte zu Baseline nach Behandlungsarm in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel ... 118
Tabelle 4-36: Patientenstatus in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel ... 119
Tabelle 4-37: Behandlungsdauer (Tage) in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel ... 120
Tabelle 4-38: Charakterisierung der Studienpopulationen – Demografische Charakteristika zu Baseline nach Behandlungsarm in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich ... 121

Tabelle 4-39: Charakterisierung der Studienpopulationen – Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich ... 121

Tabelle 4-40: Charakterisierung der Studienpopulationen – Medizinische Vorgeschichte zu Baseline nach Behandlungsarm in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich ... 122

Tabelle 4-41: Patientenstatus in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich ... 123

Tabelle 4-42: Behandlungs- und Beoachtungsdauer in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich ... 124

Tabelle 4-43: Charakterisierung der Studienpopulationen – Demografische Charakteristika zu Baseline nach Behandlungsarm in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel ... 124

Tabelle 4-44: Charakterisierung der Studienpopulationen – Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel ... 125

Tabelle 4-45: Charakterisierung der Studienpopulationen – Medizinische Vorgeschichte zu Baseline nach Behandlungsarm in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel ... 126

Tabelle 4-46: Patientenstatus in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel ... 127

Tabelle 4-47: Behandlungsdauer (Tagen) in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel ... 128

Tabelle 4-48: Charakterisierung der Studienpopulationen – Demografische Charakteristika zu Baseline nach Behandlungsarm in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich ... 129

Tabelle 4-49: Charakterisierung der Studienpopulationen – Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich ... 129

Tabelle 4-50: Charakterisierung der Studienpopulationen – Medizinische Vorgeschichte zu Baseline nach Behandlungsarm in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich ... 130

Tabelle 4-51: Patientenstatus in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich ... 131

Tabelle 4-52: Behandlungsdauer (Tagen) in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich ... 132

Tabelle 4-53: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel ... 142

Tabelle 4-54: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche ... 145

Tabelle 4-55: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden ... 146

Tabelle 4-56: Operationalisierung der Mortalität 147
Tabelle 4-57: Bewertung des Verzerrungspotenzials für Mortalität in RCT für indirekte Vergleiche 147
Tabelle 4-58: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 149
Tabelle 4-59: Operationalisierung von klinische Remission 150
Tabelle 4-60: Bewertung des Verzerrungspotenzials für klinische Remission in RCT für indirekte Vergleiche 152
Tabelle 4-61: Ergebnisse für Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore) – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 153
Tabelle 4-62: Ergebnisse für Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore) – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 155
Tabelle 4-63: Ergebnisse für Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 156
Tabelle 4-64: Ergebnisse für Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 157
Tabelle 4-65: Ergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 158
Tabelle 4-66: Ergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 160
Tabelle 4-67: Ergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 161
Tabelle 4-68: Ergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 162
Tabelle 4-69: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 163
Tabelle 4-70: Operationalisierung von klinisches Ansprechen 164
Tabelle 4-71: Bewertung des Verzerrungspotenzials für klinisches Ansprechen in RCT für indirekte Vergleiche 166
Tabelle 4-72: Ergebnisse für Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore) – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 167
Tabelle 4-73: Ergebnisse für Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore) – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 168
Tabelle 4-74: Ergebnisse für Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 170

Tabelle 4-75: Ergebnisse für Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 171
Tabelle 4-76: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 172
Tabelle 4-77: Operationalisierung von Mukosaheilung 173
Tabelle 4-78: Bewertung des Verzerrungspotenzials für Mukosaheilung in RCT für indirekte Vergleiche 175
Tabelle 4-79: Ergebnisse für Anteil der Patienten mit Mukosaheilung – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 176
Tabelle 4-80: Ergebnisse für Anteil der Patienten mit Mukosaheilung – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 177
Tabelle 4-81: Ergebnisse für Anteil der Patienten mit Mukosaheilung – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 178
Tabelle 4-82: Ergebnisse für Anteil der Patienten mit Mukosaheilung – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 179
Tabelle 4-83: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 180
Tabelle 4-84: Operationalisierung von Stuhlfrequenz 181
Tabelle 4-85: Bewertung des Verzerrungspotenzials für Stuhlfrequenz in RCT für indirekte Vergleiche 183
Tabelle 4-86: Ergebnisse für Anteil der Patienten mit Stuhlfrequenz ≤ 1 – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 184
Tabelle 4-87: Ergebnisse für Anteil der Patienten mit Stuhlfrequenz ≤ 1 – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 185
Tabelle 4-88: Ergebnisse für Anteil der Patienten mit Stuhlfrequenz ≤ 1 – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 187
Tabelle 4-89: Ergebnisse für Anteil der Patienten mit Stuhlfrequenz ≤ 1 – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 188
Tabelle 4-90: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 189
Tabelle 4-91: Operationalisierung von rektale Blutung 190
Tabelle 4-92: Bewertung des Verzerrungspotenzials für rektale Blutung in RCT für indirekte Vergleiche 192
Tabelle 4-93: Ergebnisse für Anteil der Patienten mit rektaler Blutung = 0 – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 193
Tabelle 4-94: Ergebnisse für Anteil der Patienten mit rektaler Blutung = 0 – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 194
Tabelle 4-95: Ergebnisse für Anteil der Patienten mit rektaler Blutung = 0 – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 196

Tabelle 4-96: Ergebnisse für Anteil der Patienten mit rektaler Blutung = 0 – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 197
Tabelle 4-97: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 198
Tabelle 4-98: Operationalisierung von EQ-5D VAS 199
Tabelle 4-99: Bewertung des Verzerrungspotenzials für EQ-5D VAS in RCT für indirekte Vergleiche 201
Tabelle 4-100: Ergebnisse für mittlere Veränderung der EQ-5D VAS – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 202
Tabelle 4-101: Ergebnisse für mittlere Veränderung der EQ-5D VAS – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 204
Tabelle 4-102: Ergebnisse für mittlere Veränderung der EQ-5D VAS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 205
Tabelle 4-103: Ergebnisse für mittlere Veränderung der EQ-5D VAS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 206
Tabelle 4-104: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 207
Tabelle 4-105: Operationalisierung von IBDQ 208
Tabelle 4-106: Bewertung des Verzerrungspotenzials für IBDQ in RCT für indirekte Vergleiche 210
Tabelle 4-107: Ergebnisse für mittlere Veränderung des IBDQ – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 211
Tabelle 4-108: Ergebnisse für mittlere Veränderung des IBDQ – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 213
Tabelle 4-109: Ergebnisse für mittlere Veränderung des IBDQ – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 214
Tabelle 4-110: Ergebnisse für mittlere Veränderung des IBDQ – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 215
Tabelle 4-111: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 216
Tabelle 4-112: Operationalisierung von SF-36 PCS/MCS 217
Tabelle 4-113: Bewertung des Verzerrungspotenzials für SF-36 PCS/MCS in RCT für indirekte Vergleiche 219
Tabelle 4-114: Ergebnisse für mittlere Veränderung des SF-36 PCS – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche 220
Tabelle 4-115: Ergebnisse für mittlere Veränderung des SF-36 PCS – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche 222
Tabelle 4-116: Ergebnisse für mittlere Veränderung des SF-36 PCS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase 223
Tabelle 4-117: Ergebnisse für mittlere Veränderung des SF-36 PCS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase 224

Tabelle 4-118: Ergebnisse für mittlere Veränderung des SF-36 MCS – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche ... 225
Tabelle 4-119: Ergebnisse für mittlere Veränderung des SF-36 MCS – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche ... 227
Tabelle 4-120: Ergebnisse für mittlere Veränderung des SF-36 MCS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase ... 228
Tabelle 4-121: Ergebnisse für mittlere Veränderung des SF-36 MCS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase ... 229
Tabelle 4-122: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden ... 230
Tabelle 4-123: Operationalisierung von Verträglichkeit ... 231
Tabelle 4-124: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT für indirekte Vergleiche ... 234
Tabelle 4-125: Ergebnisse für UE Gesamtraten – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche ... 235
Tabelle 4-126: Ergebnisse für UE Gesamtraten – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche ... 239
Tabelle 4-127: Ergebnisse für UE Gesamtraten - Upadacitinib gegenüber Ustekinumab - aus RCT für indirekte Vergleiche Induktionsphase ... 241
Tabelle 4-128: Ergebnisse für UE Gesamtraten – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase ... 243
Tabelle 4-129: Weitere Ergebnisse für Verträglichkeit – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche Erhaltungsphase ... 244
Tabelle 4-130: Subgruppenergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) aus RCT (U-ACHIEVE Substudie 3) mit dem zu bewertenden Arzneimittel ... 246
Tabelle 4-131: Subgruppenergebnisse für Anteil der Patienten mit Mukosaheilung aus RCT (U-ACHIEVE Substudie 3) mit dem zu bewertenden Arzneimittel ... 247
Tabelle 4-132: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien ... 249
Tabelle 4-133: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien ... 249
Tabelle 4-134: Operationalisierung von <Endpunkt xxx> ... 250
Tabelle 4-135: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien ... 250
Tabelle 4-136: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen ... 252
Tabelle 4-137: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen ... 253
Tabelle 4-138: Zusammenfassung der Ergebnisse des adjustierten indirekten Vergleichs der Studien U-ACHIEVE, U-ACCOMPLISH und UNIFI und daraus resultierender Zusatznutzen ... 260

Tabelle 4-139: Zusammenfassung der positiven und negativen Effekte von Upadacitinib gegenüber Ustekinumab aus dem adjustierten indirekten Vergleich 270
Tabelle 4-140: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens 271
Tabelle 4-141 (Anhang): Studiendesign und -methodik für Studie U-ACHIEVE Substudie 3 308
Tabelle 4-142 (Anhang): Studiendesign und -methodik für Studie UNIFI Erhaltungsstudie 328
Tabelle 4-143 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie U-ACHIEVE Substudie 1 351
Tabelle 4-144 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie U-ACHIEVE Substudie 2 363
Tabelle 4-145 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie U-ACCOMPLISH 375
Tabelle 4-146 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UNIFI Induktionsstudie 387
Tabelle 4-147 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie U-ACHIEVE Substudie 3 399
Tabelle 4-148 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UNIFI Erhaltungsstudie 411

Abbildungsverzeichnis

Seite

Abbildung 1: Netzwerkstruktur eines indirekten Vergleichs zwischen Upadacitinib und Ustekinumab über den Brückenkomparator Placebo 73

Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel 78

Abbildung 3: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel 79

Abbildung 4: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel 97

Abbildung 5: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel – indirekte Vergleiche 98

Abbildung 6: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit der ZVT (Ustekinumab) – indirekte Vergleiche ... 99

Abbildung 7: Netzwerkstruktur des indirekten Vergleichs 146

Abbildung 8: Netzwerkstruktur des indirekten Vergleichs 150

Abbildung 9: Forest-Plot für klinische Remission (Mayo Gesamtscore) zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 154

Abbildung 10: Forest-Plot für klinische Remission (Adaptierter Mayo Score) zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 159

Abbildung 11: Netzwerkstruktur des indirekten Vergleichs 164

Abbildung 12: Forest-Plot für klinisches Ansprechen (Mayo Gesamtscore) zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 168

Abbildung 13: Netzwerkstruktur des indirekten Vergleichs 173

Abbildung 14: Forest-Plot für Mukosaheilung zu Woche 8 aus RCT (U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 177

Abbildung 15: Netzwerkstruktur des indirekten Vergleichs 181

Abbildung 16: Forest-Plot für Stuhlfrequenz ≤ 1 zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 185

Abbildung 17: Netzwerkstruktur des indirekten Vergleichs 190

Abbildung 18: Forest-Plot für rektale Blutung = 0 zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 194

Abbildung 19: Netzwerkstruktur des indirekten Vergleichs 199

Abbildung 20: Forest-Plot für die mittlere Veränderung der EQ-5D VAS zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 204
Abbildung 21: Netzwerkstruktur des indirekten Vergleichs 208
Abbildung 22: Forest-Plot für die mittlere Veränderung des IBDQ zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 213
Abbildung 23: Netzwerkstruktur des indirekten Vergleichs 217
Abbildung 24: Forest-Plot für die mittlere Veränderung des SF-36 PCS zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 222
Abbildung 25: Forest-Plot für die mittlere Veränderung des SF-36 MCS zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 227
Abbildung 26: Netzwerkstruktur des indirekten Vergleichs 231
Abbildung 27: Forest-Plot für Alle UE zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 238
Abbildung 28: Forest-Plot für Schwerwiegende UE zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 238
Abbildung 29: Forest-Plot für UE, die zum Therapieabbruch führten zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 239
Abbildung 30: Forest-Plot für Todesfälle zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo 239
Abbildung 31: Flussdiagramm für Studie U-ACHIEVE Substudie 1 324
Abbildung 32: Flussdiagramm für Studie U-ACHIEVE Substudie 2 325
Abbildung 33: Flussdiagramm für Studie U-ACCOMPLISH 326
Abbildung 34: Flussdiagramm für Studie U-ACHIEVE Substudie 3 327
Abbildung 35: Flussdiagramm für Studie UNIFI (Induktionsphase) 347
Abbildung 36: Flussdiagramm für Studie UNIFI (Erhaltungsphase; randomisierte Population) 348
Abbildung 37: Flussdiagramm für Studie UNIFI (Erhaltungsphase; nicht randomisierte Population) 349

Abkürzungsverzeichnis

Abkürzung	Bedeutung
6-MP	6-Mercaptopurin
6-TG	6-Thioguanin
ALT	Alanin-Aminotransferase
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
AMIce	Arzneimittel-Informationssystem (vormals AMIS)
AMIS	Arzneimittel-Informationssystem
ANCOVA	Kovarianzanalyse (Analysis of Covariance)
ASA	Aminosalicylsäure
AST	Aspartat-Aminotransferase
AWG	Anwendungsgebiet
AZA	Azathioprin
BCG	Bacille-Calmette-Guerin
Bio-IR	Intoleranz oder unzureichendes Ansprechen auf eine vorherige biologische Therapie (Biologic-Inadequate Response)
BMI	Body Mass Index
CMH	Cochran-Mantel-Haenszel
Cochrane	Cochrane Central Register of Controlled Trials
CONSORT	Consolidated Standards of Reporting Trials
CRP	C-reaktives Protein
CTCAE	Common Terminology Criteria for Adverse Events
CYP	Cytochrom P450
CU	Colitis ulcerosa
DIMDI	Deutsches Institut für Medizinische Dokumentation
DNA	Desoxyribonukleinsäure (Deoxyribonucleic Acid)
EKG	Elektrokardiogramm
EMA	Europäische Arzneimittel-Agentur (European Medicines Agency)
EMBASE	Excerpta Medica Database
EQ-5D	EuroQol-5 dimensions
EU-CTR	European Union Clinical Trials Register
FACIT-F	Functional Assessment of Chronic Illness Therapy – Fatigue
FDA	Food and Drug Administration

Abkürzung	Bedeutung
FU	Follow-up
G-BA	Gemeinsamer Bundesausschuss
GCAP	Zytapheresebehandlung
HBsAg	HBV-Oberflächenantigen
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HIV	Humanes Immundefizienzvirus
hsCRP	Hoch sensitives C-reaktives Protein (high sensitivity C-Reactive Protein
IBD	Chronisch-entzündliche Darmerkrankung (Inflammatory Bowel Disease)
IBDQ	Inflammatory Bowel Disease Questionnaire
ICTRP	International Clinical Trials Registry Platform
IQR	Interquartilsabstand (Interquartile Range)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Interaktives Response Technologie System (Interactive Response Technology)
ITT	Intention-to-Treat
i.v.	Intravenös
JAK	Januskinase
k. A.	Keine Angabe
KI	Konfidenzintervall
LCAP	Leukocytapheresis
LOCF	Last Observation Carried Forward
LS-MW	Least-Squares-Mittelwert
MACE	Major Adverse Cardiovascular Events
MAR	Missing-at-Random
MCS	Mental Health Component Summary Score
MDRD	Modification of Diet in Renal Disease
MedDRA	Medical Dictionary for Regulatory Activities
MEDLINE	Medical Literature Analysis and Retrieval System Online
MI	Multiple Imputation
MMRM	Mixed Model for Repeated Measures

Abkürzung	Bedeutung
MTC	Mixed Treatment Comparison
MTX	Methotrexat
MW	Mittelwert
NMSC	Nicht melanozytärer Hautkrebs (Non Melanoma Skin Cancer)
NRI	Non-Responder-Imputation
NRI-C	Non-Responder-Imputation incorporating multiple imputation to handle missing data due to COVID-19
OR	Odds-Ratio
PCS	Physical Health Component Summary Score
p.o.	Peroral
PT	Preferred Term nach MedDRA
Q8W	Alle 8 Wochen
Q12W	Alle 12 Wochen
QD	Einmal täglich (Quaque Die)
RBS	Rektale Blutung (Rectal Bleeding Score)
RCT	Randomisierte kontrollierte Studie (Randomized Controlled Trial)
RD	Risikodifferenz
REML	Restriktive Maximum-Likelihood-Methode
RNA	Ribonukleinsäure (Ribonucleic Acid)
RR	Relatives Risiko
RTB-MI	Multiple Imputation Incorporating Return-to-Baseline
SAP	Statistischer Analyseplan
s.c.	Subkutan
SD	Standardabweichung (Standard Deviation)
SE	Standardfehler (Standard Error)
SF-36	Short Form – 36
SFS	Stuhlfrequenz (Stool Frequency Score)
SGB	Sozialgesetzbuch
SMQs	Standardised MedDRA Queries
SOC	Systemorganklasse (System Organ Class) nach MedDRA
STE	Surrogate Threshold Effects

Abkürzung	Bedeutung
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes UE
TB	Tuberkulose
TNF	Tumornekrosefaktor
TPN	Totale parenterale Ernährung
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UESI	UE vom speziellem Interesse
UE	Unerwünschtes Ereignis
ULN	Obere Normgrenze (Upper Limit of Normal)
VAS	Visuelle Analogskala
VerfO	Verfahrensordnung (des G-BA)
VTE	Venöse thromboembolische Ereignisse
WHO	World Health Organization (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

Ziel des vorliegenden Dossiers ist die Bewertung des medizinischen Nutzens und medizinischen Zusatznutzens von Upadacitinib zur Behandlung der mittelschweren bis schweren aktiven Colitis ulcerosa (CU) bei erwachsenen Patienten, die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben.

Die Bewertung erfolgt gegenüber der vom Gemeinsamen Bundesausschuss (G-BA) festgelegten zweckmäßigen Vergleichstherapie (ZVT) anhand patientenrelevanter Endpunkte aus den Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit. Basis dieser Nutzenbewertung ist ein adjustierter indirekter Vergleich nach Bucher, in welchem Upadacitinib mit der ZVT Ustekinumab über den Brückenkomparator Placebo verglichen wird.

Im beschriebenen Anwendungsgebiet wurden vom G-BA wie folgt zwei Populationen mit entsprechenden ZVT festgelegt:

Population A:

Erwachsene Patienten mit mittelschwerer bis schwerer aktiver Colitis ulcerosa, die auf eine konventionelle Therapie unzureichend angesprochen haben, nicht mehr darauf ansprechen oder eine Unverträglichkeit oder Kontraindikation aufweisen.

- Ein TNF-Inhibitor (Adalimumab oder Infliximab oder Golimumab) oder Vedolizumab oder Ustekinumab

Population B:

Erwachsene Patienten mit mittelschwerer bis schwerer aktiver Colitis ulcerosa, die auf ein Biologikum (TNF-Inhibitor oder Integrin-Inhibitor oder Interleukin-Inhibitor) unzureichend angesprochen haben, nicht mehr darauf ansprechen oder eine Unverträglichkeit gegen eine entsprechende Behandlung aufweisen.

- Ein Wechsel der Therapie auf Vedolizumab oder Tofacitinib oder Ustekinumab oder einen TNF-Inhibitor (Adalimumab oder Infliximab oder Golimumab), jeweils unter Berücksichtigung der Zulassung und der Vortherapie(n)

Hinweise:

Sofern Infliximab eingesetzt wird, sollte es ggf. in Kombinationstherapie mit einem Thiopurin eingesetzt werden.

Im Rahmen des indirekten Vergleichs ist eine getrennte Beantwortung dieser Fragestellungen nicht möglich, da die Daten nicht für alle Endpunktdimensionen bzw. patientenrelevante Endpunkte, insbesondere nicht für die Verträglichkeit, getrennt nach Vorbehandlung vorliegen.

Analysen zur Zielpopulation sind aber im Rahmen des indirekten Vergleichs adäquat, da ein gemeinsames Therapieziel für beide Populationen zu Grunde liegt. Da Ustekinumab für beide Fragestellungen als ZVT durch den G-BA definiert wurde und sich bei einer Behandlung unter Upadacitinib für die Erhaltungsphase in der dargestellten, deskriptiven Subgruppenanalyse, vergleichbare Ansprechraten für Bio-IR- und Non-Bio-IR-Patienten zeigen, werden beide Fragestellungen zu einer Fragestellung zusammengefasst und Ergebnisse für die Zielpopulation dargestellt.

Datenquellen

Die bibliografische Literaturrecherche wurde in den Datenbanken MEDLINE, EMBASE und Cochrane (Cochrane Central Register of Controlled Trials) durchgeführt, wobei für jede einzelne Datenbank eine sequenzielle, individuell adaptierte Suchstrategie entwickelt und separat angewandt wurde. Für die Suche wurde die Plattform OVID verwendet.

Die Studienregistersuche wurde in den Portalen clinicaltrials.gov (www.clinicaltrials.gov), International Clinical Trials Registry Platform (ICTRP) Search Portal, Suchportal der Weltgesundheitsorganisation (World Health Organization, WHO) (http://apps.who.int/trialsearch/) und im European Union Clinical Trials Registry (EU-CTR) (https://www.clinicaltrialsregister.eu/ctr-search/search) durchgeführt. Es wurde nach abgeschlossenen, laufenden und abgebrochenen Studien gesucht. Zudem wurde im Clinical Data Suchportal der Europäischen Arzneimittel-Agentur (European Medicines Agency, EMA) (https://clinicaldata.ema.europa.eu), im Arzneimittel-Informationssystem AMIce (https://www.dimdi.de/dynamic/de/arzneimittel/arzneimittel-recherchieren/amis/) sowie auf der Internetseite des G-BA nach Einträgen und Ergebnisberichten zu denjenigen Studien gesucht, die zuvor schon als relevante Studien identifiziert wurden.

Ein-/Ausschlusskriterien für Studien

Studien, die für die Beantwortung der Fragestellung geeignet sind, wurden anhand von prädefinierten Ein- und Ausschlusskriterien identifiziert. Eine zusammenfassende Übersicht der Ein- und Ausschlusskriterien findet sich in Tabelle 4-1.

Tabelle 4-1: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien mit dem zu bewertenden Arzneimittel – RCT

Kriterium	Einschluss		Ausschluss		Begründung
Population	Erwachsene Patienten mit mittelschwerer bis schwerer aktiver CU die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben	E1	Andere, nicht dem zugelassenen Anwendungsgebiet entsprechende, Patientenpopulation	A1	Population des zugelassenen Anwendungs-gebiets
Intervention(en)	Behandlung mit Upadacitinib gemäß Fachinformation	E2	Andere Therapien	A2	Zu bewertendes Arzneimittel
Vergleichs-therapie	Tumornekrosefaktor (TNF)-Inhibitor (Adalimumab, Golimumab oder Infliximab) oder Ustekinumab oder Vedolizumab oder Tofacitinib	E3	Andere Therapien oder Behandlung mit Vergleichstherapien abweichend von der Fachinformation	A3	ZVT gemäß G-BA
Endpunkte	Mindestens ein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	E4	Kein patienten-relevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	A4	Nutzendimensionen nach § 5 Absatz 2 des 5. Kapitels der VerfO
Studientyp	RCT	E5	Keine RCT	A5	RCTs bieten aufgrund ihres Designs die höchste Ergebnissicherheit. Für den Fall, dass keine RCTs identifiziert werden, wird die beste verfügbare Evidenz ermittelt.
Studiendauer	Mindestens 52 Wochen	E6	Weniger als 52 Wochen	A6	Mindestdauer gemäß G-BA-Beratung

Kriterium	Einschluss		Ausschluss		Begründung
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügt[a]	E7	• Kein Studienbericht oder keine Vollpublikation verfügbar sowie auch keine publizierten Ergebnisse auf der Webseite offizieller Studienregister[b] • Mehrfachpublikation ohne relevante Zusatzinformationen[c]	A7	Ausreichende Datenverfügbarkeit ist zur Bewertung von Ergebnissen erforderlich.

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

b: Keine publizierten Ergebnisse auf der Webseite offizieller Studienregister, die den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügen

c: Jeglicher Publikationstyp in welchem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden. Dies gilt auch für in der bibliografischen Literaturrecherche identifizierte Studienregistereinträge, welche ebenfalls in der Studienregistersuche identifiziert werden.

CONSORT: Consolidated Standards of Reporting Trials; CU: Colitis ulcerosa; G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TNF: Tumornekrosefaktor; TREND: Transparent Reporting of Evaluations with Non-randomized Design; VerfO: Verfahrensordnung (des G-BA)

Ein- und Ausschlusskriterien für die Suche nach Studien mit Upadacitinib im Vergleich zu Placebo für adjustierte indirekte Vergleiche sind in Tabelle 4-2 aufgeführt.

Tabelle 4-2: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für den indirekten Vergleich (RCT) mit dem zu bewertenden Arzneimittel

Kriterium	Einschluss		Ausschluss		Begründung
Population	Erwachsene Patienten mit mittelschwerer bis schwerer aktiver CU die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben.	E1	Andere, nicht dem zugelassenen Anwendungsgebiet entsprechende, Patientenpopulation	A1	Population des zugelassenen Anwendungs-gebiets
Intervention(en)	Behandlung mit Upadacitinib gemäß Fachinformation	E2	Andere Therapien	A2	Zu bewertendes Arzneimittel

Kriterium	Einschluss		Ausschluss		Begründung
Vergleichstherapie	Placebo	E3	Andere Therapien oder Behandlung mit Vergleichstherapien abweichend von der Fachinformation	A3	Brückenkomparator für adjustierten indirekten Vergleich
Endpunkte	Mindestens ein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	E4	Kein patienten-relevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	A4	Nutzendimensionen nach § 5 Absatz 2 des 5. Kapitels der VerfO
Studientyp	RCT	E5	Keine RCT	A5	RCTs bieten aufgrund ihres Designs die höchste Ergebnissicherheit. Für den Fall, dass keine RCTs identifiziert werden, wird die beste verfügbare Evidenz ermittelt.
Studiendauer	Mindestens 52 Wochen	E6	Weniger als 52 Wochen	A6	Mindestdauer gemäß G-BA-Beratung
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügt[a]	E7	• Kein Studienbericht oder keine Vollpublikation verfügbar sowie auch keine publizierten Ergebnisse auf der Webseite offizieller Studienregister[b] • Mehrfachpubli-kation ohne relevante Zusatz-informationen[c]	A7	Ausreichende Datenverfügbarkeit ist zur Bewertung von Ergebnissen erforderlich.

Kriterium	Einschluss		Ausschluss		Begründung
a: Studien, deren Ergebnisse auf der Webseite offizieller Studienregister publiziert wurden (z. B. clinicaltrials.gov), können ebenfalls berücksichtigt werden. b: Keine publizierten Ergebnisse auf der Webseite offizieller Studienregister, die den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügen c: Jeglicher Publikationstyp in welchem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden. Dies gilt auch für in der bibliografischen Literaturrecherche identifizierte Studienregistereinträge, welche ebenfalls in der Studienregistersuche identifiziert werden. CONSORT: Consolidated Standards of Reporting Trials; CU: Colitis ulcerosa; G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Non-randomized Design; VerfO: Verfahrensordnung (des G-BA); ZVT: Zweckmäßige Vergleichstherapie					

Ein- und Ausschlusskriterien für die Suche nach der ZVT Ustekinumab für indirekte Vergleiche (RCT) sind in Tabelle 4-3 zusammengefasst.

Tabelle 4-3: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für den indirekten Vergleich (RCT) mit der ZVT (Ustekinumab)

Kriterium	Einschluss		Ausschluss		Begründung
Population	Erwachsene Patienten mit mittelschwerer bis schwerer aktiver CU die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben.	E1	Andere, nicht dem zugelassenen Anwendungsgebiet entsprechende, Patientenpopulation	A1	Population des zugelassenen Anwendungsgebiets
Intervention(en)	Behandlung mit Ustekinumab gemäß Fachinformation	E2	Andere Therapien	A2	ZVT (Komparator für adjustierten indirekten Vergleich)
Vergleichstherapie	Placebo	E3	Andere Therapien oder Behandlung mit Vergleichstherapien abweichend von der Fachinformation	A3	Brückenkomparator für adjustierten indirekten Vergleich
Endpunkte	Mindestens ein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	E4	Kein patienten-relevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	A4	Nutzendimensionen nach § 5 Absatz 2 des 5. Kapitels der VerfO

Kriterium	Einschluss		Ausschluss		Begründung
Studientyp	RCT	E5	Keine RCT	A5	RCTs bieten aufgrund ihres Designs die höchste Ergebnissicherheit. Für den Fall, dass keine RCTs identifiziert werden, wird die beste verfügbare Evidenz ermittelt.
Studiendauer	Mindestens 52 Wochen	E6	Weniger als 52 Wochen	A6	Mindestdauer gemäß G-BA-Beratung
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügt[a]	E7	• Kein Studienbericht oder keine Vollpublikation verfügbar sowie auch keine publizierten Ergebnisse auf der Webseite offizieller Studienregister[b] • Mehrfachpublikation ohne relevante Zusatzinformationen[c]	A7	Ausreichende Datenverfügbarkeit ist zur Bewertung von Ergebnissen erforderlich.

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

b: Keine publizierten Ergebnisse auf der Webseite offizieller Studienregister, die den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügen

c: Jeglicher Publikationstyp in welchem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden. Dies gilt auch für in der bibliografischen Literaturrecherche identifizierte Studienregistereinträge, welche ebenfalls in der Studienregistersuche identifiziert werden.

CONSORT: Consolidated Standards of Reporting Trials; CU: Colitis ulcerosa; G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Non-randomized Design; VerfO: Verfahrensordnung (des G-BA); ZVT: Zweckmäßige Vergleichstherapie

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

Zur Beurteilung des Zusatznutzens von Upadacitinib liegt keine relevante direktvergleichende klinische Studie vor. Für die Bewertung des medizinischen Nutzens und medizinischen Zusatznutzens von Upadacitinib im oben genannten Anwendungsgebiet (AWG) wird daher ein adjustierter, indirekter Vergleich nach Bucher mit der ZVT Ustekinumab durchgeführt. Da die

darin berücksichtigten Studien U-ACHIEVE und UNIFI placebokontrolliert sind, wurde als Brückenkomparator der Placebo-Arm gewählt.

Zur Ableitung des Zusatznutzens von Upadacitinib liegt für den adjustierten indirekten Vergleich für das zu bewertende Arzneimittel die Studie U-ACHIEVE (M14-234) vor. Hierbei handelt es sich um eine multizentrische, randomisierte, doppelt-verblindete, placebokontrollierte Phase-IIb/III-Studie der Evidenzstufe Ib (nach §5 5. Kapitel, 1. Abschnitt, Absatz 6 der Verfahrensordnung [VerfO] des G-BA). Bei der Studie U-ACCOMPLISH (M14-675), welche im Dossier ergänzend dargestellt wird, handelt es sich um eine abgeschlossene, multizentrische, randomisierte, doppelt-verblindete, placebokontrollierte Phase-III-Studie der Evidenzstufe Ib (nach §5 5. Kapitel, 1. Abschnitt, Absatz 6 der VerfO des G-BA). Die Studie U-ACHIEVE umfasst drei Substudien: eine Phase-IIb-Dosierungsstudie für die Induktionsphase (Substudie 1), eine Phase-III-Induktionsstudie (Substudie 2) und eine Phase-III-Erhaltungsstudie (Substudie 3). Die Studie U-ACCOMPLISH ist eine weitere Phase-III-Induktionsstudie. Die Induktionsstudien werden als ergänzende Evidenz zur Rekapitulation des deutschen Versorgungskontext dargestellt und zur Ableitung des Zusatznutzens herangezogen.

Für den adjustierten indirekten Vergleich wird für die ZVT Ustekinumab die Studie UNIFI herangezogen. Hierbei handelt es sich um eine randomisierte, doppelblinde, multizentrische, placebokontrollierte Phase-III-Studie, die die Sicherheit und Wirksamkeit von Ustekinumab in der Induktions- und Erhaltungsphase bei Patienten mit CU untersucht. Die Studie entspricht der Evidenzstufe Ib (nach §5 5. Kapitel, 1. Abschnitt, Absatz 6 der VerfO des G-BA).

Die herangezogenen klinischen Studien wurden anhand des Consolidated Standards of Reporting Trials (CONSORT)-Statements (Anhang 4-E) und der Cochrane-Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse (Anhang 4-F) auf die Planungs-, Durchführungs- und Auswertungsqualität überprüft.

Zur Beurteilung des Zusatznutzens von Upadacitinib wurden vom G-BA zwei Fragestellungen (Patienten mit oder ohne unzureichendes Ansprechen auf eine biologische Therapie: Bio-IR, Non-Bio-IR) formuliert. Im Rahmen des indirekten Vergleichs ist eine getrennte Beantwortung dieser Fragestellungen nicht möglich, da die Daten nicht für alle Endpunktdimensionen bzw. patientenrelevante Endpunkte, insbesondere nicht für die Verträglichkeit, getrennt nach Vorbehandlung vorliegen. Analysen zur Zielpopulation sind aber im Rahmen des indirekten Vergleichs adäquat, da ein gemeinsames Therapieziel für beide Populationen zu Grunde liegt. Da Ustekinumab für beide Fragestellungen als ZVT durch den G-BA definiert wurde und sich bei einer Behandlung unter Upadacitinib für die Erhaltungsphase in der dargestellten, deskriptiven Subgruppenanalyse, vergleichbare Ansprechraten für Bio-IR- und Non-Bio-IR-Patienten zeigen, werden beide Fragestellungen zu einer Fragestellung zusammengefasst und Ergebnisse für die Zielpopulation dargestellt.

Im vorliegenden Dossier wird jeweils, sofern nicht anders angegeben, der letzte verfügbare Analysezeitpunkt (primär Woche 52 in der Studie U-ACHIEVE bzw. Woche 44 in der Studie UNIFI) in der Erhaltungsphase indirekt verglichen und berichtet. Die Analysezeitpunkte sind

sowohl in Bezug auf die Studiendauer, als auch die Vergleichbarkeit der Zeitpunkte adäquat. Die untersuchten Endpunkte wurden mittels validierter und krankheitsspezifischer Instrumente erhoben und gemäß der im Studienprotokoll festgelegten Vorgehensweise ausgewertet (Intention-to-Treat [ITT]-Prinzip). Das Verzerrungspotenzial auf Endpunktebene für die Endpunkte Mortalität, Morbidität, Lebensqualität und Verträglichkeit wird als niedrig verzerrt bewertet.

Die Studien U-ACHIEVE und UNIFI sind hinsichtlich demografischer und krankheitsspezifischer Charakteristika hinreichend homogen (Abschnitt 4.3.2.1.2). Bei den Studien handelt es sich um randomisierte, kontrollierte, doppelt-verblindete Phase-III-Studien mit parallelen Behandlungsarmen. In beide Studien wurden erwachsene Patienten mit mittelschwerer bis schwerer aktiver CU eingeschlossen, welche auf eine vorangegangene Therapie unzureichend angesprochen haben, nicht mehr darauf ansprechen oder eine Unverträglichkeit oder Kontraindikation aufweisen. Die Studien untersuchen die Effekte in der Erhaltungsphase bei Patienten, die während der Induktionsstudie ein klinisches Ansprechen auf die entsprechende Behandlung gezeigt haben.

Aufgrund des konsequent umgesetzten Studiendesigns (Randomisierung, verdeckte Gruppenzuteilung, doppelte Verblindung, placebokontrolliert), sowie aufgrund der umfassenden und transparenten Auswertung und Berichterstattung, ist das Verzerrungspotenzial der Studien auf Studienebene zusammenfassend als niedrig zu bewerten. Die Aussagekraft der Nachweise für einen Zusatznutzen der Studien ist zusammenfassend als hoch anzusehen.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Tabelle 4-4: Zusammenfassung der Ergebnisse des adjustierten indirekten Vergleichs der Studien U-ACHIEVE, U-ACCOMPLISH und UNIFI und daraus resultierender Zusatznutzen

Endpunkt	**Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert**		**Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert**	**Wahrscheinlichkeit und Ausmaß des Zusatznutzens**
	Upadacitinib	**Ustekinumab**		
Klinische Remission (Mayo Gesamtscore)				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit klinischer Remission	RR: 10,74 [5,47; 21,09] p < 0,0001	RR: 2,91 [1,72; 4,94] p < 0,0001	RR: 3,69 [1,57; 8,68] p = 0,0028	
Erhaltungsphase				
Anteil der Patienten mit klinischer Remission (niedrigere Dosierung)	RR: 3,78 [2,28; 6,24] p < 0,0001	RR: 1,60 [1,16; 2,21] p = 0,0046	RR: 2,36 [1,30; 4,29] p = 0,0049	
Anteil der Patienten mit klinischer Remission (höhere Dosierung)	RR: 4,41 [2,70; 7,22] p < 0,0001	RR: 1,82 [1,33; 2,49] p = 0,0002	RR: 2,42 [1,35; 4,34] p = 0,0029	

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Klinische Remission (Adaptierter Mayo Score)				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit klinischer Remission	RR: 7,39 [4,42; 12,36] p < 0,0001	RR: 3,02 [1,87; 4,89] p < 0,0001	RR: 2,45 [1,21; 4,95] p = 0,0127	
Erhaltungsphase				
Anteil der Patienten mit klinischer Remission (niedrigere Dosierung)	RR: 3,52 [2,20; 5,65] p < 0,0001	RR: 1,61 [1,17; 2,21] p = 0,0034	RR: 2,19 [1,24; 3,87] p = 0,0070	
Anteil der Patienten mit klinischer Remission (höhere Dosierung)	RR: 4,30 [2,72; 6,81] p < 0,0001	RR: 1,73 [1,27; 2,37] p = 0,0005	RR: 2,48 [1,42; 4,32] p = 0,0013	
Klinisches Ansprechen (Mayo Gesamtscore)				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit klinischem Ansprechen	RR: 3,04 [2,50; 3,70] p < 0,0001	RR: 1,97 [1,64; 2,37] p < 0,0001	RR: 1,55 [1,19; 2,02] p = 0,0011	
Erhaltungsphase				
Anteil der Patienten mit klinischem Ansprechen (niedrigere Dosierung)	RR: 2,81 [2,03; 3,89] p < 0,0001	RR: 1,53 [1,26; 1,85] p < 0,0001	RR: 1,84 [1,26; 2,69] p = 0,0017	
Anteil der Patienten mit klinischem Ansprechen (höhere Dosierung)	RR: 3,40 [2,48; 4,66] p < 0,0001	RR: 1,59 [1,32; 1,93] p < 0,0001	RR: 2,13 [1,48; 3,08] p < 0,0001	
Mukosaheilung				
Induktionsphase (ergänzend)				Anhaltspunkt für einen beträchtlichen Zusatznutzen
Anteil der Patienten mit Mukosaheilung	RR: 5,50 [3,55; 8,52] p < 0,0001	RR: 2,08 [1,36; 3,17] p = 0,0007	RR: 2,65 [1,44; 4,86] p = 0,0017	
Erhaltungsphase				
Anteil der Patienten mit Mukosaheilung (niedrigere Dosierung)	RR: 2,91 [1,79; 4,73] p < 0,0001	RR: 1,61 [1,16; 2,23] p = 0,0043	RR: 1,81 [1,01; 3,24] p = 0,0475	
Anteil der Patienten mit Mukosaheilung (höhere Dosierung)	RR: 4,14 [2,61; 6,56] p < 0,0001	RR: 1,90 [1,39; 2,60] p < 0,0001	RR: 2,17 [1,25; 3,79] p = 0,0063	

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Stuhlfrequenz				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit Stuhlfrequenz ≤ 1	RR: 2,76 [2,25; 3,38] p < 0,0001	RR: 1,49 [1,25; 1,76] p < ,0001	RR: 1,86 [1,42; 2,42] p < 0,0001	
Erhaltungsphase				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit Stuhlfrequenz ≤ 1 (niedrigere Dosierung)	RR: 2,62 [1,88; 3,65] p < 0,0001	RR: 1,22 [1,04; 1,44] p = 0,0160	RR: 2,15 [1,48; 3,10] p < 0,0001	
Anteil der Patienten mit Stuhlfrequenz ≤ 1 (höhere Dosierung)	RR: 3,11 [2,26; 4,28] p < 0,0001	RR: 1,44 [1,24; 1,67] p < 0,0001	RR: 2,16 [1,52; 3,08] p < 0,0001	
Rektale Blutungen				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit Rektaler Blutung = 0	RR: 2,33 [1,98; 2,74] p < 0,0001	RR: 1,70 [1,44; 2,00] p < 0,0001	RR: 1,37 [1,09; 1,73] p = 0,0073	
Erhaltungsphase				
Anteil der Patienten mit Rektaler Blutung = 0 (niedrigere Dosierung)	RR: 2,64 [1,91; 3,64] p < 0,0001	RR: 1,38 [1,19; 1,60] p < 0,0001	RR: 1,91 [1,34; 2,72] p = 0,0004	
Anteil der Patienten mit Rektaler Blutung = 0 (höhere Dosierung)	RR: 3,24 [2,38; 4,42] p < 0,0001	RR: 1,37 [1,18; 1,59] p < 0,0001	RR: 2,37 [1,68; 3,34] p < 0,0001	
EQ-5D VAS				
Induktionsphase (ergänzend)				Zusatznutzen nicht belegt
Mittlere Veränderung der EQ-5D VAS	Hedges‘ g: 0,78 [0,64; 0,91] p < 0,0001	Hedges‘ g: 0,41 [0,25; 0,57] p < 0,0001	Hedges‘ g: 0,37 [0,16; 0,57] p = 0,0006	
Erhaltungsphase				
Mittlere Veränderung der EQ-5D VAS (niedrigere Dosierung)	Hedges‘ g : 0,48 [0,24; 0,71] p < 0,0001	Hedges‘ g : 0,28 [0,07; 0,50] p = 0,0085	Hedges‘ g : 0,19 [-0,12; 0,50] p = 0,2310	
Mittlere Veränderung der EQ-5D VAS (höhere Dosierung)	Hedges‘ g: 0,63 [0,40; 0,86] p < 0,0001	Hedges‘ g: 0,56 [0,35; 0,77] p < 0,0001	Hedges‘ g: 0,07 [-0,24; 0,38] p = 0,6594	

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Gesundheitsbezogene Lebensqualität: IBDQ				
Induktionsphase (ergänzend)				Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen
Mittlere Veränderung des IBDQ	Hedges‘ g: 1,01 [0,87; 1,15] p < 0,0001	Hedges‘ g: 0,60 [0,44; 0,76] p < 0,0001	Hedges‘ g: 0,41 [0,20; 0,62] p = 0,0002	
Erhaltungsphase				Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen
Mittlere Veränderung des IBDQ (niedrigere Dosierung)	Hedges‘ g: 0,73 [0,50; 0,97] p < 0,0001	Hedges‘ g: 0,35 [0,14; 0,57] p = 0,0011	Hedges‘ g: 0,38 [0,06; 0,69] p = 0,0195	
Mittlere Veränderung des IBDQ (höhere Dosierung)	Hedges‘ g: 0,94 [0,71; 1,18] p < 0,0001	Hedges‘ g: 0,57 [0,35; 0,78] p < 0,0001	Hedges‘ g: 0,38 [0,06; 0,70] p = 0,0205	
Gesundheitsbezogene Lebensqualität: SF-36 PCS				
Induktionsphase (ergänzend)				Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen
Mittlere Veränderung des SF-36 PCS	Hedges‘ g: 0,82 [0,68; 0,96] p < 0,0001	Hedges‘ g: 0,49 [0,34; 0,65] p < 0,0001	Hedges‘ g: 0,33 [0,12; 0,54] p = 0,0023	
Erhaltungsphase				
Mittlere Veränderung des SF-36 PCS (niedrigere Dosierung)	Hedges‘ g: 0,69 [0,45; 0,92] p < 0,0001	Hedges‘ g: 0,19 [-0,02; 0,40] p = 0,0763	Hedges‘ g: 0,49 [0,18; 0,81] p = 0,0021	
Mittlere Veränderung des SF-36 PCS (höhere Dosierung)	Hedges‘ g: 0,82 [0,58; 1,05] p < 0,0001	Hedges‘ g: 0,49 [0,28; 0,71] p < 0,0001	Hedges‘ g: 0,33 [0,01; 0,64] p = 0,0432	
Gesundheitsbezogene Lebensqualität: SF-36 MCS				
Induktionsphase (ergänzend)				Zusatznutzen nicht belegt
Mittlere Veränderung des SF-36 MCS	Hedges‘ g : 0,61 [0,47; 0,74] p < 0,0001	Hedges‘ g : 0,29 [0,14; 0,45] p = 0,0003	Hedges‘ g : 0,32 [0,11; 0,52] p = 0,0028	
Erhaltungsphase				
Mittlere Veränderung des SF-36 MCS (niedrigere Dosierung)	Hedges‘ g : 0,38 [0,15; 0,61] p = 0,0013	Hedges‘ g : 0,29 [0,08; 0,50] p = 0,0066	Hedges‘ g : 0,08 [-0,23; 0,39] p = 0,6037	
Mittlere Veränderung des SF-36 MCS (höhere Dosierung)	Hedges‘ g : 0,49 [0,26; 0,72] p < 0,0001	Hedges‘ g : 0,28 [0,07; 0,49] p = 0,0096	Hedges‘ g : 0,21 [-0,10; 0,52] p = 0,1797	

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Gesamtraten UE				
Induktionsphase (ergänzend)				Zusatznutzen/ Schaden nicht belegt
Alle UE	RR: 1,06 [0,95; 1,19] p = 0,3099	RR: 1,06 [0,90; 1,24] p = 0,5011	RR: 1,01 [0,83; 1,22] p = 0,9517	
SUE	RR: 0,55 [0,31; 0,98] p = 0,0417	RR: 0,50 [0,25; 1,01] p = 0,0536	RR: 1,10 [0,44; 2,74] p = 0,8417	
UE, die zum Therapieabbruch führten	RR: 0,36 [0,20; 0,63] p = 0,0005	RR: k. A. [k. A.; k. A.] p = k. A.	RR: k. A. [k. A.; k. A.] p = k. A.	
Erhaltungsphase				
Alle UE (niedrigere Dosierung)	RR: 1,02 [0,92; 1,12] p = 0,7365	RR: 0,88 [0,77; 1,00] p = 0,0415	RR: 1,16 [0,99; 1,36] p = 0,0692	
Alle UE (höhere Dosierung)	RR: 1,04 [0,94; 1,14] p = 0,4303	RR: 0,98 [0,88; 1,09] p = 0,7199	RR: 1,06 [0,92; 1,23] p = 0,4319	
SUE (niedrigere Dosierung)	RR: 0,89 [0,51; 1,57] p = 0,6996	RR: 0,78 [0,39; 1,55] p = 0,4764	RR: 1,15 [0,47; 2,81] p = 0,7588	
SUE (höhere Dosierung)	RR: 0,89 [0,51; 1,57] p = 0,6894	RR: 0,88 [0,45; 1,70] p = 0,6984	RR: 1,02 [0,43; 2,43] p = 0,9718	
UE, die zum Therapieabbruch führten (niedrigere Dosierung)	RR: 0,39 [0,19; 0,80] p = 0,0099	RR: 0,46 [0,21; 0,98] p = 0,0434	RR: 0,86 [0,30; 2,42] p = 0,7698	
UE, die zum Therapieabbruch führten (höhere Dosierung)	RR: 0,70 [0,39; 1,25] p = 0,2331	RR: 0,25 [0,10; 0,65] p = 0,0044	RR: 2,83 [0,92; 8,66] p = 0,0688	

Für binäre Endpunkte werden das RR mit zugehörigem 95 %-KI und p-Wert, sowohl für die Einzelstudien als auch für den indirekten Vergleich angegeben. Für kontinuierliche Endpunkte werden das Hedges‘ g mit zugehörigem 95 %-KI und p-Wert, sowohl für die Einzelstudien als auch für den indirekten Vergleich angegeben.

Bei den Ergebnissen des indirekten Vergleichs zur Morbidität und gesundheitsbezogenen Lebensqualität stellt ein RR größer Eins bzw. ein Hedges‘ g größer Null einen Vorteil von Upadacitinib dar. Bei den Ergebnissen des indirekten Vergleichs zur Verträglichkeit stellt ein RR kleiner Eins einen Vorteil von Upadacitinib dar.

Bezogen auf die Einzelstudien, stellen für Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität ein RR größer Eins bzw. ein Hedges‘ g größer Null einen Vorteil von Upadacitinib beziehungsweise Ustekinumab gegenüber Placebo dar. Für Endpunkte zur Verträglichkeit, stellt für die Einzelstudien ein RR kleiner Eins einen Vorteil von Upadacitinib beziehungsweise Ustekinumab gegenüber

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Placebo dar. Niedrigere Dosierung: Adjustierter indirekter Vergleich der Behandlungsarme Upadacitinib 15 mg einmal täglich und Ustekinumab 90 mg alle 12 Wochen. Höhere Dosierung: Adjustierter indirekter Vergleich der Behandlungsarme Upadacitinib 30 mg einmal täglich und Ustekinumab 90 mg alle 8 Wochen. EQ-5D: EuroQol-5 dimensions, IBDQ: Inflammatory Bowel Disease Questionnaire; k. A.: Keine Angabe; KI: Konfidenzintervall; MCS: Mental Health Component Summary Score; PCS: Physical Health Component Summary Score; RR: Relatives Risiko; SF-36: Short Form – 36; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala				

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Unter der Behandlung mit Upadacitinib ergibt sich eine signifikant höhere Wahrscheinlichkeit, das patientenrelevante Haupttherapieziel einer raschen klinischen Remission sowie deren langfristiger Erhalt, und somit das vollständige Abklingen der mukosalen Entzündung zu erreichen als unter einer Behandlung mit der ZVT Ustekinumab. Zudem erfahren unter einer Behandlung mit Upadacitinib mehr Patienten eine Verbesserung der gesundheitsbezogenen Lebensqualität.

Mit einer Upadacitinib-Therapie kann somit eine umfassende Kontrolle der Krankheitsaktivität (klinische Remission und Mukosaheilung) bei Patienten mit CU erreicht werden. Dies zeigt sich auch anhand der signifikant höheren klinischen Ansprechraten als unter der ZVT Ustekinumab. Auch in Bezug auf die für den Patienten stark belastende Symptomatik der Stuhlfrequenz und rektalen Blutungen zeigen sich signifikante Vorteile für eine Behandlung mit Upadacitinib gegenüber Ustekinumab. Damit kann mit Upadacitinib eine langfristige Kontrolle der Krankheitsaktivität und der stark belastenden Symptomatik erreicht werden. Aufgrund seiner überlegenen Wirksamkeit gegenüber Ustekinumab und eines günstigen Nutzen-Risiko-Profils bietet Upadacitinib daher einen erheblichen Mehrwert für die Behandlung von Patienten mit CU.

Das Verträglichkeitsprofil von Upadacitinib ist vergleichbar mit dem von Ustekinumab und entspricht dem bekannten und konsistenten Sicherheitsprofil von Upadacitinib in den bisher zugelassenen Anwendungsgebieten. Upadacitinib ist daher wegen seines positiven Wirksamkeits- und Verträglichkeitsprofils eine Bereicherung der Therapielandschaft in Deutschland.

Insgesamt kann ein **Anhaltspunkt** für einen **erheblichen Zusatznutzen** für eine Upadacitinib-Therapie bei vorbehandelten Patienten mit CU im Vergleich zu Ustekinumab abgeleitet werden.

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).

Die Benennung der Vergleichstherapie in Modul 4 muss zur Auswahl der zweckmäßigen Vergleichstherapie im zugehörigen Modul 3 konsistent sein.

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

Ziel des vorliegenden Dossiers ist die Bewertung des medizinischen Nutzens und medizinischen Zusatznutzens von Upadacitinib zur Behandlung der mittelschweren bis schweren aktiven CU bei erwachsenen Patienten, die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben (1).

Die Bewertung erfolgt gegenüber der vom G-BA festgelegten ZVT anhand patientenrelevanter Endpunkte aus den Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit. Basis dieser Nutzenbewertung ist ein adjustierter indirekter Vergleich nach Bucher, in welchem Upadacitinib mit der ZVT Ustekinumab über den Brückenkomparator Placebo verglichen wird.

Patientenpopulation

Upadacitinib wird angewendet zur Behandlung der mittelschweren bis schweren aktiven Colitis ulcerosa bei erwachsenen Patienten, die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben.

Basis dieser Nutzenbewertung ist ein adjustierter indirekter Vergleich nach Bucher, in welchem Upadacitinib mit der ZVT Ustekinumab über den Brückenkomparator Placebo verglichen wird. Zur Beurteilung des Zusatznutzens von Upadacitinib wurden vom G-BA zwei Fragestellungen (Patienten mit oder ohne unzureichendes Ansprechen auf eine biologische Therapie: Bio-IR, Non-Bio-IR) formuliert. Im Rahmen des indirekten Vergleichs ist eine getrennte Beantwortung dieser Fragestellungen nicht möglich, da die Daten nicht für alle Endpunktdimensionen bzw. patientenrelevante Endpunkte, insbesondere nicht für die Verträglichkeit, getrennt nach Vorbehandlung vorliegen. Analysen zur Zielpopulation sind aber im Rahmen des indirekten Vergleichs inhaltlich und medizinisch adäquat, da ein gemeinsames Therapieziel zu Grunde liegt. Da Ustekinumab für beide Fragestellungen als ZVT durch den G-BA definiert wurde und sich bei einer Behandlung unter Upadacitinib für die Erhaltungsphase in der dargestellten, deskriptiven Subgruppenanalyse, vergleichbare Ansprechraten für Bio-IR- und Non-Bio-IR-Patienten zeigen, werden beide Fragestellungen zu einer Fragestellung zusammengefasst und Ergebnisse für die Zielpopulation dargestellt.

Intervention

Die zu bewertende Intervention ist Upadacitinib. Bei Upadacitinib handelt es sich um einen selektiven und reversiblen Januskinase (JAK)-Inhibitor. Die empfohlene Dosis beträgt 45 mg einmal täglich in der Induktionsphase und 15 mg bzw. 30 mg einmal täglich in der Erhaltungsphase gemäß Fachinformation (1).

Vergleichstherapie

Für das vorliegende Dossier ergeben sich für Upadacitinib für die zu betrachtenden Fragestellungen bzw. Populationen folgende ZVT (Modul 3, Abschnitt 3.1.1):

- Population A:
  Ein TNF-Inhibitor (Adalimumab oder Infliximab oder Golimumab) oder Vedolizumab oder Ustekinumab

- Population B:
  Ein Wechsel der Therapie auf Vedolizumab oder Tofacitinib oder Ustekinumab oder einen TNF-Inhibitor (Adalimumab oder Infliximab oder Golimumab), jeweils unter Berücksichtigung der Zulassung und der Vortherapie(n) (2).

Endpunkte

Die Bewertung des medizinischen Nutzens bzw. des medizinischen Zusatznutzens von Upadacitinib erfolgt anhand der Kategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit auf Grundlage patientenrelevanter Endpunkte. Eine umfassende Auflistung der im Rahmen des Dossiers betrachteten patientenrelevanten Endpunkte findet sich in Abschnitt 4.2.5.2.

Studientypen

Die höchste Ergebnissicherheit bieten RCTs, daher wird der medizinische Nutzen und der medizinische Zusatznutzen in diesem Dossier auf Basis von RCTs bewertet. Sowohl für die

Bewertung von Upadacitinib als auch die ZVT Ustekinumab liegen in der Indikation CU placebokontrollierte RCTs vor. Basierend auf diesen Studien wurde ein adjustierter, indirekter Vergleich nach Bucher durchgeführt.

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/ einer Studienergebnisdatenbank erfolgen, während ein Kongressabstrakt 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. Erstellen Sie dabei für unterschiedliche Themen der Recherche (z. B. unterschiedliche Fragestellungen) jeweils eine separate Übersicht.

Zur Identifikation relevanter Studien für die Nutzenbewertung wurde eine Suche für das zu bewertende Arzneimittel nach direkt mit der ZVT vergleichenden RCTs durchgeführt. Da keine direkt vergleichende RCT identifiziert wurde (siehe Abschnitt 4.3.1.1.5), wurden Suchen nach RCT für einen adjustierten indirekten Vergleich zur ZVT Ustekinumab durchgeführt. Studien, die für einen adjustierten indirekten Vergleich geeignet sind, wurden anhand von Ein- und Ausschlusskriterien identifiziert, welche Tabelle 4-6 bzw. Tabelle 4-7 entnommen werden können.

Tabelle 4-5: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien mit dem zu bewertenden Arzneimittel – RCT

Kriterium	Einschluss		Ausschluss		Begründung
Population	Erwachsene Patienten mit mittelschwerer bis schwerer aktiver Colitis ulcerosa die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben.	E1	Andere, nicht dem zugelassenen Anwendungsgebiet entsprechende, Patientenpopulation	A1	Population des zugelassenen Anwendungs-gebiets

Kriterium	Einschluss		Ausschluss		Begründung
Intervention(en)	Behandlung mit Upadacitinib gemäß Fachinformation	E2	Andere Therapien	A2	Zu bewertendes Arzneimittel
Vergleichs-therapie	Tumornekrosefaktor (TNF)-Inhibitor (Adalimumab, Golimumab oder Infliximab) oder Ustekinumab oder Vedolizumab oder Tofacitinib	E3	Andere Therapien oder Behandlung mit Vergleichstherapien abweichend von der Fachinformation	A3	Therapien der ZVT gemäß G-BA-Beratung
Endpunkte	Mindestens ein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	E4	Kein patienten-relevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	A4	Nutzendimensionen nach § 5 Absatz 2 des 5. Kapitels der VerfO
Studientyp	RCT	E5	Keine RCT	A5	RCTs bieten aufgrund ihres Designs die höchste Ergebnissicherheit. Für den Fall, dass keine RCTs identifiziert werden, wird die beste verfügbare Evidenz ermittelt.
Studiendauer	Mindestens 52 Wochen	E6	Weniger als 52 Wochen	A6	Mindestdauer gemäß G-BA-Beratung
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügt[a]	E7	• Kein Studienbericht oder keine Vollpublikation verfügbar sowie auch keine publizierten Ergebnisse auf der Webseite offizieller Studienregister[b] • Mehrfachpubli-kation ohne relevante Zusatz-informationen[c]	A7	Ausreichende Datenverfügbarkeit ist zur Bewertung von Ergebnissen erforderlich.

Kriterium	Einschluss		Ausschluss		Begründung
a: Studien, deren Ergebnisse auf der Webseite offizieller Studienregister publiziert wurden (z. B. clinicaltrials.gov), können ebenfalls berücksichtigt werden. b: Keine publizierten Ergebnisse auf der Webseite offizieller Studienregister, die den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügen c: Jeglicher Publikationstyp in welchem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden. Dies gilt auch für in der bibliografischen Literaturrecherche identifizierte Studienregistereinträge, welche ebenfalls in der Studienregistersuche identifiziert werden. CONSORT: Consolidated Standards of Reporting Trials; G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TNF: Tumornekrosefaktor; TREND: Transparent Reporting of Evaluations with Non-randomized Design; VerfO: Verfahrensordnung (des G-BA); ZVT: Zweckmäßige Vergleichstherapie					

Ein- und Ausschlusskriterien für die Suche nach Studien mit Upadacitinib im Vergleich zu Placebo für adjustierte indirekte Vergleiche sind in Tabelle 4-6 aufgeführt.

Tabelle 4-6: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für den indirekten Vergleich (RCT) mit dem zu bewertenden Arzneimittel

Kriterium	Einschluss		Ausschluss		Begründung
Population	Erwachsene Patienten mit mittelschwerer bis schwerer aktiver Colitis ulcerosa die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben.	E1	Andere, nicht dem zugelassenen Anwendungsgebiet entsprechende, Patientenpopulation	A1	Population des zugelassenen Anwendungs-gebiets
Intervention(en)	Behandlung mit Upadacitinib gemäß Fachinformation	E2	Andere Therapien	A2	Zu bewertendes Arzneimittel
Vergleichstherapie	Placebo	E3	Andere Therapien oder Behandlung mit Vergleichstherapien abweichend von der Fachinformation	A3	Brückenkomparator für adjustierten indirekten Vergleich
Endpunkte	Mindestens ein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	E4	Kein patienten-relevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	A4	Nutzendimensionen nach § 5 Absatz 2 des 5. Kapitels der VerfO

Kriterium	Einschluss		Ausschluss		Begründung
Studientyp	RCT	E5	Keine RCT	A5	RCTs bieten aufgrund ihres Designs die höchste Ergebnissicherheit. Für den Fall, dass keine RCTs identifiziert werden, wird die beste verfügbare Evidenz ermittelt.
Studiendauer	Mindestens 52 Wochen	E6	Weniger als 52 Wochen	A6	Mindestdauer gemäß G-BA-Beratung
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügt[a]	E7	• Kein Studienbericht oder keine Vollpublikation verfügbar sowie auch keine publizierten Ergebnisse auf der Webseite offizieller Studienregister[b] • Mehrfachpublikation ohne relevante Zusatzinformationen[c]	A7	Ausreichende Datenverfügbarkeit ist zur Bewertung von Ergebnissen erforderlich.

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

b: Keine publizierten Ergebnisse auf der Webseite offizieller Studienregister, die den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügen

c: Jeglicher Publikationstyp in welchem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden. Dies gilt auch für in der bibliografischen Literaturrecherche identifizierte Studienregistereinträge, welche ebenfalls in der Studienregistersuche identifiziert werden.

CONSORT: Consolidated Standards of Reporting Trials; G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Non-randomized Design; VerfO: Verfahrensordnung (des G-BA)

Ein- und Ausschlusskriterien für die Suche nach der ZVT Ustekinumab für indirekte Vergleiche (RCT) sind in Tabelle 4-7 zusammengefasst:

Tabelle 4-7: Einschluss- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für den indirekten Vergleich (RCT) mit der ZVT (Ustekinumab)

Kriterium	Einschluss		Ausschluss		Begründung
Population	Erwachsene Patienten mit mittelschwerer bis schwerer aktiver Colitis ulcerosa die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben.	E1	Andere, nicht dem zugelassenen Anwendungsgebiet entsprechende, Patientenpopulation	A1	Population des zugelassenen Anwendungsgebiets
Intervention(en)	Behandlung mit Ustekinumab gemäß Fachinformation	E2	Andere Therapien	A2	ZVT (Komparator für adjustierten indirekten Vergleich)
Vergleichstherapie	Placebo	E3	Andere Therapien oder Behandlung mit Vergleichstherapien abweichend von der Fachinformation	A3	Brückenkomparator für adjustierten indirekten Vergleich
Endpunkte	Mindestens ein patientenrelevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	E4	Kein patienten-relevanter Endpunkt, der zur Bewertung der Mortalität, Morbidität, Lebensqualität oder der Verträglichkeit verwendet werden kann	A4	Nutzendimensionen nach § 5 Absatz 2 des 5. Kapitels der VerfO
Studientyp	RCT	E5	Keine RCT	A5	RCTs bieten aufgrund ihres Designs die höchste Ergebnissicherheit. Für den Fall, dass keine RCTs identifiziert werden, wird die beste verfügbare Evidenz ermittelt.
Studiendauer	Mindestens 52 Wochen	E6	Weniger als 52 Wochen	A6	Mindestdauer gemäß G-BA-Beratung

Kriterium	Einschluss		Ausschluss		Begründung
Publikationstyp	Vollpublikation oder Studienbericht verfügbar, der den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügt[a]	E7	• Kein Studienbericht oder keine Vollpublikation verfügbar sowie auch keine publizierten Ergebnisse auf der Webseite offizieller Studienregister[b] • Mehrfachpublikation ohne relevante Zusatzinformationen[c]	A7	Ausreichende Datenverfügbarkeit ist zur Bewertung von Ergebnissen erforderlich.

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

b: Keine publizierten Ergebnisse auf der Webseite offizieller Studienregister, die den Kriterien des CONSORT-Statements, TREND- bzw. STROBE-Statements genügen

c: Jeglicher Publikationstyp in welchem keine (zu schon identifizierten Informationsquellen) zusätzlichen Informationen dargestellt werden. Dies gilt auch für in der bibliografischen Literaturrecherche identifizierte Studienregistereinträge, welche ebenfalls in der Studienregistersuche identifiziert werden.

CONSORT: Consolidated Standards of Reporting Trials; G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Non-randomized Design; VerfO: Verfahrensordnung (des G-BA); ZVT: Zweckmäßige Vergleichstherapie

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 (inklusive „in-process & other non-indexed citations) 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.

Die bibliografische Literaturrecherche wird in den Datenbanken MEDLINE, EMBASE und Cochrane durchgeführt, wobei für jede einzelne Datenbank eine sequenzielle, individuell adaptierte Suchstrategie entwickelt und separat angewandt wird. Für die Suche wird die Plattform OVID verwendet.

Die Ergebnisse der systematischen bibliografischen Literaturrecherche nach RCT mit dem zu bewertenden Arzneimittel sind im Abschnitt 4.3.1.1.2 dargestellt. Im Anschluss daran wurden weitere systematische Literaturrecherchen für einen indirekten Vergleich zwischen Upadacitinib und der ZVT Ustekinumab über den Brückenkomparator Placebo durchgeführt. Die Ergebnisse sind im Abschnitt 4.3.2.1 dargestellt.

Die verwendeten Suchstrategien und das Datum der Suche sind in Anhang 4-A beschrieben. Die im Volltext ausgeschlossenen Publikationen sind in Anhang 4-C aufgeführt.

4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken

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

Eine Suche in Studienregistern/ Studienergebnisdatenbanken 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 sowie Studienergebnisdatenbanken 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/ Studienergebnisdatenbanken clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal), Suchportal der WHO, Clinical Data Suchportal der European Medicines Agency (https://clinicaldata.ema.europa.eu) sowie dem Arzneimittel-Informationssystem (AMIS, https://www.pharmnet-bund.de/dynamic/de/arzneimittel-informationssystem/index.html) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern / Studienergebnisdatenbanken (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern/ Studienergebnisdatenbanken anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister/ Studienergebnisdatenbank einzeln und mit einer für das jeweilige Studienregister/ Studienergebnisdatenbank adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Für Clinical Data (Suchportal der European Medicines Agency) und das Arzneimittel-Informationssystem (AMIS) genügt hingegen die Suche nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und

Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern/ Studienergebnisdatenbanken 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.

Die Suche wurde in den Portalen clinicaltrials.gov (www.clinicaltrials.gov), ICTRP Search Portal, Suchportal der WHO (http://apps.who.int/trialsearch/) und im EU-CTR (www.clinicaltrialsregister.eu/ctr-search/search) durchgeführt. Es wird nach abgeschlossenen, laufenden und abgebrochenen Studien gesucht. Zudem wird im Clinical Data Suchportal der EMA (https://clinicaldata.ema.europa.eu) im AMIce (https://www.bfarm.de/DE/Arzneimittel/Arzneimittelinformationen/Arzneimittel-recherchieren/AMIce/Datenbankinformation-AMIce-Arzneimittel/_node.html) nach Ergebnisberichten zu denjenigen Studien gesucht, die zuvor schon als relevante Studien identifiziert wurden.

Die Ergebnisse der Studienregistersuche nach RCTs mit dem zu bewertenden Arzneimittel sind in Abschnitt 4.3.1.1.3 dargestellt. Die Ergebnisse der Studienregistersuche nach RCTs, welche einen direkten Vergleich des zu bewertenden Arzneimittels mit der ZVT erlauben, werden im Abschnitt 4.3.1.1.3 dargestellt. Im Anschluss wurde eine weitere Suche in den Studienregistern für einen indirekten Vergleich durchgeführt. Hierbei wurde nach Studien zu Upadacitinib und dem Vergleichswirkstoff Ustekinumab jeweils gegenüber Placebo gesucht. Die Ergebnisse sind im Abschnitt 4.3.2.1.1 dargestellt.

Die verwendeten Suchbegriffe und das Datum der Suchen sind im Anhang 4-B beschrieben. Die ausgeschlossenen Studien sind im Anhang 4-D aufgelistet.

4.2.3.4 Suche auf der Internetseite des G-BA

Die Internetseite des G-BA ist grundsätzlich zu durchsuchen, um sicherzustellen, dass alle vorliegenden Daten zu Studienmethodik und –ergebnissen von relevanten Studien in die Bewertung einfließen.

Auf der Internetseite des G-BA werden Dokumente zur frühen Nutzenbewertung nach §35a SGB V veröffentlicht. Diese enthalten teilweise anderweitig nicht veröffentlichte Daten zu Studienmethodik und –ergebnissen[1]. Solche Daten sind dabei insbesondere in den Modulen 4 der Dossiers pharmazeutischer Unternehmer, in IQWiG-Nutzenbewertungen sowie dem

[1] Köhler M, Haag S, Biester K, Brockhaus AC, McGauran N, Grouven U, Kölsch H, Seay U, Hörn H, Moritz G, Staeck K, Wieseler B. Information on new drugs at market entry: retrospective analysis of health technology assessment reports, journal publications, and registry reports. BMJ 2015;350:h796

Beschluss des G-BA einschließlich der Tragenden Gründe und der Zusammenfassenden Dokumentation zu erwarten.

Die Suche auf der Internetseite des G-BA 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 auf der G-BA Internetseite immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird. Die Suche ist dann sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie durchzuführen. Es genügt die Suche nach Einträgen zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

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

Beschreiben Sie nachfolgend das Vorgehen für die Suche. Benennen Sie die Wirkstoffe und die auf der Internetseite des G-BA genannten zugehörigen Vorgangsnummern, zu denen Sie eine Suche durchgeführt haben.

Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Auf der Internetseite des G-BA wird eine Suche nach Einträgen zu Studien durchgeführt, die bereits anderweitig identifiziert wurden.

Die Ergebnisse der Suche auf der Internetseite des G-BA sind im Abschnitt 4.3.2.1.1.2 dargestellt.

4.2.3.5 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2, 4.2.3.3 und 4.2.3.4 beschriebenen Rechercheschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Selektion der Treffer aus der Literaturrecherche, der Suche in Studienregistern/ Studienergebnisdatenbanken und der Suche auf der Internetseite des G-BA wird von zwei Personen unabhängig voneinander durchgeführt. Die Studien werden anhand der in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien auf ihre Relevanz geprüft. Diskrepanzen hinsichtlich der Bewertung wurden durch Diskussion gelöst.

4.2.4 Bewertung der Aussagekraft der Nachweise

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

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

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

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

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

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

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

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

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen.

Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

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

Die für die Fragestellung relevanten Studien wurden hinsichtlich ihrer Aussagekraft bewertet. Dafür erfolgte eine Einschätzung des Verzerrungspotenzials entsprechend den Vorgaben der Dossiervorlage. Die Basis hierfür bildeten die relevanten Informationen aus klinischen Studienberichten, Studienprotokollen und statistischen Analyseplänen (falls vorhanden) oder aus Volltextpublikationen der jeweiligen Studien.

Bewertung des Verzerrungspotenzials

Die Bewertung des Verzerrungspotenzials erfolgte gemäß Vorgabe zunächst auf Studienebene (also endpunktübergreifend) und anschließend getrennt für jeden Endpunkt, gegebenenfalls für jede Operationalisierung.

Gemäß Dossiervorlage wurden für die Bewertung des Verzerrungspotenzials auf Studienebene folgende Aspekte untersucht:

- 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

Unter Berücksichtigung der einzelnen verzerrenden Aspekte auf Studienebene wurde für alle RCT das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. Von einem niedrigen Verzerrungspotenzial wurde ausgegangen, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Zur Bewertung des Verzerrungspotenzials von RCT auf Endpunktebene werden – im Einklang mit der Dossiervorlage – folgende Aspekte verwendet:

- Verblindung der Endpunkterheber

- Umsetzung des ITT-Prinzips
- Ergebnisgesteuerte Berichterstattung
- Sonstige Aspekte

Unter Berücksichtigung der einzelnen verzerrenden Aspekte auf Endpunktebene wurde zusätzlich für alle Endpunkte das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. Ein „hohes“ Verzerrungspotenzial auf Studienebene führt grundsätzlich für alle Endpunkte der betreffenden Studie ebenfalls zu einer Einstufung des Verzerrungspotenzials auf Endpunktebene als „hoch“. Einzelne Abweichungen von diesem Vorgehen wurden entsprechend begründet.

Die Vorgehensweise bei der Bewertung des Verzerrungspotenzials jedes einzelnen Aspekts auf Endpunkt- und Studienebene richtete sich außerdem nach den Hinweisen der Bewertungsbögen in Anhang 4-F.

Die Ergebnisse für relevante Studien oder Endpunkte mit „hohem“ Verzerrungspotenzial wurden gleichwohl im Dossier dargestellt. Die Klassifizierung des Verzerrungspotenzials dient als Hilfe zur Beurteilung der Aussagekraft und Ergebnissicherheit der Studien.

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)[2]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[3] bzw. STROBE-Statements[4] folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

[2] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.
[3] 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.
[4] 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.

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.

Generell erfolgte die Beschreibung von Design und Methodik der eingeschlossenen Studien unter Berücksichtigung des CONSORT-Statements (Items 2b bis 14).

Entsprechend wurden gemäß Dossiervorlage für randomisierte Studien die Items 2b bis 14 (CONSORT) in Anhang 4-E ausgefüllt und Flow-Charts erstellt.

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.

Sofern zur Berechnung von Ergebnissen von Standardverfahren und –software abgewichen wird (insbesondere beim Einsatz spezieller Software oder individueller Programmierung), sind die Berechnungsschritte und ggf. verwendete Software explizit abzubilden. Insbesondere der Programmcode ist in lesbarer Form anzugeben.

Patientencharakteristika

Zur Charakterisierung der in den Studien eingeschlossenen Patientenpopulationen werden, sofern verfügbar, folgende demografische Merkmale dargestellt:

- Alter (Mittelwert mit Standardabweichung)
- Geschlecht (weiblich, männlich)
- Abstammung (weiß, schwarz oder afroamerikanisch, asiatisch, andere)

- Geografische Region (USA/Puerto Rico/Kanada, andere)
- Body Mass Index (BMI) (Mittelwert mit Standardabweichung)
- Gewicht (Mittelwert mit Standardabweichung)

Als krankheitsbeschreibende Faktoren werden, sofern verfügbar, folgende Charakteristika dargestellt:

- Dauer der Erkrankung seit Diagnose (Mittelwert mit Standardabweichung)
- Ausdehnung der Erkrankung (Proktitis, Linksseitencolitis, ausgedehnte Colitis)
- Mayo Gesamtscore (Mittelwert mit Standardabweichung, kategorielle Darstellung)
- Adaptierter Mayo Score (Mittelwert mit Standardabweichung, kategorielle Darstellung)
- Hoch sensitives C-reaktives Protein (high sensitivity C-Reactive Protein, hsCRP) (Mittelwert mit Standardabweichung bzw. Median mit Interquartile Range (IQR))
- Fäkales Calprotectin (Mittelwert mit Standardabweichung bzw. Median mit IQR)

Zu Baseline werden zur medizinischen Vorgeschichte folgende Charakteristika dargestellt:

- Intoleranz oder unzureichendes Ansprechen auf eine vorherige biologische Therapie (Bio-IR-Status) (Bio-IR, Non-Bio-IR)
- Kortikosteroideinnahme zu Baseline (ja, nein)
- Aminosalicylate zu Baseline (ja, nein)
- Immunsuppressiva zu Baseline (ja, nein)

Darüber hinaus wird die Patientendisposition mit Angabe der Gründe für eine vorzeitige Beendigung bzw. einen vorzeitigen Abbruch der Studie oder der Behandlung beschrieben. Weiterhin wird die Dauer der Behandlung betrachtet.

Patientenrelevante Endpunkte

Zur Bewertung des Ausmaßes des medizinischen Zusatznutzens wurden gemäß § 3 Absatz 1 und 2, 5. Kapitel der VerfO des G-BA Ergebnisse bezüglich der patientenrelevanten Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit herangezogen. Für das vorliegende Dossier werden Endpunkte dargestellt, für die im Rahmen eines indirekten Vergleiches mit der ZVT Ustekinumab Daten zur Verfügung stehen. Die dazu operationalisierten patientenrelevanten Endpunkte sind im Folgenden in Tabelle 4-8 dargestellt.

Tabelle 4-8: Nutzendimensionen und patientenrelevante Endpunkte zur Bewertung des Ausmaßes des medizinischen Zusatznutzens von Upadacitinib

Nutzendimension	Patientenrelevante Endpunkte
Mortalität	• Todesfälle (Betrachtung im Rahmen der UE)
Morbidität	• Klinische Remission (nach Mayo Gesamtscore und Adaptiertem Mayo Score) • Klinisches Ansprechen (nach Mayo Gesamtscore) • Mukosaheilung (nach endoskopischem Subscore des Mayo Gesamtscores und Geboes Score) • Stuhlfrequenz (SFS) (Subscore des Mayo Gesamtscores) • Rektale Blutung (RBS) (Subscore des Mayo Gesamtscores) • EuroQol-5 dimensions visuelle Analogskala (EQ-5D VAS)
Gesundheitsbezogene Lebensqualität	• Inflammatory Bowel Disease Questionnaire (IBDQ) • Short form – 36 (SF-36) Physical Health Component Summary Score (PCS) und Mental Health Component Summary Score (MCS)
Verträglichkeit	• Unerwünschte Ereignisse (UE) für die im Rahmen eines indirekten Vergleiches Daten vorliegen o Alle UE o Schwerwiegende unerwünschte Ereignisse (SUE) o UE, die zum Therapieabbruch führten • Zusätzlich ergänzend dargestellte UE für Upadacitinib Erhaltungsphase o Schwere UE o UE, die zum Tod führten o UE von speziellem Interesse (UESI)

EQ-5D: EuroQol-5 dimensions; IBDQ: Inflammatory Bowel Disease Questionnaire ; MCS: Mental Health Component Summary Score; PCS: Physical Health Component Summary Score; RBS: Rektale Blutung; SF-36: Short Form – 36; SFS: Stuhlfrequenz; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; UESI: Unerwünschtes Ereignis von speziellem Interesse; VAS: Visuelle Analogskala

Mortalität

Anzahl der Todesfälle

Bei der CU handelt es sich um eine chronische Erkrankung mit Mortalitätsraten, die sich nicht wesentlich von derjenigen in der Gesamtbevölkerung unterscheiden (3). Daher ist die Betrachtung der Mortalität als prädefinierter Studienendpunkt in dieser Indikation nicht üblich. Die im vorliegenden Dossier präsentierten Ergebnisse zur Mortalität werden somit im Rahmen der Darstellung der unerwünschten Ereignisse (UE) berichtet.

Operationalisierung

Die im vorliegenden Dossier präsentierten Ergebnisse zur Mortalität beruhen auf der Anzahl aller Todesfälle, die während der Beobachtungsdauer aufgetreten sind. Es erfolgt zusätzlich eine narrative Beschreibung der Todesfälle.

Morbidität

Zur Bewertung der Morbidität werden patientenrelevante Endpunkte zur Messung der Krankheitsaktivität (klinische Remission nach Mayo Gesamtscore, klinisches Ansprechen nach Mayo Gesamtscore und nach Adaptiertem Mayo Score, Mukosaheilung nach endoskopischem Subscore des Mayo Gesamtscore und Geboes Score, Stuhlfrequenz [Stool Frequency Score, SFS], Rektale Blutung [Rectal Bleeding Score, RBS]) und der allgemeine Gesundheitszustand [EuroQol-5 dimensions visuelle Analogskala (EQ-5D VAS)] dargestellt.

Krankheitsaktivität

Unbehandelt oder unzureichend behandelt kann die CU zu einer erheblichen gesundheitlichen Belastung der Patienten führen. Mit zunehmender Entzündung und einer steigenden Krankheitsaktivität nimmt die körperliche Funktionsfähigkeit immer weiter ab. Zudem ist eine aktive Erkrankung mit einer signifikanten Verschlechterung der Lebensqualität verbunden (4-8). Der Ausschuss für Humanarzneimittel der EMA sieht für die Indikation CU die Krankheitsaktivität als patientenrelevanten krankheitsbedingten Endpunkt an (9, 10).

Der Mayo Gesamtscore ist ein Messinstrument zur Bestimmung der Krankheitsaktivät bei Patienten mit CU. Er umfasst vier Subscores: Stuhlfrequenz, rektale Blutungen, endoskopisches Erscheinungsbild und die ärztliche Beurteilung, wodurch eine umfassende und adäquate Beurteilung der Krankheitsaktivität der CU bzw. im Umkehrschluss die Wirksamkeit der Therapie beurteilt werden kann (11). Der Mayo Gesamtscore und dessen einzelne Subscores werden daher als patientenrelevante und valide Endpunkte zur Beurteilung der Morbidität im speziellen der Krankheitsaktivität herangezogen.

Klinische Remission nach Mayo Gesamtscore und Adaptiertem Mayo Score

Nach der aktuellen deutschen S3-Leitlinie ist die Anwendung klinischer und/oder endoskopischer Scores zur Messung der Entzündungsaktivität angezeigt, um ein Therapieansprechen zu quantifizieren und patientenindividuelle Krankheitsverläufe zu objektivieren. Dies ist hilfreich, um die Therapie rechtzeitig zu eskalieren und Darmschädigungen vorzubeugen (12, 13). Hierfür stehen zahlreiche Scores zur Verfügung, welche im Rahmen von klinischen Studien in der Indikation CU häufig zur Beurteilung der Krankheitsaktivität erhoben werden (10, 12-14). Einer der meist verwendeten Scores ist dabei der Mayo Gesamtscore, inklusive des zugehörigen endoskopischen Subscores (13-15). Die verwendeten Scores sollten gemäß der europäischen Leitlinien sowohl Anzeichen und Symptome erfassen, als auch eine endoskopische Beurteilung einschließen (10, 14). Die klinische Remission wird in der aktuellen europäischen Leitlinie dementsprechend als das allumfassende Ausbleiben der Hauptsymptomatik sowie einer endoskopischen Mukosaheilung definiert (16).

Operationalisierung

Der Mayo Gesamtscore ist einer der meist verwendeten Krankheitsaktivitätsindizes in randomisierten kontrollierten Studien in der CU (10, 13, 14). Insgesamt besteht er aus vier Subscores, welche aus Sicht von Patienten und behandelnden Ärzten die wichtigsten Krankheitsdomänen umfassen: rektale Blutungen, Stuhlfrequenz, ärztliche Beurteilung und endoskopisches Erscheinungsbild (11, 14). Jede dieser Domänen wird durch ein Scoring System hinsichtlich des Schweregrads der Erkrankung, beginnend bei 0 (normal/keine Beschwerden) bis 3 (starke Beschwerden), kategorisiert, was insgesamt einen Gesamtscore von 0 bis 12 Punkten ergibt (11). Erst ab einer Gesamtpunktzahl von 3 spricht man von einer aktiven Erkrankung. So definiert eine Gesamtpunktzahl von 3 bis 5 Punkten eine milde aktive Erkrankung, eine Gesamtpunktzahl von 6 bis 10 indiziert eine moderate aktive Erkrankung und eine Gesamtpunktzahl von 11 bis 12 indiziert eine schwere aktive Erkrankung. Die klinische Remission ist definiert als ein Mayo Gesamtscore von insgesamt ≤ 2 Punkten und keiner Subdomäne> 1 Punkt (17).

Der Adaptierte Mayo Score beinhaltet die gleichen Subscores wie der Mayo Gesamtscore mit Außnahme der Arztbeurteilung und hat einen Wertebereich von 0 bis 9. Sowohl für den Mayo Gesamtscore als auch den Adaptierten Mayo Score bilden höhere Werte einen schwereren Krankheitsverlauf ab. Die klinische Remission per Adaptiertem Mayo Score ist definiert als ein Adaptierter Mayo Score ≤ 2, mit einer Stuhlfrequenz ≤ 1 und nicht höher als ihr Ausgangswert, einer rektalen Blutung von 0 und einer Endoskopie-Bewertung von ≤ 1.

Im vorliegenden Dossier wird der Anteil der Patienten mit klinischer Remission nach Mayo Gesamtscore und nach Adaptiertem Mayo Score zu Woche 8 der Induktionsphase und zu Woche 52 der Erhaltungsphase berichtet.

Patientenrelevanz

Eine Reduktion des Mayo Gesamtscores bzw. des Adaptierten Mayo Scores korreliert nachweisbar direkt mit einer vom Patienten wahrnehmbaren Verbesserung der Krankheitsaktivität im Bezug auf unterschiedliche Domänen der Erkrankung (17). Die Patientenrelevanz der durch den Mayo Score messbaren Veränderung der Krankheitsaktivität wird zudem dadurch unterstrichen, dass zwei der Kriterien (Stuhlfrequenz und rektale Blutungen) direkt durch den Patienten selbst berichtet werden. Die beiden anderen Domänen (ärztliche Beurteilung und endoskopisches Erscheinungsbild) zeigten ein sehr starke direkte Korrelation mit diesen patientenberichteten Domänen und sind daher ebenfalls als patientenrelevant einzustufen (18). Des Weiteren konnte eine höhere Krankheitsaktivität gemessen mit dem Mayo Gesamtscore oder dem Adaptierten Mayo Score direkt korreliert werden mit einer Verschlechterung der gesundheitsbezogenen Lebensqualität und dem Auftreten akuter Krankheitsschübe (18). Darüber hinaus stellt ein auffallendes endoskopisches Erscheinungsbild (gemessen durch den endoskopischen Subscore) und damit eine persistierende mukosale Entzündung ein erhebliches Risiko für den Patienten dar, hospitalisiert zu werden und erneute Schübe zu erleben. Eine Verbesserung in der Endoskopie ist daher als patientenrelevant zu betrachten (19).

Validität

Der Mayo Gesamtscore/Adaptierte Mayo Score ist, mitunter aufgrund seiner Einfachheit, das meist verwendete Messinstrument zur Beurteilung der Krankheitsaktivität bei Patienten mit CU, welches in allen großen Studien in der Indikation auf Empfehlung der EMA eingesetzt wird (10, 13, 14, 20). Der Mayo Score wird daher als standardisiertes, von internationalen Zulassungsbehörden anerkanntes Instrument angesehen, um die Entzündungsaktivität zu messen und damit die Krankheitsaktivität der CU zu kategorisieren und die Wirksamkeit von neuen Arzneimitteln in der Indikation zu belegen (10, 12, 21).

Klinisches Ansprechen nach Mayo Gesamtscore

Gemäß der deutschen S3-Leitlinie ist die Anwendung klinischer und/oder endoskopischer Aktivitätsindizes angezeigt, um ein Therapieansprechen zu quantifizieren und patientenindividuelle Krankheitsverläufe zu objektivieren (12, 13). Das klinische Ansprechen nach Mayo Gesamtscore gilt als direktes Maß der Wirksamkeit der Behandlung, die Erhebung im Rahmen von Zulassungstudien wird von der EMA empfohlen (10). Die Verwendung des Mayo Gesamtscore zur Messung des klinischen Ansprechens erlaubt eine zusammenfassende Beurteilung klinischer Symptome (Stuhlfrequenz und rektale Blutungen) und der endoskopischen Ausdehnung der Entzündungsreaktion (14, 20).

Operationalisierung

Im vorliegenden Dossier wird der Anteil der Patienten mit klinischem Ansprechen berichtet, welche zu Woche 8 in der Induktionsphase und zu Woche 44 bzw. 52 in der Erhaltungsphase ein klinisches Ansprechen, definiert als die Reduktion des Mayo Gesamtscores um ≥ 3 Punkte und einer relativen Reduktion von ≥ 30 % zur Baseline aufweisen. Außerdem ist eine Abnahme des Subscores, welcher rektale Blutungen beschreibt, von ≥ 1 oder ein absoluter Wert von ≤ 1 für diesen Subscore Teil der Definition eines klinischen Ansprechens.

Patientenrelevanz

Aufgrund des Designs und der Einteilung in vier Krankheitsdomänen, welche Stuhlfrequenz, rektale Blutungen, ärztliche Beurteilung und endoskopisches Erscheinungsbild umfassen, charakterisiert der Mayo Gesamtscore das Gesamtbild der Krankheitsaktivität. In einer aktuellen Studie konnte gezeigt werden, dass eine Reduktion des Mayo Gesamtscores um ≥ 3 Punkte stark mit einer Verbesserung der patientenrelevanten Lebensqualität korreliert (18). Somit spiegelt eine Reduktion der Gesamtpunktzahl des Mayo Gesamtscores direkt eine vom Patienten wahrnehmbare Linderung der Symptomatik und Verringerung der Krankheitsaktivität wider, woraus sich die Patientenrelevanz dieses Endpunkts ergibt.

Validität

Der Mayo Gesamtscore ist ein etabliertes Messinstrument zur Beurteilung des klinischen Ansprechens, welches auch von internationalen Zulassungsbehörden im Rahmen von regulatorischen Dossiers berücksichtigt und akzeptiert wird (10, 13, 14, 21). Aufgrund der einfachen Anwendbarkeit wird der Mayo Gesamtscore auch regelmäßig in pivotalen RCTs in

der Indikation CU zur Beurteilung der Wirksamkeit der Prüfintervention herangezogen (14, 20). Eine niedrige bzw. minimale Krankheitsaktivität wird in der Literatur durch eine Reduktion des Mayo Gesamtscores von insgesamt ≥ 3 Punkten und/oder einer Reduktion der Gesamtpunktzahl von ≥ 30 % zur Baseline und einer Abnahme des Subscores für rektale Blutungen ≥ 1 oder einen Absolutwert für rektale Blutungen ≤ 1 erreicht (14).

Mukosaheilung

Die CU wird pathophysiologisch durch eine chronische Entzündung mukosaler Zellen des Gastrointestinaltrakts verursacht, daher stellt die Heilung der Mukosa das übergeordnete Ziel der Behandlung der CU dar (13, 14, 22-24). Zahlreiche Studien bestätigen, dass eine Heilung der Mukosa mit einem günstigen klinischen Ausgang für Patienten mit CU assoziiert ist (23, 25, 26). So ist das Erreichen einer endoskopischen und histologischen Heilung der mukosalen Entzündung mit einer verbesserten klinischen Langzeitremission, einer Reduktion des Risikos einer Kolektomie und einer Verringerung des Steroidverbrauchs assoziiert (15, 19, 27, 28). Auch in systematischen Reviews konnte diese Assoziation zwischen einer Heilung der Mukosa und der Vermeidung von CU-spezifischen Komplikationen sowie dem Erreichen einer steroidfreien Remission bestätigt werden (13, 23, 26, 27, 29). Die Mukosaheilung sollte nach den Empfehlungen der deutschen S3-Leitlinie und der EMA-Leitlinie zur Durchführung von Zulassungstudien mittels endoskopischer Kontrolle nachgewiesen werden (10, 12-14).

Operationalisierung

Die Mukosaheilung wird in den eingeschlossenen Studien definiert als eine histologische, endoskopische Verbesserung der Mukosa. Um das histologische Erscheinungsbild der mukosalen Entzündung evaluieren zu können, wird der Geboes Score, der meist verwendete Score in pivotalen Studien der Indikation CU, angewendet. Liegt dieser Score bei ≤ 3,1 und der endoskopische Subscore des Mayo Gesamtscores bei 0 wird eine Mukosaheilung definiert (30). Der Geboes Score umfasst sechs Kategorien: strukturelle (architektonische) Veränderungen des Darms (Kategorie 0), chronische Entzündungsinfiltrate (Kategorie 1), Neutrophile und Eosinophile der Lamina propria (Kategorie 2), Neutrophile der Epithelzellschicht (Kategorie 3), Zerstörung der Krypten (Kategorie 4) und Ulzerationen (Kategorie 5). Der Score kann Werte zwischen 0 und 5,4 Punkten annehmen, wobei höhere Werte eine stärkere histologisch begründete Entzündung anzeigen. Ein Geboes Score ≤ 3,1 zeigt eine histologisch begründete Entzündung mit strukturellen Veränderungen des Darms an (30-32). Jede Kategorie ist in vier Unterkategorien in Abhängigkeit des Schweregrads der jeweiligen Kategorie aufgeteilt. Eine histologische Evaluierung der mukosalen Entzündung stellt einen essentiellen Messparameter dar, um die Aktivität der CU zu erfassen. Ein augenscheinlich klinischer und endoskopischer Krankheitsstillstand ist, aufgrund des Ausbleibens einer histologischen Mukosaheilung, weiterhin mit einem hohen Risiko eines Rückfalls, sowie einem hohen Krebsrisiko assoziiert (33). Diese Operationalisierung der Mukosaheilung wird regelhaft im Anwendungsgebiet in klinischen Studien eingesetzt, u. a. in pivotalen Studien zur ZVT (34-36).

Im vorliegenden Dossier wird der Anteil der Patienten mit Mukosaheilung zu Woche 8 und zu Woche 44 bzw. 52 berichtet.

Patientenrelevanz

In den letzten Jahren hat der Endpunkt Mukosaheilung aufgrund seiner klinischen Relevanz mehr und mehr an Bedeutung gewonnen (22, 23, 37). In vielen klinischen und pharmakologischen Studien im Anwendungsgebiet wird die Mukosaheilung als wichtiger Endpunkt erfasst (22, 37, 38). Eine aktuelle Studie zeigt eine Korrelation der Mukosaheilung mit dem Erreichen einer Langzeitremission und längerer Zeitintervalle ohne einen Krankheitsrückfall (22). Außerdem ist die Mukosaheilung assoziiert mit einer Reduktion der Hospitalisierungsrate, einer Vermeidung von Komplikationen und chirurgischen Eingriffen und einer signifikanten Verbesserung der Lebensqualität betroffener Patienten (20, 39-41). Zusätzlich ist auch das Krebsrisiko und die krebsverursachte Mortalität weitaus geringer (22, 42).

Validität

Zusammenfassend ist die histo-endoskopische Beurteilung der Mukosaheilung ein essentieller Endpunkt, der aufgrund von Empfehlungen in nationalen und internationalen Leitlinien sowie der Richtlinie der EMA, regelhaft in pivotalen Studien der Indikation CU erhoben wird. (10, 13, 14, 22, 29). In der aktuellen deutschen S3-Leitlinie wird explizit die große prognostische Bedeutung der Mukosaheilung im klinischen Verlauf der CU betont (12, 13).

Stuhlfrequenz (SFS) (Subscore des Mayo Gesamtscores)

Primäre Symptome der CU umfassen rektale Blutungen, Stuhldrang, Diarrhoe, abdominelle Schmerzen, fäkale Inkontinenz, Fatigue, Fieber und Gewichtsverlust (13). Aufgrund des kontinuierlichen Wiederaufflammens akuter Symptome sind Patienten der CU mit starken physischen aber auch psychischen Einschränkungen konfrontiert (5, 43, 44). Ein schwer kontrollierbarer Stuhldrang und Diarrhoe bedingen Fehlzeiten am Arbeitsplatz, Arbeitslosigkeit und als finale Konsequenz soziale Isolation und die Entwicklung von Depressionen.

Operationalisierung

Einer der vier Subscores zur Evaluierung der Entzündung durch den Mayo Gesamtscore beschreibt die Stuhlfrequenz. In einer Skala von 0 bis 3 Punkten wird der Schweregrad der Stuhlfrequenz kategorisiert. Während ein Score von 0 eine normale Stuhlfrequenz definiert, bildet ein Score von 1 eine gesteigerte Stuhlfrequenz von 1 bis 2 Stühlen mehr als normal pro Tag ab. Bei einer Stuhlfrequenz von 3 bis 4 Stühlen mehr als normal pro Tag wird ein Score von 2 definiert, ab > 5 Stühle mehr als normal pro Tag wird ein Score von 3 Punkten erhoben (18).

Im vorliegenden Dossier wird der Anteil der Patienten mit einer SFS ≤ 1 zu Woche 8 und 44 berichtet.

Patientenrelevanz

Das primäre Behandlungsziel der CU ist das Erreichen und die Stabilisierung einer klinischen Remission (13). Für einen Großteil der Patienten bedeutet dies eine Reduktion der Stuhlfrequenz, das Ausbleiben rektaler Blutungen und einer Stuhlinkontinenz sowie die Minimierung der oben angesprochenen psychischen Belastungen (11, 18). In direkter Konsequenz folgt eine signifikante Verbesserung der patientenbezogenen Lebensqualität. Daher ist das Erfassen der Stuhlfrequenz als Messparameter des Mayo Gesamtscores als patientenrelevanter Endpunkt zu betrachten. Im Rahmen der G-BA-Beratung zum AWG wurde die Stuhlfrequenz als patientenrelevanter Endpunkt vom G-BA bestätigt (45).

Validität

In mehreren wissenschaftlichen Arbeiten konnte eine starke Korrelation zwischen dem Mayo Gesamtscore und einzelnen Subscores der Kategorien Stuhlfrequenz und rektale Blutungen nachgewiesen werden (17, 18). Das Erfassen der Stuhlfrequenz als Messparameter des Mayo Gesamtscores erfolgt in allen pivotalen Studien der Indikation CU (13, 14, 20). Ein für einen Patienten gut kontrollierbarer Stuhlgang ist als valide zu erachten.

Rektale Blutung (RBS) (Subscore des Mayo Gesamtscore)

Rektale Blutungen gehören zur Hauptsymptomatik der CU (13), welche eine erhebliche Beeinträchtigungen der patientenrelevanten Lebensqualität bedingen und somit u. a. zur eingeschränkten sozialen Teilhabe und der Entwicklung einer Depression beitragen (4, 5). Die Reduktion des Mayo Gesamtscores korreliert nachweisbar direkt mit einer vom Patienten wahrnehmbaren Verbesserung der Krankheitsaktivität und damit einhergehend mit dem übergeordneten Therapieziel der Verbesserung der Lebensqualität (13, 17).

Operationalisierung

Über das Kriterium rektale Blutungen des Mayo Gesamtscores wird die Schwere der Blutung auf einer Skala von 0 (keine Blutung) bis 3 (hauptsächlich Blut) abgeschätzt (11). Im vorliegenden Dossier wird der Anteil der Patienten mit einer RBS = 0 zu Woche 8 und zu Woche 44 berichtet.

Patientenrelevanz

Das Ausbleiben rektaler Blutungen als Bestandteil einer Remission, welche als primäres Ziel einer medikamentösen Therapie anzusehen ist, hat direkte Konsequenzen auf physische und psychische Belastungen, denen Patienten mit CU ausgesetzt sind (13). Hinsichtlich der patientenbezogenen Lebensqualität tritt eine immense Verbesserung durch das Ausbleiben eines der Hauptsymptome ein. Somit ist der Endpunkt rektale Blutungen, erfasst als ein Kriterium des Mayo Gesamtscores, als patientenrelevant anzusehen. Im Rahmen der G-BA Beratung zum AWG wurde rektale Blutung als patientenrelevanter Endpunkt vom G-BA bestätigt (45). Zudem wird der Subscore RBS direkt vom Patienten beantwortet und ist daher per se als patientenrelevant zu erachten.

Validität

In allen pivotalen Studien der Indikation CU erfolgte die Dokumentation rektaler Blutungen im Zuge der Evaluation der Entzündung im Rahmen des Mayo Gesamtscores (14, 20). Dieses Messinstrument der Krankheitsaktivität wird von internationalen Zulassungsbehörden akzeptiert und von führenden Wissenschaftlern für klinische Studien in der Indikation CU empfohlen (10, 12, 21). Die Evaluation rektaler Blutungen durch ein Kriterium des Mayo Gesamtscores ist somit auch in der Indikation CU als valider Endpunkt anzusehen. Die komplette Abwesenheit von Blut im Stuhl (RBS = 0) ist als valide zu erachten.

EuroQol-5D visuelle Analogskala (EQ-5D VAS)

Der EQ-5D ist ein standardisiertes, generisches Messinstrument, welches 1990 von der EuroQol-Forschungsgruppe zur Erfassung des Gesundheitszustands entwickelt wurde (46, 47). Er kann bei einer Vielzahl an Indikationen im Rahmen von klinischen und gesundheitsökonomischen Untersuchungen eingesetzt werden und hat sich weltweit als Instrument für die klinische Forschung etabliert (48). Der EQ-5D besteht aus einem deskriptiven Teil in Form eines Fragebogens und einer Visuellen Analogskala (VAS). Der Fragebogen dient der Selbsteinschätzung des aktuellen Gesundheitsstatus durch den Patienten in den fünf Dimensionen Mobilität, Selbstversorgung, tägliche Aktivitäten, Schmerzen/Unbehagen und Ängste/Depressionen. Die VAS kann als quantitatives Maß für den Gesundheitszustand verwendet werden, die die eigene Einschätzung des Patienten widerspiegelt.

Operationalisierung

Bei der EQ-5D VAS schätzen die Befragten ihren Gesundheitszustand auf einer in der Regel 20 cm langen Skala im Wertebereich von 0 (schlechtester vorstellbarer Gesundheitszustand) bis 100 (bester vorstellbarer Gesundheitszustand) ein. Höhere Werte entsprechen hier einem besseren Gesundheitszustand (47, 48). Im vorliegenden Dossier wird die mittlere Veränderung der EQ-5D VAS zu Woche 8 und zu Woche 44 bzw. 52 im Vergleich zu Baseline dargestellt.

Patientenrelevanz

Der EQ-5D wird indikationsunabhängig eingesetzt und die VAS gilt als verlässliches Messinstrument für die Erfassung und insbesondere für die Veränderung des Gesundheitszustands. Die Erhebung des patientenrelevanten allgemeinen Gesundheitszustands durch die EQ-5D VAS ist adäquat, da der Patient hierbei direkt selbst über diesen Auskunft erteilt; entsprechend ist die Patientenrelevanz von Endpunkten, die mittels der EQ-5D VAS erhoben wurden, auch durch den G-BA anerkannt (9, 49).

Validität

Der EQ-5D ist ein häufig verwendetes generisches Messinstrument zur Bestimmung des Gesundheitszustands (50) und gilt in der Indikation CU wie in vielen anderen Indikationen als valides und verlässliches Messinstrument (8, 51, 52)

Gesundheitsbezogene Lebensqualität

Die CU kann weitreichende und schwerwiegende Auswirkungen auf die Lebensqualität des Patienten haben (4-8, 51, 53). Bei vielen CU-Patienten vergehen Jahre bevor ihre Erkrankung korrekt diagnostiziert wird. Diese Patienten haben nachweislich schwere mukosale Entzündungen und ein erhöhtes Risiko für Komplikationen (12, 54). Die Schwere der Erkrankung beeinflusst wiederum die Lebensqualität der Patienten. Mit zunehmender Ausdehnung der Entzündung und steigender Krankheitsaktivität wird die Lebensqualität enorm eingeschränkt (4-8). Eine langfristige Verbesserung der gesundheitsbezogenen Lebensqualität bzw. die Optimierung der gesundheitsbezogenen Lebensqualität ist daher bei der Therapie der CU von hoher Relevanz (13, 14).

Allgemein ist die gesundheitsbezogene Lebensqualität gemäß den allgemeinen Methoden des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) eine „patientenrelevante Zielgröße zur Feststellung krankheits- und behandlungsbedingter Veränderungen“ (55). Die gesundheitsbezogene Lebensqualität stellt ein breites Konzept dar, welches verschiedene Gesundheitsaspekte, wie beispielsweise Fatigue, körperliche Funktionalität und Emotionalität, umfasst. Diese können mit Hilfe von generischen oder krankheitsspezifischen Messinstrumenten gemessen werden. Die dabei am häufigsten verwendeten Messinstrumente sind generische Fragebögen wie z. B. der Short Form – 36 (SF-36) (8, 10, 51).

Inflammatory Bowel Disease Questionnaire (IBDQ)

Die Beurteilung der gesundheitsbezogenen Lebensqualität durch generische Instrumente kann sehr sinnvoll durch Instrumente ergänzt werden, die spezifisch auf das jeweilige Krankheitsbild abgestimmt sind. Der IBDQ ist ein etablierter Fragebogen, der die gesundheitsbezogene Lebensqualität von Patienten mit chronisch-entzündlicher Darmerkrankung (Inflammatory Bowel Disease, IBD [Colitis ulcerosa und Morbus Crohn]) erfasst.

Operationalisierung

Der IBDQ umfasst 32 Fragen, welche vier untergeordnete Komponenten beschreiben: Darmsymptome (10 Elemente), systemische Symptome (5 Elemente), emotionale Funktion (12 Elemente) und Sozialfunktion (5 Elemente). Jede Frage kann auf einer Skala von 1 (schlechteste Situation) bis 7 (beste Situation) bewertet werden. Die Gesamtpunktzahl reicht von 32 bis 224 Punkten, wobei eine höhere Gesamtpunktzahl eine bessere gesundheitsbezogene Lebensqualität indiziert (56, 57). Jede untergeordnete Komponente des IBDQ wird ihrerseits anhand Gesamtpunktzahlen von 10 bis 70, 5 bis 35, 12 bis 84 und 5 bis 35 berechnet (58, 59).

Im vorliegenden Dossier wird die mittlere Veränderung des IBDQ zu Woche 8 und zu Woche 44 bzw. 52 im Vergleich zu Baseline dargestellt.

Patientenrelevanz

Patienten mit CU sind von einer erheblichen Einschränkung ihrer Lebensqualität betroffen (5, 7, 8, 60). Die gesundheitsbezogene Lebensqualität dieser Patienten kann mit Hilfe des IBDQ durch den Patienten selbst subjektiv eingeschätzt werden. Daher wird der indikationsspezifische Fragebogen IBDQ vom G-BA als grundsätzlich geeignet eingestuft, was seine Patientenrelevanz verdeutlicht (45).

Validität

Eine durch die Patienten selbst angewendete Form sowie eine Kurzform des Fragebogens wurden validiert und als valide und verlässlich eingestuft. Der Validierungsstatus des IBDQ wurde 2004 zusammengestellt und 2006 für die deutsche Version bestätigt (61, 62). Auch gemäß der aktuellen deutschen S3-Leitlinie und den Empfehlungen des Ausschusses für Humanarzneimittel der EMA stellt der IBDQ ein valides Instrument zur Erfassung der gesundheitsbezogenen Lebensqualität bei Patienten mit CU dar (10, 13).

SF-36 (PCS und MCS)

Beim SF-36 handelt es sich um ein generisches und international eingesetztes Messinstrument zur Selbsteinschätzung der gesundheitsbezogenen Lebensqualität durch den Patienten, welches in klinischen sowie Interventions- und Bevölkerungsstudien zum Einsatz kommt (63). Er wurde erstmals 1989 vorgestellt und umfasst zwei Komponenten zur physischen und mentalen Gesundheit der Patienten (64). Der SF-36 ist weltweit das am häufigsten eingesetzte patientenbeurteilte Instrument (60).

Operationalisierung

Im SF-36 Version 2 werden mittels 36 Fragen acht Dimensionen (körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperliche Schmerzen, allgemeiner Gesundheitszustand, Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden) auf einer Skala von 0 (geringste Lebensqualität) bis 100 (höchste Lebensqualität) erfasst, wobei höhere Werte für eine bessere gesundheitsbezogene Lebensqualität stehen. Aus diesen Dimensionen wurden zwei Summenscores berechnet: der Physical Health Component Summary Score (PCS) und der Mental Health Component Summary Score (MCS). Die Berechnung des PCS und des MCS erfolgt über ein Norm Based Scoring auf einen Mittelwert von 50 und eine Standardabweichung von 10, ein definiertes Minimum und Maximum haben die Summenscores daher nicht (63).

Tabelle 4-9: Domänen und Summenskalen des SF-36-Fragebogens

Domäne	**Summenskala**
Körperliche Funktionsfähigkeit	PCS
Körperliche Rollenfunktion	
Körperliche Schmerzen	
Allgemeiner Gesundheitszustand	

Domäne	Summenskala
Vitalität Soziale Funktionsfähigkeit Emotionale Rollenfunktion Psychisches Wohlbefinden	MCS
MCS: Mental Health Component Summary Score; PCS: Physical Health Component Summary Score; SF-36: Short Form – 36	

Im vorliegenden Dossier wird die mittlere Veränderung des SF-36 (PCS und MCS), zu Woche 8 der Induktionsphase und zu Woche 44 bzw. 52 im Vergleich zu Baseline dargestellt.

Patientenrelevanz

Patienten mit CU sind von einer erheblichen Einschränkung ihrer Lebensqualität betroffen (5, 7, 8, 60). Das Ausmaß dieser Einschränkung kann mit Hilfe des SF-36 durch den Patienten subjektiv eingeschätzt werden. Von Seiten des G-BA wird unter den generischen Fragebögen zur Erfassung der Lebensqualität der SF-36 als bevorzugt empfohlen (45). Dementsprechend wurde die gesundheitsbezogene Lebensqualität, erhoben anhand des SF-36, bereits in anderen Verfahren zur frühen Nutzenbewertung vom G-BA berücksichtigt (9, 49).

Validität

In klinischen Studien zur CU ist der SF-36 das am häufigsten verwendete generische Messinstrument der Lebensqualität (65). Seine Validität und Zuverlässigkeit wurde vielfach geprüft (10). Auch für die Indikation CU ist der SF-36 als ein verlässliches und validiertes Messinstrument anzuerkennen (10, 60).

Verträglichkeit

Die Beurteilung der Verträglichkeit des Arzneimittels ist ein wesentlicher Bestandteil der klinischen Prüfung und umfasst UE, die bei einem Patienten nach Verabreichung des Prüfpräparats auftreten. Die Verringerung von therapieassoziierten Nebenwirkungen ist grundsätzlich als patientenrelevanter Aspekt eines zu bewertenden Arzneimittels anzusehen. Jedoch sind nicht alle untersuchten UE direkt patientenrelevant, da beispielsweise Änderungen von Laborparametern nicht per se für den Patienten spürbar und nicht zwingend symptomatisch sind. Die Patientenrelevanz zum Endpunkt Verträglichkeit wird vom G-BA verlangt und anerkannt, um den Nutzen und Schaden eines Arzneimittels abzuwägen.

Zur Beurteilung der Verträglichkeit werden die Gesamtraten für folgende UE, für die im Rahmen eines indirekten Vergleiches Daten vorliegen, zu Woche 8 und Woche 44 bzw. 52 dargestellt:

- Alle UE

- Schwerwiegende UE (SUE)
- UE, die zum Therapieabbruch führten

Darüber hinaus werden, um ein vollständiges Bild des Verträglichkeitsprofils von Upadacitinib abbilden zu können, ergänzend folgende Gesamtraten für UE für Upadacitinib in der Erhaltungsphase dargestellt:

- Schwere UE
- UE, die zum Tod führten

Darüber hinaus werden, um ein vollständiges Bild des Verträglichkeitsprofils von Upadacitinib abbilden zu können, für die Erhaltungsphase von Upadacitinib ergänzend folgende im Prüfplan präspezifizierte UE von speziellem Interesse (UESI) dargestellt:

- Schwerwiegende Infektionen
- Opportunistische Infektionen (ohne Tuberkulose und Herpes Zoster)
- Mögliche Malignität (alle Typen)
    - Malignität
    - Nicht melanozytärer Hautkrebs (Non Melanoma Skin Cancer, NMSC)
    - Malignität außer NMSC
    - Lymphom
- Hepatische Erkrankungen
- Gastrointestinale Perforation
- Anämie
- Neutropenie
- Lymphopenie
- Herpes Zoster
- Erhöhte Kreatinphosphokinase
- Nierendysfunktion
- Aktive Tuberkulose

- Adjudizierte kardiovaskuläre Ereignisse (Major Adverse Cardiovascular Events, MACE)
    - Kardiovaskulärer Tod
    - Nicht tödlicher Myokardinfarkt
    - Nicht tödlicher Schlaganfall
- Andere adjudizierte kardiovaskuläre Ereignisse
- Unbestimmte/unbekannte Todesursache
- Venöse thromboembolische Ereignisse (VTE)
    - Tiefe Beinvenenthrombose
    - Lungenembolie
- Andere Venenthrombosen
- Arterielle thromboembolische Ereignisse (nicht kardial, nicht neurologisch)

Da im Rahmen des indirekten Vergleiches für die ZVT keine eindeutigen Daten für UE ohne erkrankungsbezogene Ereignisse vorliegen, wird auf eine Darstellung von UE unter Ausschluss erkrankungsbezogener Ereignisse verzichtet.

Analysezeitpunkte

Bei der CU handelt es sich um eine Erkrankung mit chronischem Verlauf, daher ist die Erfassung der Nachhaltigkeit der Effekte von wesentlicher Bedeutung. Die Behandlung der CU mit dem zu bewertenden Arzneimittel und der ZVT ist in zwei Phasen gegliedert: die Induktionsphase und die Erhaltungsphase. Gemäß den Leitlinien der EMA bildet eine Behandlungs- und Beobachtungsdauer von 44 Wochen beziehungsweise 52 Wochen in der Erhaltungsphase die klinische Praxis ab (10).

Im vorliegenden Dossier wird für alle Endpunkte zur Morbidität, gesundheitsbezogenen Lebensqualität und Verträglichkeit primär der letztverfügbare Analysezeitpunkt zu Woche 44 beziehungsweise Woche 52 der Erhaltungsphase dargestellt. Für die Studie U-ACHIEVE Substudie 3 mit dem zu bewertenden Arzneimittel wird primär der Analysezeitpunkt zu Woche 52 erhoben, für die Studie UNIFI Erhaltungsstudie, die die ZVT umfasst, wird der primäre Analysezeitpunkt zu Woche 44 erhoben. Um eine bessere Vergleichbarkeit mit dem letztverfügbaren Analysezeitpunkt zwischen den Studien zu gewährleisten, wird für Endpunkte – sofern verfügbar – der gemeinsame Analysezeitpunkt zu Woche 44 berichtet. Dies ist nur für eine kleine Auswahl an Endpunkten möglich, da der überwiegende Anteil der Endpunkte in den

Erhaltungsstudien nur zu Woche 52 oder nur zu Woche 44 erhoben wurde. Insgesamt ist der Vergleich auch zwischen den unterschiedlichen Zeitpunkten zu Woche 44 und Woche 52 als adäquat anzusehen, da dieser die klinische Praxis abbildet. Sowohl für Upadacitinib als auch für Ustekinumab, gibt es keine Hinweise auf relevante Unterschiede zwischen den Zeitpunkten zu Woche 44 und Woche 52. Verlaufsgrafiken für die klinische Remission anhand des Partiellen Mayo Score zeigen konstante Ansprechraten zwischen den genannten Zeitpunkten bei der U-ACHIEVE Substudie 3 (66) als auch bei der UNIFI Erhaltungsstudie (67).

Die Studiendauer ist mit einem Erhebungszeitpunkt von 52 Wochen bei Upadacitinib und einem Erhebungszeitpunkt von 44 Wochen bei Ustekinumab sowohl in Bezug auf die Studiendauer, als auch die Vergleichbarkeit der zwei Zeitpunkte adäquat.

Zusätzlich wird für alle Endpunkte der Analysezeitpunkt der Induktionsstudien zu Woche 8 dargestellt, um die Induktionsphase, als Teil der Behandlung mit dem zu bewertendem Arzneimittel und der ZVT, entsprechend dem Versorgungsalltag für die gesamte Labelpopulation, abzubilden.

Upadacitinib

Im vorliegenden Dossier werden für das zu bewertende Arzneimittel zwei Analysezeitpunkte dargestellt.

Die ersten acht Wochen der Behandlung stellen die Induktionsphase dar, die durch die Studien U-ACHIEVE Substudie 1 und Substudie 2 sowie die Studie U-ACCOMPLISH abgebildet wird. Für diese Studien werden im vorliegenden Dossier die Ergebnisse zum Analysezeitpunkt zu Woche 8 dargestellt. Patienten aus den Induktionsstudien U-ACHIEVE Substudie 1 und Substudie 2 sowie der Induktionsstudie U-ACCOMPLISH, die zu Woche 8 ein klinisches Ansprechen aufweisen und für die Behandlung in der Erhaltungsphase geeignet waren, werden in die Studie U-ACHIEVE Substudie 3 eingeschlossen, die die Erhaltungsphase über 52 Wochen abbildet. Für die Erhaltungsstudie U-ACHIEVE Substudie 3 wird primär der Analysezeitpunkt zu Woche 52 dargestellt. Vereinzelte Endpunkte werden auch zu Woche 44 erhoben und zu diesem Analysezeitpunkt berichtet.

Ustekinumab

Die Induktionsphase der ersten acht Wochen Behandlung wird durch die Studie UNIFI Induktionsstudie abgebildet, welcher im vorliegenden Dossier als Analysezeitpunkt für die Induktionsphase berichtet wird. Patienten, die in der Induktionsphase ein klinisches Ansprechen auf eine Behandlung mit Ustekinumab zeigen, werden in die UNIFI Erhaltungsstudie eingeschlossen und im Rahmen der Erhaltungsphase für weitere 44 Wochen behandelt. Im vorliegenden Dossier wird für alle Endpunkte zur Morbidität, gesundheitsbezogenen Lebensqualität und Verträglichkeit der letztverfügbare Analysezeitpunkt zu Woche 44 der Erhaltungsphase dargestellt.

Imputation und Umgang mit fehlenden Werten

Upadacitinib

Für die Analysen der binären Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität werden im Rahmen der Erhaltungsphase im vorliegenden Dossier fehlende Werte mit einer modifizierten NRIC-Analyse (Non-Responder-Imputation Incorporating Multiple Imputation to Handle Missing Data due to COVID-19) imputiert. Patienten mit fehlenden Daten aufgrund einer Infektion mit COVID-19 oder aufgrund von logistischen Einschränkungen durch die Pandemie, werden mittels multipler Imputation (MI) ersetzt. Patienten mit fehlenden Werten, die nicht in Zusammenhang mit COVID-19 stehen, werden für den entsprechenden Untersuchungszeitpunkt und Endpunkt als Non-Responder angesehen. Darüber hinaus werden Patienten, die eine Rescue-Therapie erhalten, als Non-Responder angesehen. Dieses Vorgehen ist im statistischen Analyseplan (SAP) für binäre Endpunkte präspezifiziert. Binäre Endpunkte der Induktionsstudien werden, wie in den entsprechenden SAP präspezifiziert, per NRI-Methode (U-ACHIEVE Substudie 1) beziehungsweise per NRIC-Analyse (U-ACHIEVE Substudie 2, U-ACCOMPLISH) imputiert.

Für die Analysen stetiger Endpunkte werden fehlende Werte für die Erhaltungsphase im vorliegenden Dossier mit einer RTB-MI-Analyse (Multiple Imputation Incorporating Return-to-Baseline to Handle Visits After UC-related Rescue Medication-Use) imputiert. Hierbei werden potenzielle Verletzungen der Missing-at-Random (MAR)-Annahme aufgrund einer Rescue-Therapie wie folgt adressiert: Daten werden nach einer Rescue-Therapie als fehlende Werte gekennzeichnet und alle fehlenden Werte werden mittels MI imputiert. Für die imputierten Datensätze werden, unter der Annahme, dass Wirksamkeitsdaten nach Einleitung einer Rescue-Therapie sich wieder ihrem Ausgangswert angleichen, fehlende Angaben zur Veränderung gegenüber dem Ausgangswert durch einen Wert einer Normalverteilung $N(0, V_c)$ ersetzt. V_c ist hierbei die Varianz der Veränderung zu Baseline, geschätzt aus allen beobachteten Werten unabhängig von den Behandlungsgruppen und unter Ausschluss der Werte nach Einleitung einer Rescue-Therapie. Dieses Vorgehen zum Umgang mit fehlenden Werten bei stetigen Endpunkten ist im SAP präspezifiziert. Stetige Endpunkte der Induktionsstudien werden, wie in den entsprechenden SAP präspezifiziert, per LOCF (Last Observation Carried Forward)-Methode (U-ACHIEVE Substudie 1) beziehungsweise über Mixed Model for Repeated Measures (MMRM)-Modelle, welche alle beobachteten Datenpunkte berücksichtigen (U-ACHIEVE Substudie 2, U-ACCOMPLISH), adressiert.

Ustekinumab

Für die Analysen der binären Endpunkte werden fehlende Daten als ein Nichterreichen des entsprechenden Endpunktes gewertet. Fehlende Werte werden demzufolge mittels NRI-Methode imputiert. Dieses Vorgehen ist im SAP präspezifiziert.

Für die Analysen stetiger Endpunkte werden für fehlende Werte der zuletzt beobachtete Wert fortgeschrieben. Fehlende Werte werden demzufolge mittels LOCF-Methode imputiert. Eine Ausnahme hiervon bildet der Mayo Gesamtscore, für welchen im Falle fehlender Werte, der

zuletzt verfügbare Wert eines Mayo Subscores eingetragen wird. Dieses Vorgehen ist im SAP präspezifiziert.

Statistische Methoden

Upadacitinib

Die Analyse der binären Endpunkte zur Morbidität, gesundheitsbezogenen Lebensqualität und der Verträglichkeit, basiert auf einem unstratifizierten Cochran-Mantel-Haenszel (CMH)-Test. Zum Vergleich der Behandlungsgruppen werden das relative Risiko (RR), das Odds-Ratio (OR) sowie die Risikodifferenz (RD) mit dem jeweiligen 95 %-Konfidenzintervall (KI) und p-Wert als Effektmaße, auf Basis der Häufigkeiten aus dem Studienbericht berechnet. Fehlende Werte werden wie im vorherigen Abschnitt beschrieben ersetzt. Für die Bewertung des Behandlungseffekts im Rahmen der Zusatznutzenableitung wird das RR herangezogen.

Analog zum Vorgehen im Studienbericht erfolgt für das vorliegende Dossier die Auswertung stetiger Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität, welche longitudinal erhoben wurden, mittels MMRM mit Behandlungsgruppe, Visite, der Interaktion aus Behandlung und Visite sowie den Stratifizierungsvariablen Bio-IR-Status (Bio-IR, Non-Bio-IR) zu Baseline der Induktionsstudien, Vorhandensein einer klinischen Remission zu Beginn der Erhaltungsphase (ja, nein) und Einnahme von Kortikosteroiden zu Beginn der Erhaltungsphase (ja, nein) als feste Effekte, sowie dem zugehörigen Baseline-Wert als Kovariable. Stetige Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität, für die nur ein post-Baseline-Wert vorliegt, werden mittels einer Kovarianzanalyse (ANCOVA) mit Behandlungsgruppe, sowie den Stratifizierungsvariablen Bio-IR-Status (Bio-IR, Non-Bio-IR) zu Baseline der Induktionsstudien, Vorhandensein einer klinischen Remission zu Beginn der Erhaltungsphase (ja, nein) und Einnahme von Kortikosteroiden zu Beginn der Erhaltungsphase (ja, nein) als feste Effekte, sowie dem zugehörigen Baseline-Wert als Kovariable analysiert. Fehlende Werte für stetige Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität werden wie im vorherigen Abschnitt beschrieben ersetzt. Im vorliegenden Dossier werden post hoc die standardisierten Mittelwertdifferenzen (Hedges‘ g) berechnet und dargestellt. Die Berechnung des Hedges‘ g basiert auf der inversen Varianz-Methode auf Basis der Populationsgröße N, der mittleren Veränderung und der zugehörigen Standardabweichung (Standard Deviation, SD). Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht vorliegt, erfolgte eine Schätzung über den Standardfehler (Standard Error, SE) mittels der Formel $SD = SE * \sqrt{N}$.

Ustekinumab

Die Analyse der binären Endpunkte zur Morbidität, gesundheitsbezogenen Lebensqualität und der Verträglichkeit, erfolgt analog zu Upadacitinib.

Die Analyse der stetigen Endpunkte des Mayo Gesamtscores und modifizierte Varianten des Mayo Scores, erfolgt mittels einer ANCOVA mit dem zugehörigen Baseline-Wert, Vorhandensein einer klinischen Remission zu Beginn der Erhaltungsphase,

Induktionsbehandlung und Behandlungsgruppe der Erhaltungsphase als Kovariablen. Für weitere stetige Endpunkte wird diese ANCOVA mit den beschriebenen Kovariablen auf Basis der van der Waerden Methode durchgeführt.

Analysepopulation

Upadacitinib

Zur Beurteilung des Zusatznutzens von Upadacitinib wurden vom G-BA zwei Fragestellungen (Patienten mit oder ohne unzureichendes Ansprechen auf eine biologische Therapie: Bio-IR, Non-Bio-IR) formuliert. Im Rahmen des indirekten Vergleichs ist eine getrennte Beantwortung dieser Fragestellungen nicht möglich, da die Daten nicht für alle Endpunktdimensionen bzw. patientenrelevante Endpunkte, insbesondere nicht für die Verträglichkeit, getrennt nach Vorbehandlung vorliegen. Analysen zur Zielpopulation sind aber im Rahmen des indirekten Vergleichs adäquat, da ein gemeinsames Therapieziel zu Grunde liegt. Da Ustekinumab für beide Fragestellungen als ZVT durch den G-BA definiert wurde und sich bei einer Behandlung unter Upadacitinib für die Erhaltungsphase in der dargestellten, deskriptiven Subgruppenanalyse, vergleichbare Ansprechraten für Bio-IR- und Non-Bio-IR-Patienten zeigen, werden beide Fragestellungen zu einer Fragestellung zusammengefasst und Ergebnisse für die Zielpopulation dargestellt.

Die Auswertung der binären und stetigen Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität erfolgt anhand der Population aller randomisierten Patienten, die eine Dosis entsprechend ihrer zugewiesenen Studienmedikation erhalten haben. Die Darstellung der Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität basiert in Induktions- und Erhaltungsstudien jeweils auf der ITT-Population Diese umfasst alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben und ein Ansprechen auf Upadacitinib zu Woche 8 zeigten, und wird entsprechend der Randomisierung ausgewertet. Sie bildet die präspezifizierte Analysepopulation für die Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität.

Die Auswertung aller Endpunkte zur Verträglichkeit erfolgt für Induktions- und Erhaltungsstudien anhand der Safety Population, diese schließt alle randomisierten Patienten ein, die mindestens eine Dosierung der Studienmedikation erhalten haben und wird entsprechend der erhaltenen Studienmedikation ausgewertet. Diese entspricht der für alle Sicherheitsendpunkte präspezifizierten Analysepopulation. Für Endpunkte zur Verträglichkeit werden die Anteile der Patienten mit mindestens einem UE zum Zeitpunkt zu Woche 8 (Induktionsphase) und Woche 52 (Erhaltungsphase) dargestellt.

Ustekinumab

Die Auswertung der binären und stetigen Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität erfolgt nach dem ITT-Prinzip. Für die Induktionsphase werden alle randomisierten Patienten in den Auswertungen berücksichtigt. Für die Erhaltungsphase werden

alle randomisierten Patienten, die ein klinisches Ansprechen auf eine 8-wöchige Ustekinumab-Behandlung hatten, in den Auswertungen berücksichtigt.

Die Auswertung aller Endpunkte zur Verträglichkeit erfolgt für die Induktions-, sowie die Erhaltungsphase auf Basis der Safety Population. Diese beinhaltet alle Patienten die mindestens eine Dosierung der Studienmedikation erhalten haben.

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 nach der Knapp-Hartung-Methode mit der Paule-Mandel-Methode zur Heterogenitätsschätzung[5] erfolgen. Im Fall von sehr wenigen Studien ist die Heterogenität nicht verlässlich schätzbar. Liegen daher weniger als 5 Studien vor, ist auch die Anwendung eines Modells mit festem Effekt oder eine qualitative Zusammenfassung in Betracht zu ziehen. Kontextabhängig können auch alternative Verfahren wie z. B. Bayes'sche Verfahren oder Methoden aus dem Bereich der generalisierten linearen Modelle in Erwägung gezogen werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär 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[6] werden.

[5] Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. Cochrane Database Syst Rev 2015: 25-27.
[6] 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.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand geeigneter statistische Maße auf Vorliegen von Heterogenität[7, 5] erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik $\geq 0{,}05$), 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 4.2.3.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

Die Meta-Analyse ist ein statistisches Verfahren um Ergebnisse mehrerer Studien, welche die gleiche Frage bearbeiten, quantitativ zu einem Ergebnis zusammenzufassen und dadurch die Aussagekraft (Genauigkeit der Effektschätzer) gegenüber Einzelstudien zu erhöhen. Im vorliegenden Anwendungsgebiet liegen für die Induktionsphase insgesamt drei Studien vor: Die Substudie 1 und die Substudie 2 der Studie U-ACHIEVE (M14-234), sowie die Studie U-ACCOMPLISH (M14-675). Für die Darstellung der Induktionsphase sollen diese drei Induktionsstudien meta-analytisch zusammengefasst werden, um gepoolte Effektschätzer dieser Studien zu erhalten.

Vor Durchführung der Meta-Analyse werden die genannten Studien auf Vergleichbarkeit hinsichtlich der Fragestellung, des Studiendesigns, der Studienpopulation, der Intervention und der Endpunkte geprüft.

Im Hinblick auf die Faktoren zum Studiendesign, der Operationalisierung der Endpunkte oder der Ein- und Ausschlusskriterien der zugrundeliegenden, zulassungsrelevanten Induktionsstudien kann von einer hinreichenden Ähnlichkeit der Studien ausgegangen werden. Alle drei Studien wurden als Induktionsstudien für die nachfolgende Erhaltungsstudie designt und basieren auf vergleichbaren Studienprotokollen mit identischen Ein- und Ausschlusskriterien sowie gleichen Operationalisierungen der Endpunkte. Hinsichtlich der Fragestellung und Intervention sind die Studien ebenso vergleichbar, da diese die Sicherheit und Wirksamkeit von Upadacitinib in der zugelassenen Dosierung von 45 mg im Vergleich zu Placebo untersuchten. Bezüglich der Patientenpopulation schlossen die Studien erwachsene Patienten mit mittelschwerer bis schwerer aktiver CU ein. Es ist davon auszugehen, dass beide Studien auch hinsichtlich der Patientenpopulation homogen sind.

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

Um dies zu untersuchen, werden für die Populationen der Induktionsstudien sowohl die demografischen als auch die krankheitsspezifischen Charakteristika, die zu Studienbeginn erhoben wurden, sowie die Vorbehandlungen, die zu Studienbeginn erhoben wurden, dargestellt und es wird auf Unterschiede zwischen den Studien in den einzelnen Merkmalen eingegangen (siehe Abschnitt 4.3.2.1.2).

Methodik der Meta-Analysen

Die im Folgenden beschriebene Meta-Analyse basiert auf den aggregierten Daten zu den Effektschätzern aus den Studienberichten der Induktionsstudien (1).

Da die relevanten Studien auf vergleichbaren Protokollen basierten, sind, aus klinischer und methodischer Sicht, keine systematischen Unterschiede in den Patientencharakteristika, Interventionen oder Endpunkten zwischen den Studien als mögliche Ursachen für Heterogenität anzunehmen.

Für Endpunkte, die in allen zugrundeliegenden Studien erhoben wurden, werden, soweit eine Zusammenfassung der Ergebnisse sinnvoll ist (im Hinblick auf die Vergleichbarkeit der Studien und die Einschätzung der statistischen Heterogenität), meta-analytische Effektschätzer berechnet. Für binäre Endpunkte oder stetige Endpunkte mit adäquater Responseschwelle werden RR, OR und RD mit zugehörigem 95 %-KI berechnet. Für stetige Endpunkte werden die Ergebnisse mittels meta-analytisch berechnetem Hedges‘ g und zugehörigem 95 %-KI präsentiert, welche auch für die Einschätzung der Relevanz der Ergebnisse herangezogen wird.

Für die Einschätzung der statistischen Heterogenität zwischen den Studien wird sowohl das Ergebnis des p-Wertes des statistischen Tests auf Heterogenität (Q-Test) als auch eine quantitative Darstellung in Form des I^2-Maßes dargestellt. Liegt der p-Wert des Heterogenitätstests unter 0,05 und liegen keine deutlich gleichgerichteten Effekte vor, werden nur Einzelergebnisse der Studien ohne die gepoolten Analysen interpretiert (55). Weiterhin werden die Ergebnisse der einzelnen Endpunkte der Einzelstudien als auch die meta-analytisch zusammengefassten Ergebnisse mit zugehörigen Konfidenzintervallen in Form von Forest Plots visualisiert. Für die Nutzenbewertung liegen nur sehr wenige Studien vor, die meta-analytisch zusammengefasst werden können (weniger als fünf). Gemäß dem IQWiG-Methodenpapier 6.1 (55) wird in einem solchen Fall das Modell mit festem Effekt gewählt.

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.

Im vorliegenden Dossier werden keine Sensitivitätsanalysen dargestellt.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es 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 (zu ergänzenden Kriterien zur Darstellung siehe Abschnitt 4.3.1.3.2).

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 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. 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.

Subgruppenanalysen dienen der Identifizierung von möglichen Effektmodifikatoren und sind hilfreich bei der Identifikation von Patientengruppen mit unterschiedlichen Nutzenprofilen. Die Beurteilung der zu bewertenden Studienmedikation gegenüber einer ZVT kann nur im Rahmen eines indirekten Vergleichs erfolgen. Eine valide Untersuchung möglicher Effektmodifikatoren setzt voraus, dass sowohl für Upadacitinib als auch für die ZVT gleiche Subgruppen definiert und operationalisiert wurden. Darüber hinaus besteht die Problematik, dass potenzielle Subgruppeneffekte nicht zwangsläufig eine Effektmodifikation von Upadacitinib gegenüber einer ZVT beschreiben, sondern die Subgruppeneffekte aus den Einzelstudien aber nicht aus dem indirekten Vergleich resultieren. Zusätzlich werden die zu vergleichenden Populationen durch Subgruppenanalysen verkleinert, was im Rahmen eines indirekten Vergleichs zu einer geringeren Aussagekraft der Ergebnisse führt.

Aus diesem Grund wird im Folgenden auf die Darstellung von Subgruppenanalysen verzichtet.

Allerdings werden, soweit verfügbar, deskriptive Subgruppenanalysen für die Subgruppe Bio-IR-Status in die Erhaltungsphase präsentiert, um eine Einschätzung der Ergebnisse hinsichtlich bio-naiven und bio-erfahrenen Patienten zu ermöglichen.

Eine deskriptive Darstellung der Subgruppenanalysen für die Subgruppe Bio-IR-Status findet sich im Abschnitt 4.3.2.1.4.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Es besteht dabei internationaler Konsens, dass Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator (häufig als nicht adjustierte indirekte Vergleiche bezeichnet) regelhaft keine valide Analysemethode darstellen[8]. Eine Ausnahme kann das Vorliegen von dramatischen Effekten sein. An Stelle von nicht adjustierten indirekten Vergleichen sollen je nach Datenlage einfache adjustierte indirekte Vergleiche[9] oder komplexere Netzwerk-Meta-Analysen (auch als „Mixed Treatment Comparison [MTC] Meta-Analysen“ oder „Multiple Treatment Meta-Analysen“ bezeichnet) für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung

[8] Bender R, Schwenke C, Schmoor C, Hauschke D. Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR [online]. [Zugriff: 31.10.2016]. URL: http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf.
[9] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50(6): 683-691.

sowohl direkter als auch indirekter Vergleiche berechnet werden. Aktuelle Verfahren wurden beispielsweise von Lu und Ades (2004)[10] und Rücker (2012)[11] vorgestellt.

Alle Verfahren für indirekte Vergleiche gehen im Prinzip von den gleichen zentralen Annahmen aus. Hierbei handelt es sich um die Annahmen der Ähnlichkeit der eingeschlossenen Studien, der Homogenität der paarweisen Vergleiche und der Konsistenz zwischen direkter und indirekter Evidenz 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[12].

Das Ergebnis eines indirekten Vergleichs kann maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen. Als Brückenkomparatoren sind dabei insbesondere Interventionen zu berücksichtigen, für die sowohl zum bewertenden Arzneimittel als auch zur zweckmäßigen Vergleichstherapie mindestens eine direkt vergleichende Studie vorliegt (Brückenkomparatoren ersten Grades).

Insgesamt ist es notwendig, die zugrunde liegende Methodik für alle relevanten Endpunkte genau und reproduzierbar zu beschreiben und die zentralen Annahmen zu untersuchen[13, 14, 15]

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

- *Benennung aller potentiellen Brückenkomparatoren ersten Grades und ggf. Begründung für die Auswahl.*
- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayes'schen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten, die Art der Untersuchung der Konvergenz der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Ähnlichkeit der eingeschlossenen Studien.*

[10] Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.
[11] Rücker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods 2012; 3(4): 312-324.
[12] Schöttker B, Lühmann D, Boulkhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.
[13] 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.
[14] Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study BMJ 2011; 343 :d4909
[15] Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. Res Synth Methods 2013; 4(4): 291-323.

- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenzannahme im Netzwerk.*
- *Bilden Sie den Code des Computerprogramms inklusive der einzulesenden Daten in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Da keine direkt vergleichende Studie zwischen Upadacitinib und einer der festgelegten ZVT identifiziert werden konnte, erfolgt die Bewertung des Zusatznutzens anhand eines adjustierten indirekten Vergleichs nach Bucher (68). Die Behandlung mit Upadacitinib wird im indirekten Vergleich der Behandlung mit Ustekinumab aus randomisierten placebokontrollierten Studien gegenübergestellt. In einer systematischen Literaturrecherche konnte mit der Studie UNIFI eine RCT identifiziert werden, die über den Brückenkomparator Placebo einen indirekten Vergleich zwischen dem zu bewertenden Arzneimittel und der ZVT-Option Ustekinumab zulässt.

Eine schematische Skizzierung der Netzwerkstruktur dieses indirekten Vergleichs wird in der Abbildung 1 dargestellt.

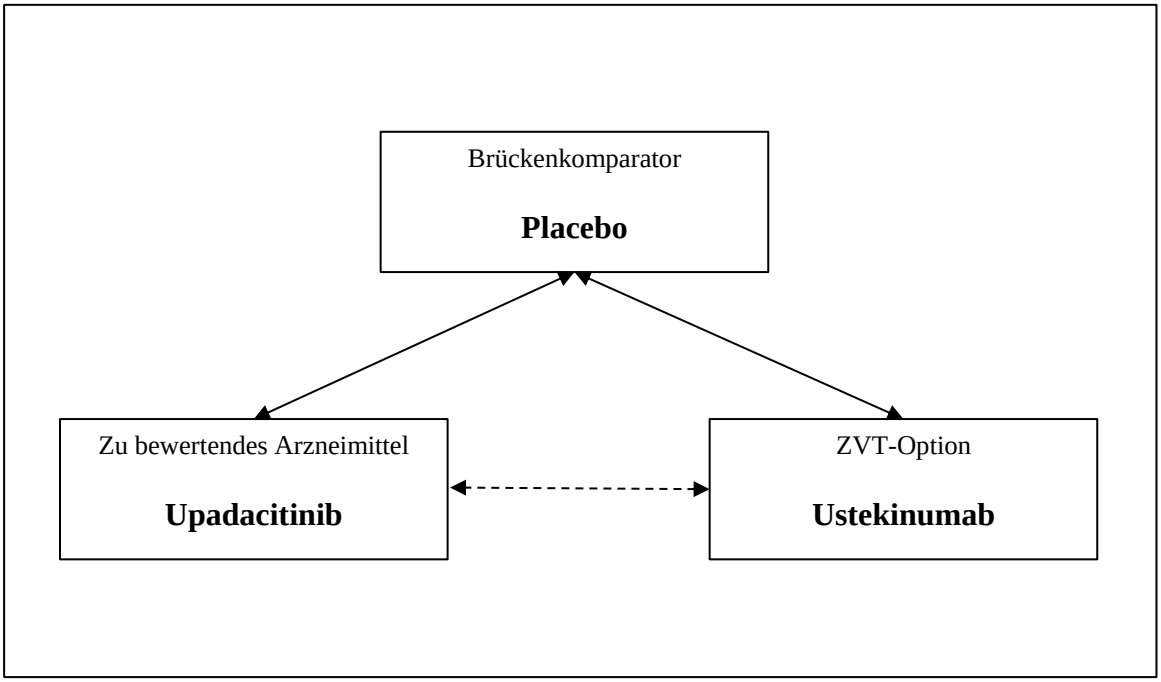

Abbildung 1: Netzwerkstruktur eines indirekten Vergleichs zwischen Upadacitinib und Ustekinumab über den Brückenkomparator Placebo

Die methodische Basis eines adjustierten indirekten Vergleichs von verschiedenen Behandlungsmöglichkeiten bildet ein Brückenkomparator über einen gemeinsamen

Vergleichsarm, anhand dessen sich die Ergebnisse zur Wirksamkeit und Sicherheit beider Interventionen quantifizieren lassen (69).

Im vorliegenden Dossier wird ein einfacher adjustierter indirekter Vergleich nach Bucher durchgeführt (68).

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

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

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

Das 95 %-KI für $\ln \theta_{BC}$ berechnet sich als

$$\ln \theta_{BC} \pm 1{,}96 \times \sqrt{\text{Var} (\ln \theta_{BC})} \quad (3)$$

Die in (1) und (3) beschriebenen Effektmaße θ gelten für den jeweiligen Schätzer des Effektmaßes (RR), wobei die Log-Transformation im Falle kontinuierlicher, normalverteilter Variablen für die Berechnung von Effektschätzer (Hedges‘ g) beziehungsweise dessen Varianz entfällt.

Als zentrale Annahmen für einen einfachen indirekten Vergleich gelten die Ähnlichkeits- und die Homogenitätsannahme. Die Ähnlichkeitsannahme beschreibt eine möglichst gute Vergleichbarkeit der zu berücksichtigenden Studien hinsichtlich möglicher Effektmodifikatoren wie Studien- und Patientencharakteristika. Die Homogenitätsannahme steht für möglichst gleichgerichtete Effekte und damit verbundenen homogenen Studienergebnissen (70). Eine detaillierte Beschreibung der für die Annahmen relevanten Studien- sowie Patientencharakteristika (Demografie, Krankheitsmerkmale und medizinische Vorgeschichte) findet sich im Abschnitt 4.3.2.1.2.1.

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)

Falls für die Bewertung des Zusatznutzens mehrere Komparatoren (z.B. Wirkstoffe) herangezogen werden, sind die Aussagen zum Zusatznutzen primär gegenüber der Gesamtheit der gewählten Komparatoren durchzuführen (z. B. basierend auf Meta-Analysen unter gemeinsamer Betrachtung aller direkt vergleichender Studien). Spezifische methodische Argumente, die gegen eine gemeinsame Analyse sprechen (z. B. statistische oder inhaltliche Heterogenität), sind davon unbenommen. Eine zusammenfassende Aussage zum Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie ist in jedem Fall erforderlich.

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, Angabe zu geplanten und durchgeführten Datenschnitten und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-10: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	**Zulassungs-studie (ja/nein)**	**Sponsor (ja/nein)**	**Status (abgeschlossen / abgebrochen / laufend)**	**Studiendauer ggf. Datenschnitt**	**Therapiearme**
U-ACCOMPLISH (M14-675)	ja	ja	abgeschlossen	8 Wochen (Induktionsphase)	Upadacitinib 45 mg QD Placebo
U-ACHIEVE (M14-234)	ja	ja	abgeschlossen	8 Wochen (Induktionsphase)	Upadacitinib 45 mg QD Placebo
				52 Wochen (Erhaltungsphase)	Upadacitinib 15 mg/30 mg QD Placebo
QD: Einmal täglich; RCT: Randomisierte kontrollierte Studie					

Geben Sie an, welchen Stand die Information in Tabelle 4-10 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.

Aufgelistet sind alle RCTs mit finanzieller Beteiligung des pharmazeutischen Unternehmers für das zu bewertende Arzneimittel, die in der Indikation laufend sind oder durchgeführt wurden (Stand: 08. Juli 2022).

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-10 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-11: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
U-ACCOMPLISH (M14-675)	Kein Vergleich zur ZVT (placebokontrolliert)
U-ACHIEVE (M14-234)	Kein Vergleich zur ZVT (placebokontrolliert)
RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie	

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.

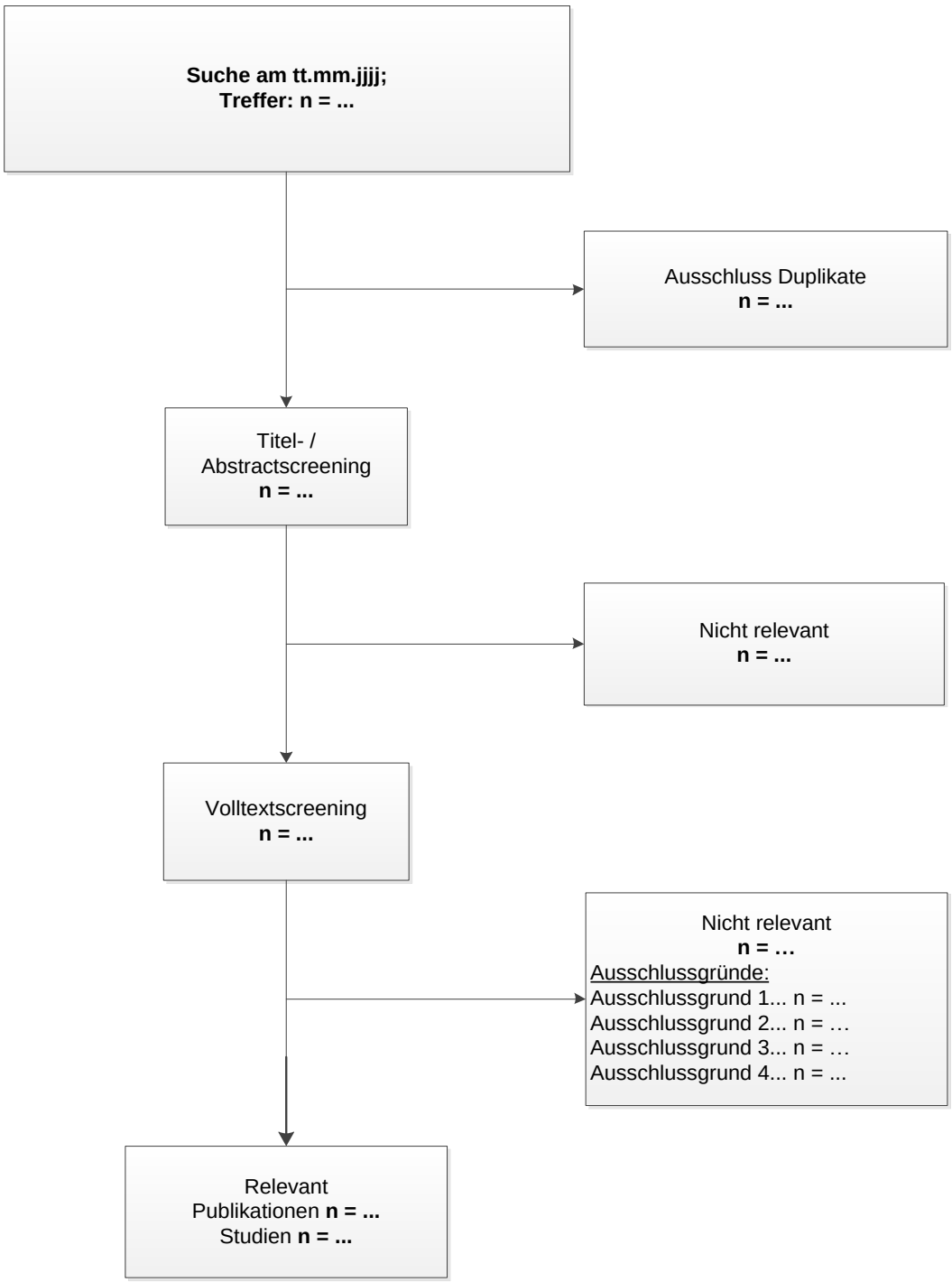

Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die bibliografische Literaturrecherche zu RCTs mit Upadacitinib fand am 02. Juni 2022 statt und erzielte insgesamt 221 Treffer. Nach Duplikatausschluss (n = 77) verblieben 144 Publikationen, welche unter Berücksichtigung der Ein- und Ausschlusskriterien (Abschnitt 4.2.2) gesichtet wurden. Im Zuge des Titel- und Abstractscreenings wurden

143 Publikationen als nicht relevant eingestuft. Eine Publikation wurde im Volltext gesichtet und anschließend ausgeschlossen (Abbildung 3; Anhang 4-C).

Es wurden im Rahmen der bibliografischen Literaturrecherche keine relevanten RCTs mit dem zu bewertenden Arzneimittel im Anwendungsgebiet identifiziert.

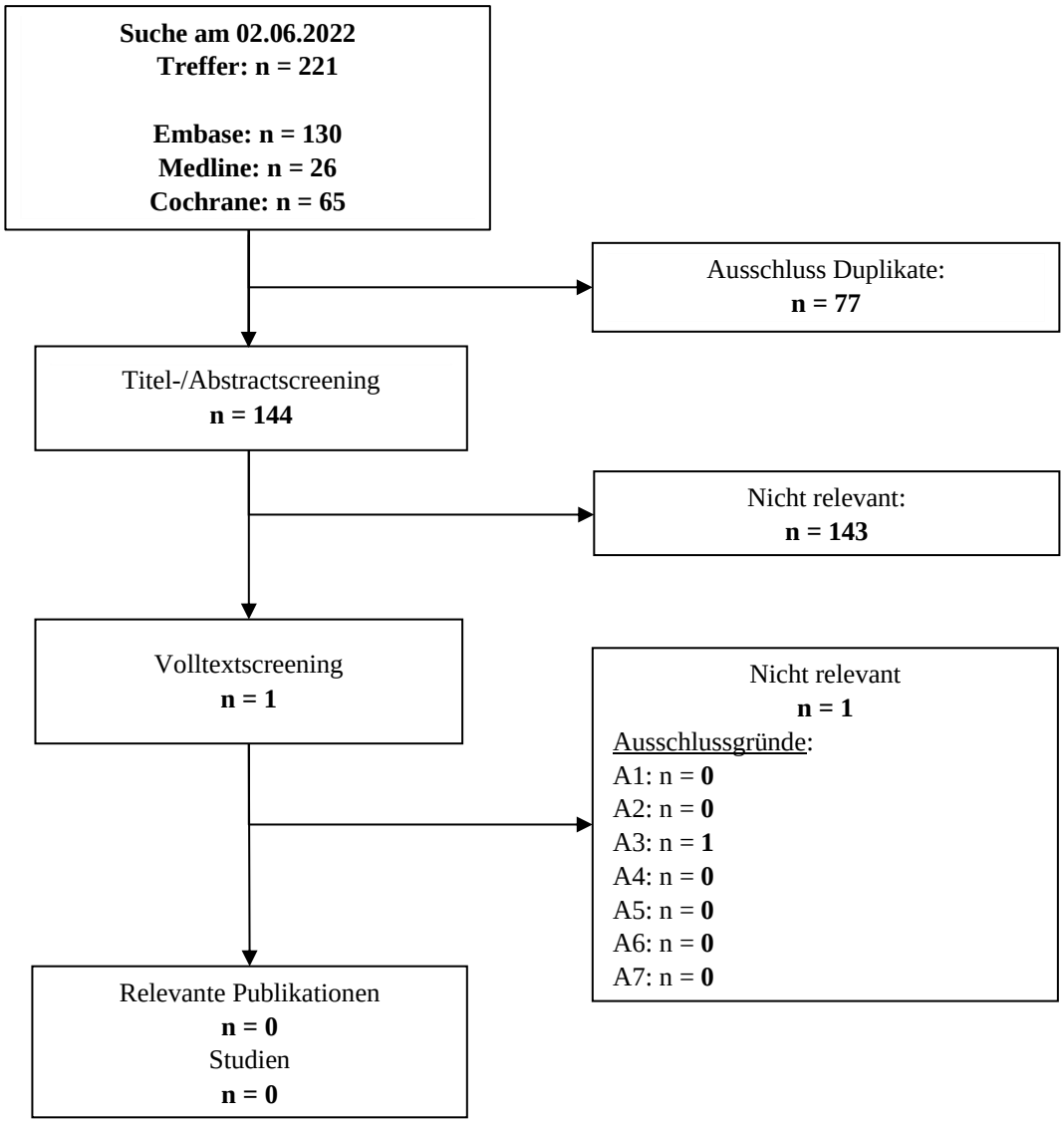

Abbildung 3: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank 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 4-10) 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 4-12: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern/ Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	**Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank 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)**
Keine Studie identifiziert.				

a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

RCT: Randomisierte kontrollierte Studie

Geben Sie an, welchen Stand die Information in Tabelle 4-12 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 Studienregistersuche ergab keine relevante RCT mit dem zu bewertenden Arzneimittel, die einen direkten Vergleich mit der ZVT erlaubt.

Stand der Recherche ist der 02. Juni 2022.

4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-10) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/ Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

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

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

Tabelle 4-13: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	**Relevante Quellen[a]**	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)**	**Studie durch Suche in Studienregistern /Studienergebnisdatenbanken identifiziert (ja/nein)**
Keine Studie identifiziert.				

a: Quellen aus der Suche auf der Internetseite des G-BA
G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie

Geben Sie an, welchen Stand die Information in Tabelle 4-13 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.

Es wurden keine relevanten Studien mit Upadacitinib im Anwendungsgebiet identifiziert (Stand: 02. Juni 2022).

4.3.1.1.5 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, 4.3.1.1.3 und 4.3.1.1.4) 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 aller verfügbaren Quellen in Abschnitt 4.3.1.4 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 4-14: 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)	Studienberichte (ja/nein [Zitat])	Register-einträge[c] (ja/nein [Zitat])	Publikation und sonstige Quellen[d] (ja/nein [Zitat])
placebokontrolliert						
Keine Studie identifiziert.						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Keine Studie identifiziert.						

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 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (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: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.

G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie

Es wurden keine für die Nutzenbewertung relevanten, direkt vergleichenden RCT-Studien mit dem zu bewertenden Arzneimittel und der ZVT identifiziert.

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. Falls Teilpopulationen berücksichtigt werden, ist die Charakterisierung der Studienpopulation auch für diese Teilpopulation durchzuführen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Geben Sie bei den Datenschnitten auch den Anlass des Datenschnittes an. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.

Es wurden keine für die Nutzenbewertung relevanten, direkt vergleichenden RCT-Studien mit dem zu bewertenden Arzneimittel und der ZVT identifiziert. Die Darstellung der Studien erfolgt im Rahmen eines indirekten Vergleichs in Abschnitt 4.3.2.1.

Tabelle 4-15: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	**Studiendesign** <RCT, doppelblind/einfach, verblindet/offen, parallel/cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer/ Datenschnitte** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
Nicht zutreffend.						
RCT: Randomisierte kontrollierte Studie						

Tabelle 4-16: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	**<Gruppe 1>**	**<Gruppe 2>**	***ggf. weitere Spalten mit Behandlungscharakteristika*** ***z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
Nicht zutreffend.			
RCT: Randomisierte kontrollierte Studie			

Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie **Gruppe**	**N**	**Alter (Jahre)**	**Geschlecht w/m (%)**	***ggf. weitere Spalten mit Populationscharakteristika*** ***z. B. Dauer der Erkrankung, Schweregrad, Therapieabbrecher, Studienabbrecher, weitere Basisdaten projektabhängig***
Nicht zutreffend.				
RCT: Randomisierte kontrollierte Studie				

Beschreiben Sie die Studien zusammenfassend. In der Beschreibung der Studien sollten Informationen zur Behandlungsdauer sowie zu geplanter und tatsächlicher Beobachtungsdauer enthalten sein. Sofern sich die Beobachtungsdauer zwischen den relevanten Endpunkten unterscheidet, sind diese unterschiedlichen Beobachtungsdauern endpunktbezogen anzugeben. Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich. 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.

Es wurden keine für die Nutzenbewertung relevanten, direkt vergleichenden RCT-Studien mit dem zu bewertenden Arzneimittel und der ZVT identifiziert.

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 4-18: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
Nicht zutreffend.							
RCT: Randomisierte kontrollierte Studie							

Begründen Sie für jede Studie die abschließende Einschätzung.

Es wurden keine für die Nutzenbewertung relevanten, direkt vergleichenden RCT-Studien mit dem zu bewertenden Arzneimittel und der ZVT identifiziert.

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 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität	Morbidität	Gesundheitsbezogene Lebensqualität	Verträglichkeit
Nicht zutreffend.				
RCT: Randomisierte kontrollierte Studie				

Es wurden keine für die Nutzenbewertung relevanten, direkt vergleichenden RCT-Studien mit dem zu bewertenden Arzneimittel und der ZVT identifiziert.

4.3.1.3.1 <Endpunkt xxx> – 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 inkl. Angaben zur Häufigkeit von und zum Umgang mit nicht oder nicht vollständig beobachteten Patienten (bei Verlaufsbeobachtungen pro Messzeitpunkt)
– 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.

Unterschiedliche Beobachtungszeiten zwischen den Behandlungsgruppen sollen durch adäquate Analysen (z.B. Überlebenszeitanalysen) adressiert werden, und zwar für alle Endpunkte (einschließlich UE nach den nachfolgend genannten Kriterien), für die eine solche Analyse aufgrund deutlich unterschiedlicher Beobachtungszeiten erforderlich ist.

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden. Dabei ist für jeden Endpunkt, für den eine solche Analyse durchgeführt wird, eine separate Kaplan-Meier-Kurve darzustellen.

Zu mit Skalen erhobenen patientenberichteten Endpunkten (z.B. zur gesundheitsbezogenen Lebensqualität oder zu Symptomen) sind immer auch die Werte im Studienverlauf anzugeben, auch als grafische Darstellung, sowie eine Auswertung, die die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt (z.B. als Symptomlast über die Zeit, geschätzt mittels MMRM-Analyse [falls aufgrund der Datenlage geeignet]). Die Auswertung von Responderanalysen mittels klinischer Relevanzschwellen bei komplexen Skalen soll nach dem folgenden Vorgehen erfolgen:

1. Falls in einer Studie Responderanalysen unter Verwendung einer MID präspezifiziert sind und das Responsekriterium mindestens 15 % der Skalenspannweite des verwendeten

Erhebungsinstruments entspricht, sind diese Responderanalysen für die Bewertung darzustellen.

2. Falls präspezifiziert Responsekriterien im Sinne einer MID unterhalb von 15 % der Skalenspannweite liegen, bestehen in diesen Fällen und solchen, in denen gar keine Responsekriterien präspezifiziert wurden, aber stattdessen Analysen kontinuierlicher Daten zur Verfügung stehen, verschiedene Möglichkeiten. Entweder können post hoc spezifizierte Analysen mit einem Responsekriterium von genau 15 % der Skalenspannweite dargestellt werden. Alternativ können Analysen der kontinuierlichen Daten dargestellt werden, für die Relevanzbewertung ist dabei auf ein allgemeines statistisches Maß in Form von standardisierten Mittelwertdifferenzen (SMDs, in Form von Hedges' g) zurückzugreifen. Dabei ist eine Irrelevanzschwelle als Intervall von - 0,2 bis 0,2 zu verwenden: Liegt das zum Effektschätzer korrespondierende Konfidenzintervall vollständig außerhalb dieses Irrelevanzbereichs, wird davon ausgegangen, dass die Effektstärke nicht in einem sicher irrelevanten Bereich liegt. Dies soll gewährleisten, dass der Effekt hinreichend sicher mindestens als klein angesehen werden kann.

3. Liegen sowohl geeignete Responderanalysen (Responsekriterium präspezifiziert mindestens 15 % der Skalenspannweite oder post hoc genau 15 % der Skalenspannweite) als auch Analysen stetiger Daten vor, sind die Responderanalysen darzustellen.

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen:

1. Gesamtrate UE,

2. Gesamtrate schwerwiegender UE (SUE),

3. Gesamtrate der Abbrüche wegen UE,

4. Gesamtraten von UE differenziert nach Schweregrad, sofern dies in der/den relevante/n Studie/n erhoben wurde (z.B. gemäß CTCAE und/oder einer anderen etablierten bzw. validierten indikationsspezifischen Klassifikation) einschließlich einer Abgrenzung schwerer und nicht schwerer UE,

5. zu den unter 1, 2 und 4 genannten Kategorien (UE ohne weitere Differenzierung, SUE, UE differenziert nach Schweregrad) soll zusätzlich zu den Gesamtraten die Darstellung nach Organsystemen und Einzelereignissen (als System Organ Class [SOCs] und Preferred Terms [PT] nach MedDRA) jeweils nach folgenden Kriterien erfolgen:

- UE (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10% der Patienten in einem Studienarm aufgetreten sind

- Schwere UE (z.B. CTCAE-Grad ≥ 3) und SUE: Ereignisse, die bei mindestens 5% der Patienten in einem Studienarm aufgetreten sind

- zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patienten UND bei mindestens 1 % der Patienten in einem Studienarm aufgetreten sind.

6. A priori definierte UE von besonderem Interesse [AESI]) sowie prädefinierte SOC-übergreifende UE-Auswertungen (z.B. als Standardised MedDRA Queries, SMQs) sollen unabhängig von der Ereignisrate dargestellt werden und zwar differenziert nach Schweregrad (dargestellt als Gesamtrate und differenziert nach Schweregrad, nicht schwer, schwer, schwerwiegend).

7. zu Kategorie 3: Die Abbruchgründe auf SOC/PT-Ebene müssen vollständig, jedoch nur deskriptiv dargestellt werden.

Sofern bei der Erhebung unerwünschter Ereignisse erkrankungsbezogenen Ereignisse (z. B. Progression, Exazerbation) berücksichtigt werden (diese Ereignisse also in die UE-Erhebung eingehen), sollen für die Gesamtraten (UE, schwere UE und SUE) zusätzliche UE-Analysen durchgeführt werden, bei denen diese Ereignisse unberücksichtigt bleiben. Alle Auswertungen zu UE können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine Darstellung ausschließlich in Modul 5 ist nicht ausreichend. Davon unbenommen sind die Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE), sowie die für die Gesamtaussage zum Zusatznutzen herangezogenen Ergebnisse im vorliegenden Abschnitt darzustellen.

Auswertungen zu den im Abschnitt 4.3.1.2.1 aufgeführten Datenschnitten sollen vollständig, d.h. für alle erhobenen relevanten Endpunkte, durchgeführt und vorgelegt werden. Das gilt auch dann wenn ein Datenschnitt ursprünglich nur zur Auswertung einzelner Endpunkte geplant war. Auf die Darstellung der Ergebnisse einzelner Endpunkte eines Datenschnitts bzw. eines gesamten Datenschnitts kann verzichtet werden, wenn hierdurch kein wesentlicher Informationsgewinn gegenüber einem anderen Datenschnitt zu erwarten ist (z. B. wenn die Nachbeobachtung zu einem Endpunkt bereits zum vorhergehenden Datenschnitt nahezu vollständig war oder ein Datenschnitt in unmittelbarer zeitlicher Nähe zu einem anderen Datenschnitt liegt).

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 4-20: 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 4-21: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> 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**
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

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

Tabelle 4-22: Ergebnisse für <Endpunkt xxx> aus RCT mit dem zu bewertenden Arzneimittel

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

Nicht zutreffend.

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.

Nicht zutreffend.

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

4.3.1.3.2 Subgruppenanalysen – RCT

Für die Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1. [16]

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

- Subgruppenanalysen sind nur für die Merkmale (z.B. Alter) durchzuführen, bei denen die resultierenden Subgruppen jeweils mindestens 10 Patienten umfassen.

[16] unbesetzt

- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.

- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ($p < 0,05$) dargestellt werden.

- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.

- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.

- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmodifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.

- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.

- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen. Stellen Sie dabei zunächst tabellarisch dar, zu welchen der in Abschnitt 4.2.5.5 genannten Effektmodifikatoren Subgruppenanalysen zu den relevanten Endpunkten vorliegen, und ob diese a priori geplant und im Studienprotokoll festgelegt waren oder posthoc durchgeführt wurden.

Orientieren Sie sich an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-23 Matrix der durchgeführten Subgruppenanalysen

Endpunkt Studie	**Alter**	**Geschlecht**	**<Effektmo-difikator-a>**	**<Effektmo-difikator-b>**	**<Effektmo-difikator-c>**	**<Effektmo-difikator-d>**
Gesamtmortalität						
<Studie 1>	●	●	●	○	○	○
<Studie 2>	●	●	○	n.d.	n.d.	n.d.
<Endpunkt 2>						
...						
●: A priori geplante Subgruppenanalyse. ○: Post hoc durchgeführte Subgruppenanalyse. n.d.: Subgruppenanalyse nicht durchgeführt.						

Stellen Sie anschließend in Tabelle 4-24 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante (p < 0,05) Interaktionsterme.

Tabelle 4-24: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für <Studie> und <Effektmodifikator>

Endpunkt Studie	**Alter**	**Geschlecht**	**<Effektmo-difikator-a>**	**<Effektmo-difikator-b>**	**<Effektmo-difikator-c>**	**<Effektmo-difikator-d>**
Gesamtmortalität						
<Studie 1>	p=0,345	p=0,321	**p=0,003**	**p=0,041**	p=0,981	p=0,212
<Studie 2>	p=0,634	p=0,212	**p<0,001**	k. A.	k. A.	k. A.
<Endpunkt 2>						
...						
k. A.: Keine Angabe						

Stellen Sie schließlich alle Subgruppenergebnisse dar.

Sofern eine Effektmodifikation für mehr als ein Subgruppenmerkmal vorliegt, kann eine Untersuchung auf eine Wechselwirkung höherer Ordnung sinnvoll sein. Dies gilt insbesondere dann, wenn diese Effektmodifikation konsistent über mehrere Endpunkte besteht. Zur Interpretation der Ergebnisse sollte dann für diese Endpunkte zusätzlich eine Subgruppenanalyse durchgeführt werden, die die Merkmale mit Effektmodifikation kombiniert. Beispiel: Für die Endpunkte Mortalität, gesundheitsbezogene Lebensqualität und schwere unerwünschte Ereignisse liegt sowohl für das Merkmal Geschlecht (mit den Ausprägungen „weiblich“ und „männlich“) als auch für das Merkmal Schweregrad (mit den Ausprägungen „niedrig“ und „hoch“) eine Effektmodifikation vor. Die zusätzliche Subgruppenanalyse erfolgt dann für die 3 genannten Endpunkte für das kombinierte Merkmal Geschlecht/Schweregrad mit den 4 Ausprägungen weiblich/niedrig, weiblich/hoch, männlich/niedrig und männlich/hoch.

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.

Nicht zutreffend.

4.3.1.4 Liste der eingeschlossenen Studien – RCT

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).

Nicht zutreffend.

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 sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der Internetseite des G-BA*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

4.3.2.1.1.1 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, Angabe zu geplanten und durchgeführten Datenschnitten und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-25: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	**Zulassungs-studie (ja/nein)**	**Sponsor (ja/nein)**	**Status (abgeschlossen / abgebrochen / laufend)**	**Studiendauer ggf. Datenschnitt**	**Therapiearme**
U-ACCOMPLISH (M14-675)	ja	ja	abgeschlossen	8 Wochen (Induktionsphase)	Upadacitinib 45 mg QD Placebo
U-ACHIEVE (M14-234)	ja	ja	abgeschlossen	8 Wochen (Induktionsphase)	Upadacitinib 45 mg QD Placebo
				52 Wochen (Erhaltungsphase)	Upadacitinib 15 mg/30 mg QD Placebo
QD: Einmal täglich; RCT: Randomisierte kontrollierte Studie					

Geben Sie an, welchen Stand die Information in Tabelle 4-25 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.

Stand der Information in Tabelle 4-25 ist der 08. Juli 2022.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-25 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-26: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
Nicht zutreffend.	
RCT: Randomisierte kontrollierte Studie	

4.3.2.1.1.2 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.

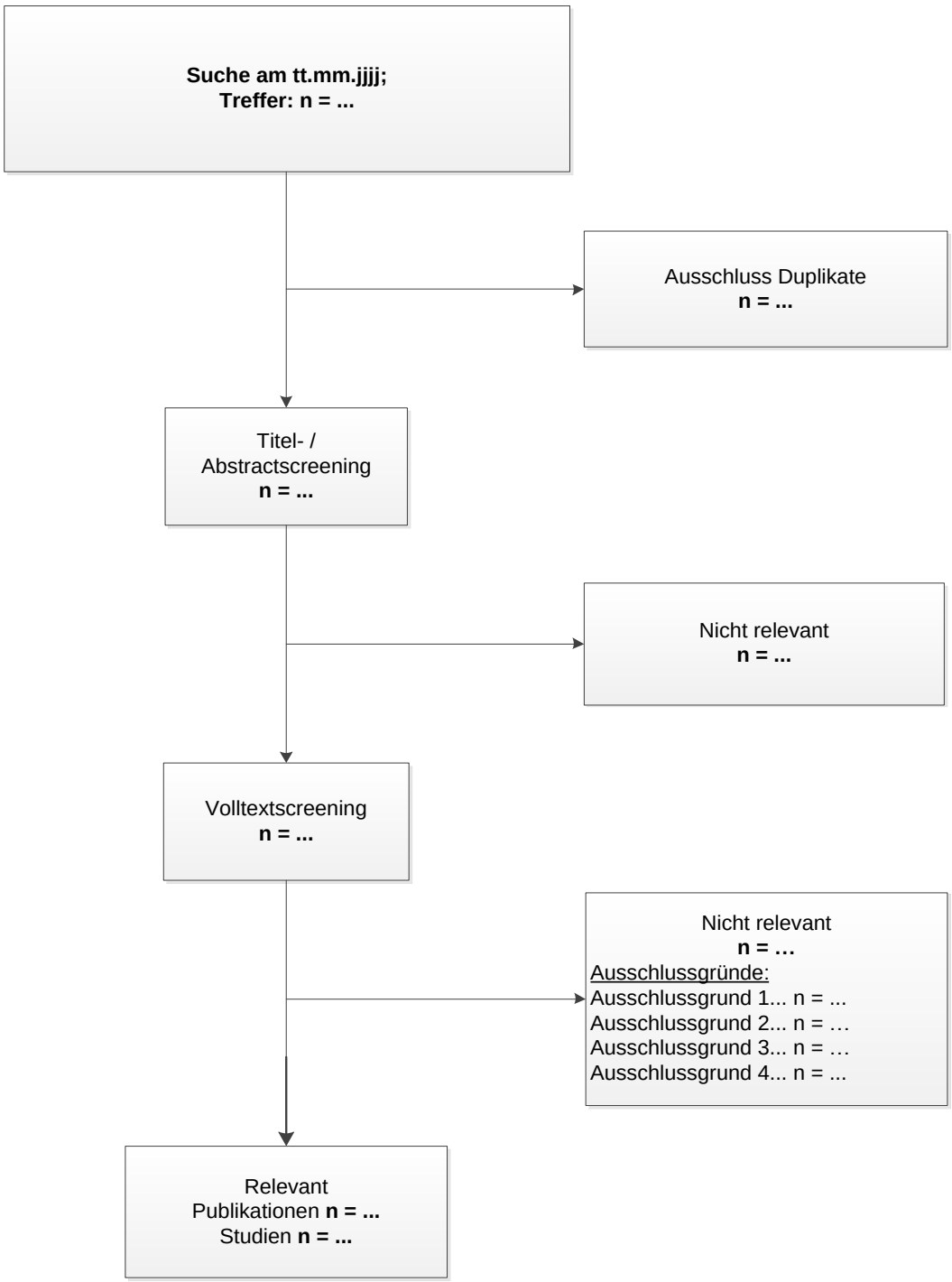

Abbildung 4: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die bibliografische Literaturrecherche zu RCTs für indirekte Vergleiche mit Upadacitinib fand am 02. Juni 2022 statt und erzielte insgesamt 221 Treffer. Nach Duplikatausschluss (n = 77) verblieben 144 Publikationen, welche unter Berücksichtigung der Ein- und Ausschlusskriterien (Abschnitt 4.2.2) gesichtet wurden. Im Zuge des Titel- und Abstractscreenings wurden

141 Publikationen als nicht relevant eingestuft. Drei Publikationen wurden im Volltext gesichtet und eingeschlossen (Abbildung 5).

Es wurde im Rahmen der bibliografischen Literaturrecherche eine RCT (3 Publikationen) mit dem zu bewertenden Arzneimittel identifiziert, die gemäß Fragestellung (Abschnitt 4.2.1) für die Nutzenbewertung herangezogen werden kann.

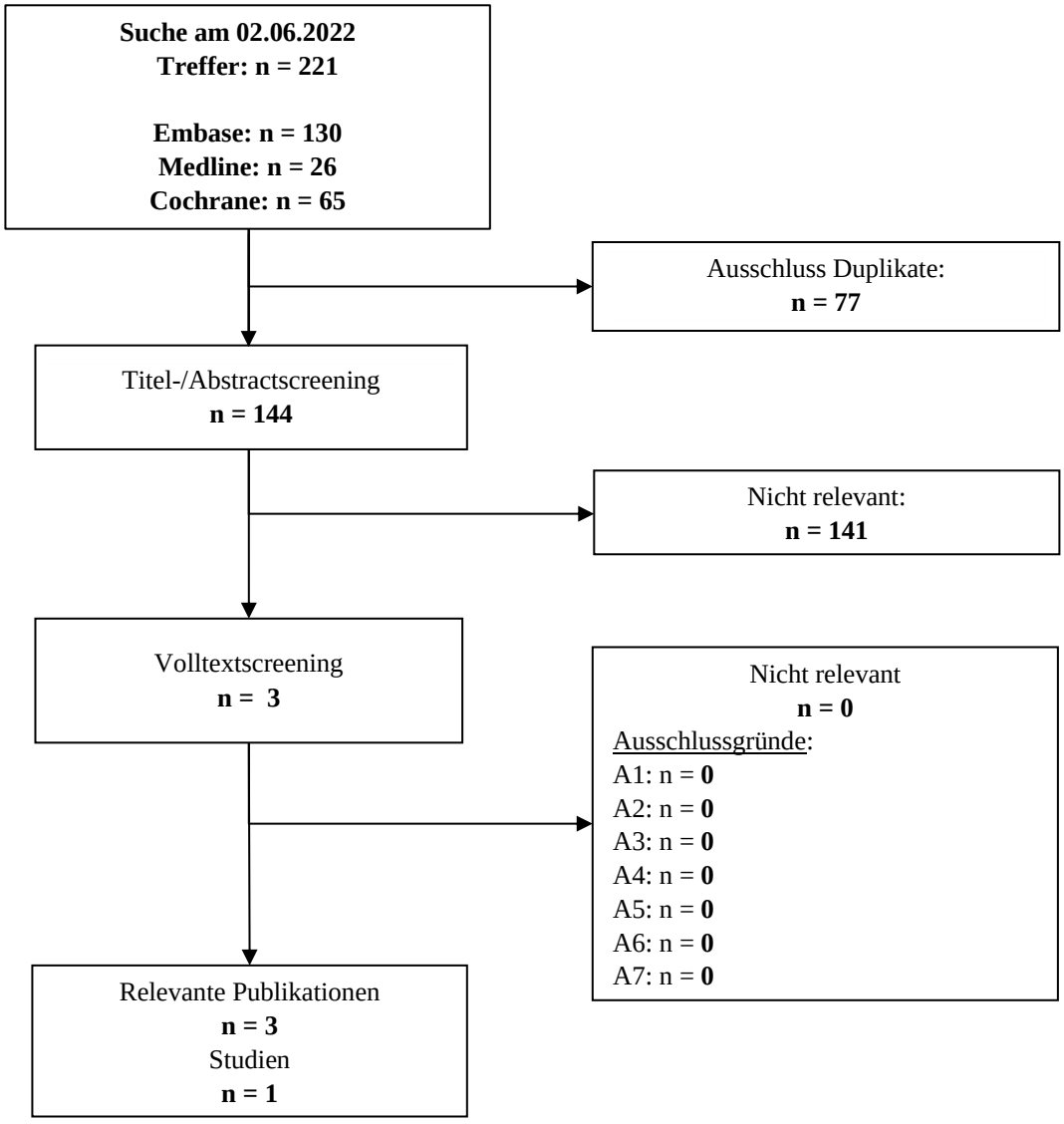

Abbildung 5: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel – indirekte Vergleiche

Die bibliografische Literaturrecherche zu RCTs für indirekte Vergleiche mit der ZVT (Ustekinumab) fand am 02. Juni 2022 statt und erzielte insgesamt 1.136 Treffer. Nach Duplikatausschluss (n = 256) verblieben 880 Publikationen, welche unter Berücksichtigung der Ein- und Ausschlusskriterien (Abschnitt 4.2.2) gesichtet wurden. Im Zuge des Titel- und Abstractscreenings wurden 871 Publikationen als nicht relevant eingestuft. 9 Publikationen wurden im Volltext gesichtet, 8 Publikationen wurden als nicht relevant ausgeschlossen, eine Publikation wurde eingeschlossen (Abbildung 5; Anhang 4-C).

Es wurde im Rahmen der bibliografischen Literaturrecherche eine RCT (1 Publikationen) mit der ZVT (Ustekinumab) identifiziert, die gemäß Fragestellung (Abschnitt 4.2.1) für adjustierte indirekte Vergleiche nach Bucher zum zu bewertenden Arzneimittel für die Nutzenbewertung herangezogen werden kann.

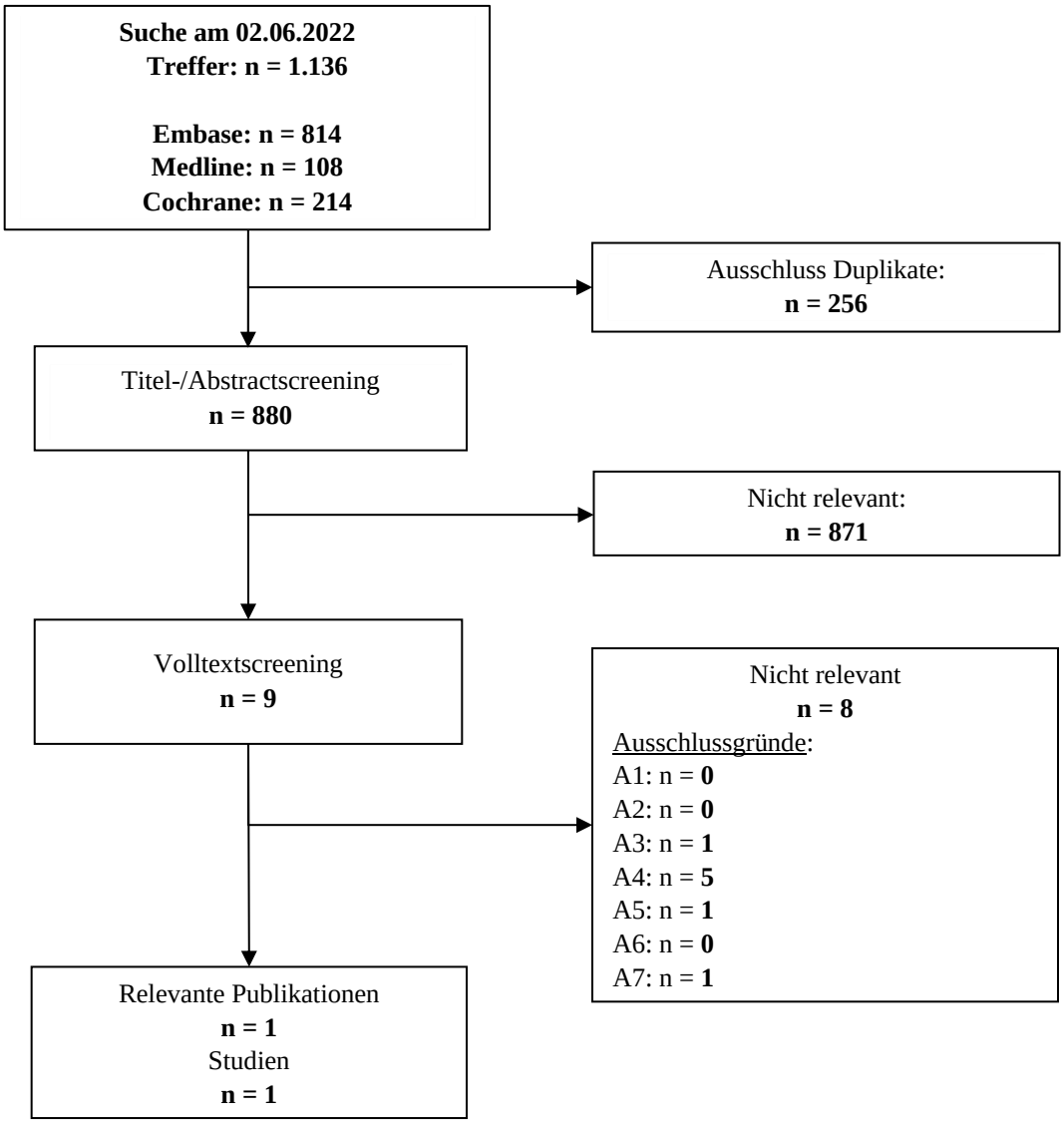

Abbildung 6: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit der ZVT (Ustekinumab) – indirekte Vergleiche

4.3.2.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank 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 4-10) 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 4-27: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern/ Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank 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)
Upadacitinib				
U-ACCOMPLISH (M14-675)	Clinicaltrials.gov (NCT03653026) (71) EudraCT (2016-000642-62) (72) ICTRP (JPRN-JapicCTI-184131; NCT03653026) (73, 74)	ja	nein	abgeschlossen
U-ACHIEVE (M14-234)	Clinicaltrials.gov (NCT02819635) (75) EudraCT (2016-000641-31) (76) ICTRP (NCT02819635) (77)	ja	ja	abgeschlossen
Ustekinumab (ZVT)				
UNIFI	Clinicaltrials.gov (NCT02407236) (78) EudraCT (2014-005606-38) (79) ICTRP (NCT02407236) (80) AMIce (81)	nein	ja	abgeschlossen

Studie	Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank 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)
a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. ICTRP: International Clinical Trials Registry Platform: RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie				

Geben Sie an, welchen Stand die Information in Tabelle 4-27 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 Studienregistersuche ergab zwei relevante RCTs mit dem zu bewertenden Arzneimittel sowie eine relevante RCT mit der ZVT (Ustekinumab) für einen adjustierten indirekten Vergleich nach Bucher über den Brückenkomparator Placebo.

Stand der Recherche ist der 02. Juni 2022.

4.3.2.1.1.4 Studien aus der Suche auf der Internetseite des G-BA

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-10) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/ Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

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

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

Tabelle 4-28: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	Relevante Quellen[a]	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Studie durch Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert (ja/nein)
Upadacitinib				
Es wurden keine relevanten Quellen identifiziert.				
Ustekinumab (ZVT)				
Es wurden keine relevanten Quellen identifiziert.				
a: Quellen aus der Suche auf der Internetseite des G-BA G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie				

Geben Sie an, welchen Stand die Information in Tabelle 4-28 *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.*

Es wurden keine relevanten Quellen auf der Internetseite des G-BA identifiziert.

Der Stand der Information ist der 02. Juni 2022.

4.3.2.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.2.1.1.1, 4.3.2.1.1.2, 4.3.2.1.1.3 und 4.3.2.1.1.4) 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 aller verfügbaren Quellen in Abschnitt 4.3.2.1.5 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 4-29: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen[a]		
	Studie zur Zulassung des zu bewertenden Arznei-mittels (ja/nein)	Gespon-serte Studie[b] (ja/nein)	Studie Dritter (ja/nein)	Studien-berichte (ja/nein [Zitat])	Register-einträge[c] (ja/nein [Zitat])	Publikation und sonstige Quellen[d] (ja/nein [Zitat])
Upadacitinib						
Placebokontrolliert						
U-ACCOMPLISH (M14-675)	ja	ja	nein	ja (82)	ja (71-74)	nein
U-ACHIEVE (M14-234)	ja	ja	nein	ja (83)	ja (75-77)	ja (66, 84, 85)
Ustekinumab (ZVT)						
Placebokontrolliert						
UNIFI	nein	nein	ja	nein	ja (78-80)	ja (35, 81)

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 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.2.1.5 (Liste der eingeschlossenen Studien – indirekte Vergleiche) 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: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.

G-BA: Gemeinsamer Bundesausschuss; RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie

Um den Versorgungskontext möglichst genau ab zu bilden wurden auch Induktionsstudien eingeschlossen. Es wurden somit insgesamt zwei relevante RCTs mit dem zu bewertenden Arzneimittel Upadacitinib sowie eine relevante RCT mit der ZVT (Ustekinumab) für einen adjustierten indirekten Vergleich nach Bucher über den Brückenkomparator Placebo identifiziert.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche identifiziert wurden und bewerten Sie darüber hinaus deren Ähnlichkeit. Begründen Sie darauf basierend den Ein- bzw. Ausschluss von Studien für die von Ihnen durchgeführten indirekten Vergleiche. Bewerten Sie das Verzerrungspotenzial der für indirekte Vergleiche herangezogenen Studien. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

4.3.2.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Falls Teilpopulationen berücksichtigt werden, ist die Charakterisierung der Studienpopulation auch für diese Teilpopulation durchzuführen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Geben Sie bei den Datenschnitten auch den Anlass des Datenschnittes an. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.

Tabelle 4-30: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel und der ZVT für indirekte Vergleiche

Studie	**Studiendesign** <RCT, doppelblind/ einfach, verblindet/offen, parallel/cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer/ Datenschnitte** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
U-ACHIEVE (M14-234)	Randomisierte, doppelblinde, parallele, multizentrische, placebokontrollierte Phase-IIb/III-Studie bestehend aus drei Substudien	Erwachsene Patienten (≥ 16 – ≤ 75 Jahre) mit mittelschwerer bis schwerer aktiver CU, welche auf eine konventionelle oder biologische Therapie nicht oder nicht mehr adäquat ansprechen oder eine entsprechende Intoleranz aufweisen, jugendliche Patienten von ≥ 16 bis ≤ 18 Jahre müssen ein Körpergewicht von ≥ 40kg und ein Tanner-Stadium von 5 aufweisen und können in Substudie 2 eingeschlossen werden, sofern von lokaler Regulierungs- bzw. Gesundheits-behörde zugelassen, in Deutschland waren nur Patienten	<u>Substudie 1</u> (Induktionsphase): Upadacitinib 7,5 mg QD (n = 47), Upadacitinib 15 mg QD (n = 49), Upadacitinib 30 mg QD (n = 52 plus n = 65), Upadacitinib 45 mg QD (n = 56 plus n = 67), Placebo QD (n = 46) <u>Substudie 2</u> (Induktionsphase): Upadacitinib 45 mg QD (n = 319), Placebo QD (n = 155) <u>Substudie 3</u> (Erhaltungsphase): Upadacitinib 7,5 mg QD (n = 20), Upadacitinib 15 mg QD (n = 324), Upadacitinib 30 mg QD (n = 316), Placebo (n = 386)	Screening: 5 Wochen Behandlung: <u>Substudie 1:</u> 8 Wochen <u>Substudie 2:</u> 8 Wochen plus weitere 8 Wochen für Patienten ohne Ansprechen nach den ersten 8 Wochen <u>Substudie 3:</u> 52 Wochen Follow-up: 30 Tage Studiendauer: bis zu 77 Wochen unter Berücksichtigung von Screening, Induktionsphase (inklusive möglicher Verlängerung um weitere 8 Wochen), Erhaltungsphase und Follow-up Finaler Datenschnitt (Substudie 3):	International (Argentinien, Australien, Belgien, Bosnien und Herzegowina, Brasilien, Chile, China, Deutschland, Estland, Finnland, Frankreich, Griechenland, Irland, Israel, Italien, Japan, Kanada, Kolumbien, Korea, Kroatien, Lettland, Litauen, Malaysia, Mexiko, Niederlande, Norwegen, Österreich, Polen, Portugal, Russland, Schweiz, Serbien, Singapur, Slowakei, Spanien, Südafrika, Taiwan, Tschechische Republik, Türkei, Ukraine, Ungarn, Vereinigtes Königreich, Vereinigte Staaten	<u>Primärer Endpunkt:</u> • Anteil der Patienten mit klinischer Remission nach Adaptiertem Mayo Score zu Woche 8 (Substudie 1 und Substudie 2) bzw. Woche 52 (Substudie 3) <u>Patientenrelevante sekundäre Endpunkte</u> in Substudie 1, 2 und 3: • Endoskopische Verbesserung und endoskopisches Ansprechen • Klinisches Ansprechen nach Mayo Gesamt-score und nach partiellem Mayo Score

Studie	**Studiendesign** <RCT, doppelblind/ einfach, verblindet/offen, parallel/cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer/ Datenschnitte** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
		≥ 18 Jahre zugelassen.		30. April 2021 (Wirksamkeits-analyse) Letzter Patient letzte Visite: 13. Dezember 2021 (finale Analyse zur Verträglichkeit)	[einschließlich Puerto Rico]) Zeitraum: Substudie 1: 09/2016 – 06/2018 Substudie 2: 10/2018 – 09/2020 Substudie 3: 12/2016 – 12/2021	• Stuhlfrequenz • Histologische Verbesserung • Mukosaheilung • FACIT-F • IBDQ • Gesundheits-bezogene Lebensqualität • Verträglichkeit
U-ACCOMPLISH (M14-675)	Randomisierte, doppelblinde, parallele, multizentrische, placebokontrollierte Phase-III-Induktionsstudie	Erwachsene Patienten (≥ 16 – ≤ 75 Jahre) mit mittelschwerer bis schwerer aktiver CU, welche auf eine konventionelle oder biologische Therapie nicht oder nicht mehr adäquat ansprechen oder eine entsprechende Intoleranz aufweisen, jugendliche Patienten von ≥ 16 bis ≤ 18 Jahre müssen ein Körpergewicht von ≥ 40kg und ein Tanner-Stadium von	Upadacitinib 45 mg QD (n = 345), Placebo QD (n = 177)	Screening: 5 Wochen Behandlung: 8 Wochen plus weitere 8 Wochen für Patienten ohne Ansprechen nach den ersten 8 Wochen Follow-up: 30 Tage Studiendauer: bis zu 25 Wochen unter Berücksichtigung von Screening und Induktionsphase (inklusive möglicher Verlängerung um	International (Argentinien, Australien, Belgien, Bosnien und Herzegowina, Brasilien, Chile, China, Deutschland, Estland, Frankreich, Griechenland, Israel, Italien, Japan, Kanada, Kolumbien, Korea, Kroatien, Lettland, Litauen, Malaysia, Mexiko, Norwegen, Österreich, Polen, Portugal, Russland, Schweiz, Serbien, Singapur, Slowakei,	Primärer Endpunkt: • Anteil der Patienten mit klinischer Remission nach Adaptierter Mayo Score zu Woche 8 Patientenrelevante sekundäre Endpunkte: • Endoskopische Verbesserung und endoskopisches Ansprechen • Klinisches Ansprechen nach

Studie	**Studiendesign** <RCT, doppelblind/ einfach, verblindet/offen, parallel/cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer/ Datenschnitte** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
		5 aufweisen und können eingeschlossen werden, sofern von lokaler Regulierungs- bzw. Gesundheitsbehörde zugelassen, in Deutschland waren nur Patienten ≥ 18 Jahre zugelassen.		weitere 8 Wochen) und Follow-up	Spanien, Südafrika, Taiwan, Tschechische Republik, Türkei, Ukraine, Ungarn, Vereinigtes Königreich, Vereinigte Staaten [einschließlich Puerto Rico]) Zeitraum: 12/2018 – 01/2021	Mayo Gesamt-score und nach partiellem Mayo Score • Stuhlfrequenz • Histologische Verbesserung • Mukosaheilung • FACIT-F • IBDQ • Gesundheits-bezogene Lebensqualität • Verträglichkeit
Studie für den indirekten Vergleich gegenüber der ZVT Ustekinumab						
UNIFI	Randomisierte, doppelblinde, parallele, multizentrische, placebo-kontrollierte Phase-III-Studie bestehend aus einer Induktionsstudie und einer Erhaltungsstudie	Erwachsene Patienten (≥ 18 Jahre) mit mittelschwerer bis schwerer aktiver CU mit einem Mayo Gesamtscore von ≥ 6 und einem Endoskopie-Subscore von ≥ 2, welche auf eine konventionelle oder biologische Therapie nicht angesprochen haben	<u>Induktionsstudie:</u> Ustekinumab 6 mg/kg (n = 322), Ustekinumab 130 mg (n = 320), Placebo (n = 319) <u>Erhaltungsstudie:</u> Ustekinumab 90 mg/alle 12 Wochen (n = 172), Ustekinumab 90 mg/alle 8 Wochen	Screening: 8 Wochen Behandlung: <u>Induktionsstudie:</u> 8 Wochen plus weitere 8 Wochen für Patienten ohne Ansprechen nach den ersten 8 Wochen <u>Erhaltungsstudie:</u> 44 Wochen Verlängerungsstudie: weitere 3 Jahre	International (Australien, Belgien, Bulgarien, Dänemark, Deutschland, Frankreich, Israel, Italien, Japan, Kanada, Korea, Niederlande, Neuseeland, Österreich, Polen, Rumänien, Russland, Serbien, Slowakei, Tschechische	<u>Primärer Endpunkt:</u> • Anteil der Patienten mit klinischer Remission nach Mayo Gesamt-Score zu Woche 8 (Induktion) bzw. Woche 44 (Erhaltung) <u>Patientenrelevante sekundäre Endpunkte</u>

Studie	**Studiendesign** <RCT, doppelblind/ einfach, verblindet/offen, parallel/cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer/ Datenschnitte** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
		oder eine entsprechende Intoleranz aufweisen	(n = 176), Placebo (n = 175)	Studiendauer: bis zu 244 Wochen unter Berücksichtigung von Screening, Induktionsphase (inklusive möglicher Verlängerung um weitere 8 Wochen), Erhaltungsphase und Verlängerungsstudie	Republik, Ukraine, Ungarn, Vereinigtes Königreich, Vereinigte Staaten)	Induktionsstudie: • Endoskopische Verbesserung • Mukosaheilung • Klinisches Ansprechen Erhaltungsstudie: • Erhaltung des klinischen Ansprechens • Endoskopische Verbesserung • Kortikosteroid-freie klinische Remission Weitere patientenrelevante Endpunkte: • Veränderung des partiellen Mayo Score • IBDQ • Verträglichkeit

CU: Colitis ulcerosa; FACIT-F: Functional Assessment of Chronic Illness Therapy – Fatigue; IBDQ: Inflammatory Bowel Disease Questionnaire; QD: Einmal täglich; RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie

Tabelle 4-31: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel und der ZVT für indirekte Vergleiche

Studie	**Interventionsarm**	**Vergleichsarm**	***ggf. weitere Spalten mit Behandlungscharakteristika*** ***z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
U-ACHIEVE (M14-234)	<u>Induktionsphase</u> Substudie 1: Upadacitinib 7,5 mg, 15 mg, 30 mg oder 45 mg jeweils p.o. QD Substudie 2: 45 mg p.o. QD <u>Erhaltungsphase</u> Substudie 3: Upadacitinib 7,5 mg, 15 mg oder 30 mg jeweils p.o. QD	<u>Induktionsphase</u> Substudie 1 und Substudie 2: Placebo zu Upadacitinib p.o. QD <u>Erhaltungsphase</u> Substudie 3: Placebo zu Upadacitinib p.o. QD	• Vorbehandlung: Patienten, die mit einer konventionellen oder biologischen Therapie behandelt wurden und auf diese nicht oder nicht mehr adäquat ansprechen oder eine entsprechende Intoleranz aufweisen • Jegliche Medikationen oder Vakzinierungen, die innerhalb von 30 Tagen vor Baseline eingenommen wurden oder werden, wurden erfasst. • Jegliche Behandlungen mit Antibiotika, Kortikosteroiden, Immunsuppressiva, Aminosalicylaten oder biologischen Wirkstoffen innerhalb der letzten 90 Tage wurden erfasst. • Medikationen jeglicher vorheriger CU-bezogener Behandlungen mit Kortikosteroiden, Immunsuppressiva, Aminosalicylaten, Antibiotika oder biologischen Wirkstoffen wurden erfasst. • Wash-out-Phase: Bei Medikationen, die entsprechend der Ausschlusskriterien während der Behandlung in der Studie nicht eingenommen werden, mussten vor Studienbeginn die Wash-out-Zeiträume entsprechend der Protokollvorgaben eingehalten werden. • Begleitmedikation: CU-bezogene Begleitmedikationen mit oralen Kortikosteroiden, Antibiotika, Aminosalicylaten oder MTX konnten in Substudie 1 und Substudie 2 beibehalten werden, in Substudie 3 war ein Ausschleichen von oralen Kortikosteroiden gemäß einem Zeitplan vorgesehen. CU-bezogene Antibiotika-Therapien sollten nach Ermessen des Prüfarztes zu Studienbeginn von Substudie 3 abgesetzt werden.

Studie	**Interventionsarm**	**Vergleichsarm**	***ggf. weitere Spalten mit Behandlungscharakteristika*** ***z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
			• Rescue-Therapie: In oder nach Woche 4 können Patienten in Substudie 3, die nicht mehr ansprechen und eine medizinische Behandlung benötigen, jedoch nicht die Einschlusskriterien der Studie M14-533 erfüllen, eine Rescue-Therapie erhalten. Diese wird in Form der Einleitung oder Erhöhung der Dosierung zugelassener CU-bezogener Medikamente nach Ermessen des Prüfarztes gegeben, um neue oder sich verschlimmernde CU-Symptome zu behandeln. Die zulässigen CU-bezogenen Medikamente sind lokal wirkende, orale oder intravenöse Kortikosteroide, Aminosalicylate, MTX oder CU-bezogene Antibiotika. • Behandlungen mit biologischen Wirkstoffen, anderen JAK-Inhibitoren, Vakzinierungen und weiteren Medikationen gemäß Studienprotokoll während der Studie waren nicht erlaubt.
U-ACCOMPLISH (M14-675)	Upadacitinib 45 mg p.o. QD	Placebo zu Upadacitinib p.o. QD	• Vorbehandlung: Patienten, die mit einer konventionellen oder biologischen Therapie behandelt wurden und auf diese nicht oder nicht mehr adäquat ansprechen oder eine entsprechende Intoleranz aufweisen. • Sämtliche Medikationen oder Vakzinierungen, die innerhalb von 30 Tagen vor Baseline eingenommen wurden oder werden, wurden erfasst. • Behandlungen mit Antibiotika, Kortikosteroiden, Immunsuppressiva, Aminosalicylaten oder biologischen Wirkstoffen innerhalb der letzten 90 Tage wurden erfasst. • Medikationen jeglicher vorheriger CU-bezogener Behandlungen mit oralen Kortikosteroiden, Immunsuppressiva, Aminosalicylaten, Antibiotika oder biologischen Wirkstoffen wurden erfasst.

Studie	**Interventionsarm**	**Vergleichsarm**	***ggf. weitere Spalten mit Behandlungscharakteristika*** ***z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
			• Wash-out-Phase: Bei Medikationen, die entsprechend den Ausschlusskriterien während der Behandlung in der Studie nicht eingenommen werden, mussten vor Studienbeginn die Wash-out-Zeiträume entsprechend den Protokollvorgaben eingehalten werden. • Begleitmedikation: CU-bezogene Begleitmedikationen mit oralen Kortikosteroiden, Antibiotika, Aminosalicylaten oder MTX konnten beibehalten werden. • Behandlungen mit biologischen Wirkstoffen, anderen JAK-Inhibitoren, Vakzinierungen und weiteren Medikationen gemäß Studienprotokoll während der Studie waren nicht erlaubt.
Studie für den indirekten Vergleich gegenüber der ZVT Ustekinumab			
UNIFI	<u>Induktionsstudie:</u> Ustekinumab 130 mg i.v. oder Ustekinumab ~ 6 mg/kg i.v. (Dosierung: < 55 kg: 260 mg; 55 – 90 kg: 390 mg > 90 kg: 520 mg) <u>Erhaltungsstudie:</u> Ustekinumab 90 mg s.c. alle 12 Wochen oder Ustekinumab 90 mg s.c. alle 8 Wochen	<u>Induktionsstudie:</u> Placebo zu Ustekinumab i.v. <u>Erhaltungsstudie:</u> Placebo zu Ustekinumab s.c. alle 8 bzw. 12 Wochen	Zu Wahrung der Verblindung erhalten alle Patienten zu den Studienvisiten zu Woche 8, 12, 24, 32, 36 und 40 in der Erhaltungsstudie die Studienmedikation entsprechend der Randomisierung. • Vorbehandlung: Patienten, die mit einer konventionellen oder biologischen Therapie behandelt wurden und auf diese nicht oder nicht mehr adäquat ansprechen oder eine entsprechende Intoleranz aufweisen • Sämtliche Medikationen innerhalb von 30 Tagen vor der ersten Gabe der Studienmedikation wurden erfasst. • Sämtliche begleitende Medikationen und Therapien wurden erfasst. • CU-bezogene Medikationen mit oralen Kortikosteroiden, Aminosalicylaten, 6-MP, AZA oder MTX dürfen in der Induktionsphase nicht begonnen oder die Dosierung erhöht werden. Patienten, die bereits eine der aufgeführten Therapie erhielten, wurden mit der gleichbleibenden Dosierung weiterbehandelt.

Studie	**Interventionsarm**	**Vergleichsarm**	***ggf. weitere Spalten mit Behandlungscharakteristika*** ***z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
			• CU-bezogene Medikationen mit Aminosalicylaten, 6-MP, AZA oder MTX sollten in der Erhaltungsphase nicht begonnen oder die Dosierung erhöht werden. Patienten, die bereits eine der aufgeführten Therapie erhielten, wurden mit der gleichbleibenden Dosierung weiterbehandelt. • Eine Behandlung mit oralen Kortikosteroiden sollte in der Erhaltungsphase nicht begonnen werden und laufende Behandlungen sollen zu Baseline ausgeschlichen werden. Die Erhöhung der Dosierung oraler Kortikosteroide wird als Rescue-Therapie gewertet. • Rescue-Therapie: Jegliche Einleitung oder Dosiserhöhung einer medikamentösen Therapie, die bei einem Verlust des Ansprechens erforderlich ist, wird als Rescue-Medikation gewertet.
6-MP: 6-Mercaptopurin; AZA: Azathioprin; CU: Colitis ulcerosa; i.v.: Intravenös; JAK: Januskinase; MTX: Methotrexat; p.o.: Peroral; QD: Einmal täglich; RCT: Randomisierte kontrollierte Studie; s.c.: Subkutan; ZVT: Zweckmäßige Vergleichstherapie			

Im vorliegenden Anwendungsgebiet umfasst die Zulassung von Upadacitinib drei Dosierungen: 45 mg in der Induktionsphase sowie 30 mg und 15 mg in der Erhaltungsphase. Im Rahmen der Zulassungsstudien U-ACHIEVE (M14-234) und U-ACCOMPLISH (M14-675) wurden weitere Dosierungen von Upadacitinib (siehe Tabelle 4-30 und Tabelle 4-31) untersucht, die im Folgenden nicht weiter berücksichtigt werden, da sie nicht der Zulassung entsprechen.

Die Zulassung von Ustekinumab im vorliegenden Anwendungsgebiet umfasst in der Induktionsphase die Dosierung von ~ 6 mg Ustekinumab i.v. pro kg Körpergewicht mit den folgenden Dosierungen:

Tabelle 4-32: Dosierung von Ustekinumab nach Körpergewicht

Körpergewicht des Patienten zum Zeitpunkt der Dosierung	**Empfohlene Dosis**
≤ 55 kg	260 mg Ustekinumab i.v.
> 55 – ≤ 85 kg	390 mg Ustekinumab i.v.
> 85 kg	520 mg Ustekinumab i.v.
i.v.: Intravenös	

Im Rahmen der Studie UNIFI wurde außerdem die Dosierung von 130 mg Ustekinumab i.v. in der Induktionsstudie untersucht (siehe Tabelle 4-30 und Tabelle 4-31). Diese Dosierung wird im Folgenden nicht weiter berücksichtigt, da sie nicht der Zulassung entspricht.

Induktionsstudien

Upadacitinib

U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH

Tabelle 4-33: Charakterisierung der Studienpopulationen – Demografische Charakteristika zu Baseline nach Behandlungsarm in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	Placebo N = 46	Upadacitinib 45 mg N = 319	Placebo N = 154	Upadacitinib 45 mg N = 341	Placebo N = 174
Alter, Jahre						
MW (SD)	39,8 (14,20)	42,3 (13,29)	43,6 (14,04)	44,4 (14,63)	42,1 (14,74)	42,2 (14,35)
Geschlecht, n (%)						
weiblich	19 (33,9)	17 (37,0)	121 (37,9)	57 (37,0)	127 (37,2)	67 (38,5)
männlich	37 (66,1)	29 (63,0)	198 (62,1)	97 (63,0)	214 (62,8)	107 (61,5)
Abstammung, n (%)						
weiß	38 (67,9)	37 (80,4)	206 (64,6)	100 (64,9)	234 (68,6)	124 (71,3)
schwarz oder afroamerikanisch	1 (1,8)	0	12 (3,8)	4 (2,6)	11 (3,2)	6 (3,4)
asiatisch	16 (28,6)	8 (17,4)	95 (29,8)	46 (29,9)	94 (27,6)	41 (23,6)
andere[a]	1 (1,8)	1 (2,2)	6 (1,9)	4 (2,6)	2 (0,6)	3 (1,7)
BMI, kg/m²						
n (fehlend)	54 (2)	46 (0)	318 (1)	153 (1)	341 (0)	174 (0)
MW (SD)	24,2 (5,85)	25,0 (5,54)	24,7 (5,10)	25,7 (6,68)	25,1 (5,33)	25,4 (5,94)

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	Placebo N = 46	Upadacitinib 45 mg N = 319	Placebo N = 154	Upadacitinib 45 mg N = 341	Placebo N = 174
Gewicht, kg						
MW (SD)	70,8 (18,29)	74,3 (17,63)	71,3 (17,39)	74,0 (19,20)	74,0 (18,59)	73,7 (20,12)
Geografische Region, n (%)						
USA	16 (28,6)	14 (30,4)	35 (11,0)	21 (13,6)	73 (21,4)	43 (24,7)
andere	40 (71,4)	32 (69,6)	284 (89,0)	133 (86,4)	268 (78,6)	131 (75,3)

a: Indianisch/Ureinwohner Alaskas, hawaiianisch/pazifisch, andere

Population: ITT

BMI: Body Mass Index; ITT: Intention-to-Treat; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

Tabelle 4-34: Charakterisierung der Studienpopulationen – Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	Placebo N = 46	Upadacitinib 45 mg N = 319	Placebo N = 154	Upadacitinib 45 mg N = 341	Placebo N = 174
Dauer der Erkrankung, Jahre						
MW (SD)	7,95 (6,794)	7,55 (6,659)	8,621 (7,1719)	9,116 (8,7610)	7,278 (6,4482)	7,387 (7,2040)

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	**Placebo N = 46**	**Upadacitinib 45 mg N = 319**	**Placebo N = 154**	**Upadacitinib 45 mg N = 341**	**Placebo N = 174**
Ausdehnung der Erkrankung, n (%)						
Proktitis	0	0	0	0	1 (0,3)	0
Linksseitencolitis	26 (46,4)	18 (39,1)	158 (49,5)	74 (48,1)	164 (48,1)	88 (50,6)
Ausgedehnte Colitis/Pankolitis	30 (53,6)	28 (60,9)	161 (50,5)	80 (51,9)	176 (51,6)	86 (49,4)
Mayo Gesamtscore zu Baseline						
n (fehlend)	55 (1)	46 (0)	318 (1)	154 (0)	340 (1)	174 (0)
MW (SD)	9,0 (1,38)	9,4 (1,37)	9,30 (1,449)	9,31 (1,470)	9,30 (1,453)	9,37 (1,416)
Mayo Gesamtscore zu Baseline, n (%)						
≤ 9	n.a.	n.a.	162 (50,9)	79 (51,3)	160 (47,1)	86 (49,4)
> 9	n.a.	n.a.	156 (49,1)	75 (48,7)	180 (52,9)	88 (50,6)
fehlend	n.a.	n.a.	1	0	1	0
Adaptierter Mayo Score zu Baseline						
n (fehlend)	55 (1)	46 (0)	318	154 (0)	340	174
MW (SD)	6,7 (1,24)	7,0 (1,11)	7,01 (1,236)	6,96 (1,171)	6,99 (1,214)	7,03 (1,235)

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	Placebo N = 46	Upadacitinib 45 mg N = 319	Placebo N = 154	Upadacitinib 45 mg N = 341	Placebo N = 174
Adaptierter Mayo Score zu Baseline, n (%)						
≤ 7	38 (69,1)	27 (58,7)	195 (61,3)	94 (61,0)	205 (60,3)	103 (59,2)
> 7	17 (30,9)	19 (41,3)	123 (38,7)	60 (39,0)	135 (39,7)	71 (40,8)
fehlend	1	0	1	0	1	0
hsCRP, mg/l						
MW (SD)	12,683 (15,0436)	9,559 (10,6251)	9,412 (15,2639)	12,223 (21,2272)	9,283 (15,3126)	10,782 (19,9449)
Median (IQR)	6,300 (k. A.)	5,390 (k. A.)	4,100 (k. A.)	4,675 (k. A.)	3,820 (k. A.)	4,730 (k. A.)
Fäkales Calprotectin, mg/kg						
n (fehlend)	45 (11)	38 (8)	277	137	296	152
MW (SD)	2.953,9 (3.503,08)	3.298,9 (4.900,68)	3.910,1 (5.698,72)	3.135,5 (3.986,79)	3.130,5 (4.719,82)	3.126,7 (4.742,09)
Median (IQR)	1.666,0 (k. A.)	2.100,5 (k. A.)	1.780,0 (k. A.)	1.902,0 (k. A.)	1.654,5 (k. A.)	1.552,5 (k. A.)

Population: ITT

hsCRP: Hoch sensitives C-reaktives Protein; IQR: Interquartile Range; ITT: Intention-to-Treat; k. A.: Keine Angabe; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

Tabelle 4-35: Charakterisierung der Studienpopulationen – Medizinische Vorgeschichte zu Baseline nach Behandlungsarm in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	Placebo N = 46	Upadacitinib 45 mg N = 319	Placebo N = 154	Upadacitinib 45 mg N = 341	Placebo N = 174
Vorherige biologische Therapie						
Bio-IR	42 (75,0)	34 (73,9)	168 (52,7)	78 (50,6)	172 (50,4)	89 (51,1)
Non-Bio-IR	14 (25,0)	12 (26,1)	151 (47,3)	76 (49,4)	169 (49,6)	85 (48,9)
- Bio-naiv	k. A.	k. A.	145 (45,5)	72 (46,8)	168 (49,3)	81 (46,5)
Kortikosteroide zu Baseline, n (%)						
ja	29 (51,8)	26 (56,5)	124 (38,9)	61 (39,6)	120 (35,2)	72 (41,4)
nein	27 (48,2)	20 (43,5)	195 (61,1)	93 (60,4)	221 (64,8)	102 (58,6)
Aminosalicylate zu Baseline, n (%)						
ja	25 (44,6)	26 (56,5)	220 (69,0)	103 (66,9)	233 (68,3)	120 (69,0)
nein	31 (55,4)	20 (43,5)	99 (31,0)	51 (33,1)	108 (31,7)	54 (31,0)
Immunsuppressiva zu Baseline, n (%)						
ja	3 (5,4)	1 (2,2)	2 (0,6)	3 (1,9)	1 (0,3)	3 (1,7)
nein	53 (94,6)	45 (97,8)	317 (99,4)	151 (98,1)	340 (99,7)	171 (98,3)

Population: ITT

Bio-IR: Intoleranz oder unzureichendes Ansprechen auf eine vorherige biologische Therapie; ITT: Intention-to-Treat; k. A.: Keine Angabe; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie

Tabelle 4-36: Patientenstatus in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	Placebo N = 46	Upadacitinib 45 mg N = 319	Placebo N = 154	Upadacitinib 45 mg N = 341	Placebo N = 174
Patienten mit Abbruch der Studienmedikation, n (%)	6 (10,7)	5 (10,9)	12 (3,8)	19 (12,3)	11 (3,2)	13 (7,5)
Primärer Abbruchgrund						
UE	4 (7,1)	3 (6,5)	7 (2,2)	7 (4,5)	5 (1,5)	5 (2,9)
Abbruch auf Wunsch des Patienten	0	0	1 (0,3)	2 (1,3)	6 (1,8)	4 (2,3)
Lost-to-follow-up	0	0	1 (0,3)	0	0	0
fehlende Wirksamkeit	2 (3,6)	2 (4,3)	2 (0,6)	9 (5,8)	0	4 (2,3)
andere	0	0	1 (0,3)	1 (0,6)	0	0
Patienten mit Abbruch der Studie, n (%)	k. A.	k. A.	13 (4,1)	20 (13,0)	11 (3,2)	13 (7,5)
Primärer Abbruchgrund						
UE	k. A.	k. A.	6 (1,9)	9 (5,8)	5 (1,5)	6 (3,4)
Abbruch auf Wunsch des Patienten	k. A.	k. A.	2 (0,6)	3 (1,9)	6 (1,8)	4 (2,3)
Lost-to-follow-up	k. A.	k. A.	1 (0,3)	1 (0,6)	0	0
COVID-19-Infektion	k. A.	k. A.	0	0	0	0

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	Placebo N = 46	Upadacitinib 45 mg N = 319	Placebo N = 154	Upadacitinib 45 mg N = 341	Placebo N = 174
Logistische Restriktionen durch COVID-19	k. A.	k. A.	1 (0,3)	0	0	0
andere	k. A.	k. A.	3 (0,9)	7 (4,5)	0	3 (1,7)
Population: ITT ITT: Intention-to-Treat; k. A.: Keine Angabe; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; UE: Unerwünschtes Ereignis						

Tabelle 4-37: Behandlungsdauer (Tage) in RCT (Induktionsstudien: U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH) mit dem zu bewertenden Arzneimittel

Merkmal	U-ACHIEVE Substudie 1		U-ACHIEVE Substudie 2		U-ACCOMPLISH	
	Upadacitinib 45 mg N = 56	Placebo N = 46	Upadacitinib 45 mg N = 319	Placebo N = 155	Upadacitinib 45 mg N = 344	Placebo N = 177
Behandlungsdauer (Tage)						
MW (SD)	54,5 (8,59)	54,1 (9,83)	56,1 (7,52)	53,1 (11,96)	55,8 (7,01)	54,3 (11,10)
Median	56,0	56,0	56,0	56,0	56,0	56,0
Min – Max	15 – 66	17 – 63	5 – 92	1 – 74	4 – 78	6 – 68
Population: ITT (U-ACHIEVE Substudie 1), Safety Population (U-ACHIEVE Substudie 2, U-ACCOMPLISH) ITT: Intention-to-Treat; MW: Mittelwert; N: Anzahl Patienten; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung						

Ustekinumab

UNIFI Induktionsstudie

Tabelle 4-38: Charakterisierung der Studienpopulationen – Demografische Charakteristika zu Baseline nach Behandlungsarm in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Induktionsstudie	
	Ustekinumab 6 mg/kg N = 322	**Placebo N = 319**
Alter, Jahre		
MW (SD)	41,7 (13,7)	41,2 (13,5)
Geschlecht, n (%)		
weiblich	127 (39,4)	122 (38,2)
männlich	195 (60,6)	197 (61,8)
Abstammung, n (%)		
weiß	243 (75,5)	248 (77,7)
schwarz oder afroamerikanisch	0	3 (0,9)
asiatisch	49 (15,2)	48 (15,0)
andere	13 (4,0)	8 (2,5)
unbekannt/nicht erhoben	17 (5,3)	12 (3,8)
Gewicht, kg		
MW (SD)	73,0 (19,3)	72,9 (16,8)

Population: ITT

ITT: Intention-to-Treat; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ZVT: Zweckmäßige Vergleichstherapie

Tabelle 4-39: Charakterisierung der Studienpopulationen – Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Induktionsstudie	
	Ustekinumab 6 mg/kg N = 322	**Placebo N = 319**
Dauer der Erkrankung, Jahre		
MW (SD)	8,2 (7,8)	8,0 (7,2)

Merkmal	UNIFI Induktionsstudie	
	Ustekinumab 6 mg/kg N = 322	**Placebo N = 319**
Ausdehnung der Erkrankung, n (%)		
n (fehlend)	320 (2)	316 (3)
Proktitis/Linksseitencolitis	168 (52,5)	167 (52,8)
Ausgedehnte Colitis/Pankolitis	152 (47,5)	149 (47,2)
Mayo Gesamtscore zu Baseline		
n (fehlend)	321 (1)	319 (0)
MW (SD)	8,9 (1,5)	8,9 (1,6)
Mayo Gesamtscore zu Baseline, n (%)		
≥ 6 – ≤ 10	276 (86,0)	263 (83,0)
> 10	45 (14,0)	54 (17,0)
fehlend	1	2
hsCRP, mg/l		
n (fehlend)	320 (2)	316 (3)
MW (SD)	k. A.	k. A.
Median (IQR)	4,8 (1,8 – 13,7)	4,7 (1,4 – 10,0)
Fäkales Calprotectin, mg/kg		
n (fehlend)	300 (22)	289 (30)
MW (SD)	k. A.	k. A.
Median (IQR)	1.506,5 (621,5 – 3.192,5)	1.224,0 (496,0 – 2.224,0)

Population: ITT

hsCRP: Hoch sensitives C-reaktives Protein; IQR: Interquartile Range; ITT: Intention-to-Treat; k. A.: Keine Angabe; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ZVT: Zweckmäßige Vergleichstherapie

Tabelle 4-40: Charakterisierung der Studienpopulationen – Medizinische Vorgeschichte zu Baseline nach Behandlungsarm in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Induktionsstudie	
	Ustekinumab 6 mg/kg N = 322	**Placebo N = 319**
Vorherige biologische Therapie		
Bio-IR	166 (51,6)	161 (50,5)

Merkmal	UNIFI Induktionsstudie	
	Ustekinumab 6 mg/kg N = 322	**Placebo N = 319**
Non-Bio-IR	156 (48,4)	158 (49,5)
- Bio-naiv	147 (45,7)	151 (47,3)
Kortikosteroide zu Baseline, n (%)		
ja	168 (52,2)	157 (49,2)
nein	k. A.	k. A.
Aminosalicylate zu Baseline, n (%)		
ja	238 (73,9)	207 (64,9)
nein	k. A.	k. A.
Immunsuppressiva[a] zu Baseline, n (%)		
ja	89 (27,6)	89 (27,9)
nein	k. A.	k. A.

a: Azathioprin, Mercaptopurin, Methotrexat

Population: ITT

Bio-IR: Intoleranz oder unzureichendes Ansprechen auf eine vorherige biologische Therapie; ITT: Intention-to-Treat; k. A.: Keine Angabe; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie

Tabelle 4-41: Patientenstatus in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Induktionsstudie	
	Ustekinumab 6 mg/kg N = 322	**Placebo N = 319**
Patienten mit Abbruch der Studie oder der Studienmedikation, n (%)	15 (4,7)	23 (7,2)
UE	1 (0,3)	3 (0,9)
Tod	1 (0,3)	0
Abbruch auf Wunsch des Patienten	7 (2,2)	17 (5,3)
Lost-to-follow-up	1 (0,3)	0
Entscheidung des Prüfarzt	0	1 (0,3)
andere	5 (1,6)	2 (0,6)

Merkmal	UNIFI Induktionsstudie	
	Ustekinumab 6 mg/kg N = 322	Placebo N = 319
Population: ITT ITT: Intention-to-Treat; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; UE: Unerwünschtes Ereignis; ZVT: Zweckmäßige Vergleichstherapie		

Tabelle 4-42: Behandlungs- und Beobachtungsdauer in RCT UNIFI (Induktionsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Induktionsstudie	
	Ustekinumab 6 mg/kg N = 320	Placebo N = 319
Behandlungen (Anzahl der Gaben)		
MW (SD)	1 (k. A.)	1 (k. A.)
Beobachtungsdauer FU (Wochen)		
MW (SD)	8,6 (k. A.)	8,7 (k. A.)
Population: ITT FU: Follow-up; ITT: Intention-to-Treat; k. A.: Keine Angabe; MW: Mittelwert; N: Anzahl Patienten; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ZVT: Zweckmäßige Vergleichstherapie		

Erhaltungsstudien

Upadacitinib

U-ACHIEVE Substudie 3

Tabelle 4-43: Charakterisierung der Studienpopulationen – Demografische Charakteristika zu Baseline nach Behandlungsarm in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg N = 148	Upadacitinib 30 mg N = 154	Placebo N = 149
Alter, Jahre			
MW (SD)	42,6 (14,10)	42,6 (14,75)	43,3 (14,37)
Geschlecht, n (%)			
weiblich	53 (35,8)	68 (44,2)	64 (43,0)
männlich	95 (64,2)	86 (55,8)	85 (57,0)

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg N = 148	Upadacitinib 30 mg N = 154	Placebo N = 149
Abstammung, n (%)			
weiß	97 (65,5)	101 (65,6)	93 (62,4)
schwarz oder afroamerikanisch	7 (4,7)	3 (1,9)	6 (4,0)
asiatisch	44 (29,7)	48 (31,2)	42 (28,2)
andere[a]	0	2 (1,3)	8 (5,4)
BMI, kg/m²			
n (fehlend)	148 (0)	153 (1)	149 (0)
MW (SD)	25,5 (5,93)	25,3 (6,52)	25,0 (5,35)
Gewicht, kg			
MW (SD)	74,4 (20,60)	73,1 (20,77)	72,2 (18,16)
Geografische Region, n (%)			
USA	26 (17,6)	30 (19,5)	32 (21,5)
andere	122 (82,4)	124 (80,5)	117 (78,5)

a: Indianisch/Ureinwohner Alaskas, hawaiianisch/pazifisch, andere

Population: ITT

BMI: Body Mass Index; ITT: Intention-to-Treat; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

Tabelle 4-44: Charakterisierung der Studienpopulationen – Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg N = 148	Upadacitinib 30 mg N = 154	Placebo N = 149
Dauer der Erkrankung, Jahre			
MW (SD)	8,915 (8,0973)	8,214 (7,6173)	8,748 (7,9964)
Ausdehnung der Erkrankung, n (%)			
Proktitis	0	0	0
Linksseitencolitis	66 (44,6)	68 (44,2)	79 (53,0)
Ausgedehnte Colitis/Pankolitis	82 (55,4)	86 (55,8)	70 (47,0)
Mayo Gesamtscore zu Baseline			
n (fehlend)	148 (0)	152 (2)	149 (0)
MW (SD)	9,25 (1,391)	9,42 (1,498)	9,33 (1,404)

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg N = 148	Upadacitinib 30 mg N = 154	Placebo N = 149
Mayo Gesamtscore zu Baseline, n (%)			
≤ 9	75 (50,7)	73 (48,0)	74 (49,7)
> 9	73 (49,3)	79 (52,0)	75 (50,3)
fehlend	0	2	0
Adaptierter Mayo Score zu Baseline, n (%)			
n (fehlend)	148 (0)	152 (2)	149 (0)
MW (SD)	6,98 (1,201)	7,08 (1,301)	7,00 (1,231)
Adaptierter Mayo Score zu Baseline, n (%)			
≤ 7	89 (60,1)	88 (57,9)	87 (58,4)
> 7	59 (39,9)	64 (42,1)	62 (41,6)
fehlend	0	2	0
hsCRP, mg/l			
MW (SD)	8,362 (12,4433)	8,626 (14,7480)	9,827 (15,8896)
Median (IQR)	3,810 (k. A.)	4,080 (k. A.)	4,250 (k. A.)
Fäkales Calprotectin, mg/kg			
n (fehlend)	130 (18)	129 (25)	127 (22)
MW (SD)	3.141,7 (4.694,00)	2.737,3 (4.326,71)	3.620,3 (5.222,27)
Median (IQR)	1.718,0 (k. A.)	1.465,0 (k. A.)	1.991,0 (k. A.)

Population: ITT

hsCRP: Hoch sensitives C-reaktives Protein; IQR: Interquartile Range; ITT: Intention-to-Treat; k. A.: Keine Angabe; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

Tabelle 4-45: Charakterisierung der Studienpopulationen – Medizinische Vorgeschichte zu Baseline nach Behandlungsarm in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg N = 148	Upadacitinib 30 mg N = 154	Placebo N = 149
Vorherige biologische Therapie, n (%)			
Bio-IR	71 (48,0)	73 (47,4)	81 (54,4)

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg N = 148	**Upadacitinib 30 mg N = 154**	**Placebo N = 149**
Non-Bio-IR	77 (52,0)	81 (52,6)	68 (45,6)
- Bio-naiv	75 (50,7)	77 (50,0)	65 (43,6)
Anzahl vorheriger biologischer Therapien, n (%)			
≤ 1	26 (36,6)	26 (35,6)	24 (29,6)
> 1	45 (63,4)	47 (64,4)	57 (70,4)
Kortikosteroide zu Baseline, n (%)			
ja	55 (37,2)	57 (37,0)	60 (40,3)
nein	93 (62,8)	97 (63,0)	89 (59,7)
Aminosalicylate zu Baseline, n (%)			
ja	99 (66,9)	106 (68,8)	99 (66,4)
nein	49 (33,1)	48 (31,2)	50 (33,6)
Immunsuppressiva zu Baseline, n (%)			
ja	1 (0,7)	1 (0,6)	0
nein	147 (99,3)	153 (99,4)	149 (100)

Population: ITT

ITT: Intention-to-Treat; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie;
RCT: Randomisierte kontrollierte Studie

Tabelle 4-46: Patientenstatus in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg N = 148	**Upadacitinib 30 mg N = 154**	**Placebo N = 149**
Patienten mit Abbruch der Studienmedikation, n (%)	49 (33,1)	33 (21,4)	98 (65,8)
Primärer Abbruchgrund			
UE	4 (2,7)	8 (5,2)	14 (9,4)
Abbruch auf Wunsch des Patienten	1 (0,7)	4 (2,6)	1 (0,7)
Fehlende Wirksamkeit	35 (23,6)	12 (7,8)	74 (49,7)
Lost-to-follow-up	0	1 (0,6)	0
COVID-19-Infektion	0	1 (0,6)	0
Logistische Restriktionen durch COVID-19	0	1 (0,6)	0

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg N = 148	**Upadacitinib 30 mg N = 154**	**Placebo N = 149**
andere	9 (6,1)	6 (3,9)	9 (6,0)
Patienten mit Abbruch der Studie, n (%)	49 (33,1)	33 (21,4)	98 (65,8)
Primärer Abbruchgrund			
UE	5 (3,4)	10 (6,5)	13 (8,7)
Abbruch auf Wunsch des Patienten	1 (0,7)	5 (3,2)	4 (2,7)
Lost-to-follow-up	0	1 (0,6)	0
COVID-19-Infektion	0	0	0
Logistische Restriktionen durch COVID-19	0	1 (0,6)	0
andere	43 (29,1)	16 (10,4)	81 (54,4)

Population: ITT

ITT: Intention-to-Treat; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; UE: Unerwünschtes Ereignis

Tabelle 4-47: Behandlungsdauer (Tagen) in RCT U-ACHIEVE (Erhaltungsphase, Substudie 3) mit dem zu bewertenden Arzneimittel

Merkmal	Erhaltungsstudie U-ACHIEVE (Substudie 3, Kohorte 1)		
	Upadacitinib 15 mg	**Upadacitinib 30 mg**	**Placebo**
Safety Population A[a]	**N = 148**	**N = 154**	**N = 149**
Behandlungsdauer (Tage)			
MW (SD)	294,5 (115,04)	320,4 (100,25)	214,2 (132,04)
Median	363,0	364,0	218,0
Min – Max	29 – 399	8 – 448	2 – 425
Safety Population C[b]	**N = 250**	**N = 251**	**N = 245**
Behandlungsdauer (Tage)			
MW (SD)	266,5 (110,92)	289,2 (101,43)	191,0 (123,47)
Median	311,0	355,0	168,0
Min – Max	2 – 399	8 – 448	2 – 425

Population: Safety Population

a: Die Safety Population A umfasst die ersten 451 Patienten der Kohorte 1 der Erhaltungsstudie (U-ACHIEVE, Substudie 3) ab Studienprotokoll Amendment 3.

b: Die Safety Population C umfasst alle Patienten der Kohorte 1 der Erhaltungsstudie (U-ACHIEVE, Substudie 3).

MW: Mittelwert; N: Anzahl Patienten; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung

Ustekinumab

UNIFI Erhaltungsstudie

Tabelle 4-48: Charakterisierung der Studienpopulationen – Demografische Charakteristika zu Baseline nach Behandlungsarm in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Erhaltungsstudie		
	Ustekinumab 90 mg/alle 12 Wo. N = 172	**Ustekinumab 90 mg/alle 8 Wo. N = 176**	**Placebo N = 175**
Alter, Jahre			
MW (SD)	40,7 (13,47)	39,5 (13,32)	42,0 (13,85)
Geschlecht, n (%)			
weiblich	76 (44,2)	82 (46,6)	68 (38,9)
männlich	96 (55,8)	94 (53,4)	107 (61,1)
Gewicht, kg			
MW (SD)	73,3 (18,91)	72,0 (19,12)	71,7 (14,61)

Population: ITT

ITT: Intention-to-Treat; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ZVT: Zweckmäßige Vergleichstherapie

Tabelle 4-49: Charakterisierung der Studienpopulationen – Krankheitsspezifische Charakteristika zu Baseline nach Behandlungsarm in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Erhaltungsstudie		
	Ustekinumab 90 mg/alle 12 Wo. N = 172	**Ustekinumab 90 mg/alle 8 Wo. N = 176**	**Placebo N = 175**
Dauer der Erkrankung, Jahre			
MW (SD)	8,6 (8,31)	8,1 (6,57)	7,5 (6,80)
Ausdehnung der Erkrankung, n (%)			
n (fehlend)	172 (0)	174 (2)	175 (0)
Linksseitencolitis/Proktitis	92 (53,5)	95 (54,6)	89 (50,9)
Ausgedehnte Colitis/Pankolitis	80 (46,5)	79 (45,4)	86 (49,1)
Mayo Gesamtscore zu Baseline			
n (fehlend)	172 (0)	174 (2)	175 (0)
MW (SD)	8,9 (1,58)	8,9 (1,55)	8,7 (1,52)

Merkmal	UNIFI Erhaltungsstudie		
	Ustekinumab 90 mg/alle 12 Wo. N = 172	**Ustekinumab 90 mg/alle 8 Wo. N = 176**	**Placebo N = 175**
Mayo Gesamtscore zu Baseline, n (%)			
≥ 6 – ≤ 10	150 (87,2)	147 (84,5)	156 (89,1)
> 10	22 (12,8)	27 (15,5)	19 (10,9)
fehlend	0	2	0
hsCRP, mg/l			
n (fehlend)	170 (2)	174 (2)	174 (1)
MW (SD)	k. A.	k. A.	k. A.
Median (IQR)	3,3 (1,15 – 9,16)	4,0 (1,37 – 12,70)	3,4 (1,39 – 9,72)
Fäkales Calprotectin, mg/kg			
n (fehlend)	156 (16)	156 (20)	163 (12)
MW (SD)	k. A.	k. A.	k. A.
Median (IQR)	1.258,0 (598,5 – 2.594,5)	1.510,0 (566,5 – 2.868,5)	1.487,0 (604,0 – 2.625,0)

Population: ITT

hsCRP: Hoch sensitives C-reaktives Protein; IQR: Interquartile Range; ITT: Intention-to-Treat; k. A.: Keine Angabe; MW: Mittelwert; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ZVT: Zweckmäßige Vergleichstherapie

Tabelle 4-50: Charakterisierung der Studienpopulationen – Medizinische Vorgeschichte zu Baseline nach Behandlungsarm in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Erhaltungsstudie		
	Ustekinumab 90 mg/alle 12 Wo. N = 172	**Ustekinumab 90 mg/alle 8 Wo. N = 176**	**Placebo N = 175**
Vorherige biologische Therapie			
Bio-IR	70 (40,7)	91 (51,7)	88 (50,3)
Non-Bio-IR	102 (59,3)	85 (48,3)	87 (49,7)
- Bio-naiv	95 (55,2)	79 (44,9)	84 (48,0)
Kortikosteroide zu Baseline, n (%)			
ja	83 (48,3)	95 (54,0)	95 (54,3)
nein	k. A.	k. A.	k. A.

Merkmal	UNIFI Erhaltungsstudie		
	Ustekinumab 90 mg/alle 12 Wo. N = 172	**Ustekinumab 90 mg/alle 8 Wo. N = 176**	**Placebo N = 175**
Aminosalicylate zu Baseline, n (%)			
ja	134 (77,9)	112 (63,6)	124 (70,9)
nein	k. A.	k. A.	k. A.
Immunsuppressiva zu Baseline, n (%)			
ja	44 (25,6)	46 (26,1)	49 (28,0)
nein	k. A.	k. A.	k. A.

Population: ITT

ITT: Intention-to-Treat; k. A.: Keine Angabe; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie

Tabelle 4-51: Patientenstatus in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Erhaltungsstudie		
	Ustekinumab 90 mg/alle 12 Wo. N = 172	**Ustekinumab 90 mg/alle 8 Wo. N = 176**	**Placebo N = 175**
Patienten mit Abbruch der Studie, n (%)	11 (6,4)	8 (4,5)	10 (5,7)
Lost-to-follow-up	0	1 (0,6)	1 (0,6)
Tod	0	0	0
Rücknahme der Einwilligungserklärung	10 (5,8)	5 (2,8)	8 (4,6)
Entscheidung des Sponsors	0	1 (0,6)	1 (0,6)
andere	1 (0,6)	1 (0,6)	0

Population: ITT

ITT: Intention-to-Treat; N: Anzahl Patienten; n: Anzahl Patienten in der jeweiligen Kategorie; RCT: Randomisierte kontrollierte Studie; ZVT: Zweckmäßige Vergleichstherapie

Tabelle 4-52: Behandlungsdauer (Tagen) in RCT UNIFI (Erhaltungsstudie) mit der ZVT für den indirekten Vergleich

Merkmal	UNIFI Erhaltungsstudie		
	Ustekinumab 90 mg/alle 12 Wo. N = 172	**Ustekinumab 90 mg/alle 8 Wo. N = 176**	**Placebo N = 175**
Behandlungen (Anzahl der Gaben)			
MW (SD)	7,3 (k. A.)	7,4 (k. A.)	7,1 (k. A.)
Beobachtungsdauer FU (Wochen)			
MW (SD)	41,8 (k. A.)	42,2 (k. A.)	42,3 (k. A.)
Population: Safety Analysis Set FU: Follow-up; k. A.: Keine Angabe; MW: Mittelwert; N: Anzahl Patienten; RCT: Randomisierte kontrollierte Studie; SD: Standardabweichung; ZVT: Zweckmäßige Vergleichstherapie			

Beschreiben Sie die Studien zusammenfassend. In der Beschreibung der Studien sollten Informationen zur Behandlungsdauer sowie zu geplanter und tatsächlicher Beobachtungsdauer enthalten sein. Sofern sich die Beobachtungsdauer zwischen den relevanten Endpunkten unterscheidet, sind diese unterschiedlichen Beobachtungsdauern endpunktbezogen anzugeben. Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich. 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

Studiendesign der RCT für den indirekten Vergleich

Das Studiendesign entspricht dem Standard in der Indikation und sah sowohl bei Upadacitinib, als auch bei Ustekinumab eine Induktionsphase und eine anschließende Re-Randomisierung der Therapieansprecher in der Erhaltungsphase vor. Auch im Versorgungsalltag wird die Erhaltungsphase bei Patienten nach der Induktion nur fortgeführt, wenn diese ausreichend angesprochen haben. Die erneute Re-Randomisierung in der Erhaltungsphase ist zwischen den Studien ausreichend ähnlich. Die Studiendauer ist mit einem Erhebungszeitpunkt von 52 Wochen bei Upadacitinib und einem Erhebungszeitpunkt von 44 Wochen bei Ustekinumab sowohl in Bezug auf die Studiendauer, als auch die Vergleichbarkeit der zwei Zeitpunkte adäquat. Die EMA-Guideline empfiehlt eine Studiendauer in der vorliegenden Indikation von 52 Wochen, welche sowohl die Induktions- als auch die Erhaltungsphase einschließen sollten.

Upadacitinib

U-ACHIEVE und U-ACCOMPLISH

Bei der Studie U-ACHIEVE (M14-234) handelt es sich um eine randomisierte, doppelblinde, multizentrische, placebokontrollierte Phase-IIb/III-Studie, die drei Substudien umfasst. In der Studie wird die Sicherheit und Wirksamkeit von Upadacitinib bei Patienten mit mittelschwerer bis schwerer aktiver CU in der Induktions- und Erhaltungsphase untersucht. Die Substudie 1 und Substudie 2 umfassen die Induktionsphase. Die Patienten aus den Substudien 1 und 2 mit klinischem Ansprechen in der Induktionsphase wurden in die Substudie 3 eingeschlossen, in der sie im Rahmen der Erhaltungsphase mit Upadacitinib oder Placebo weiter behandelt wurden. In die Substudie 3 wurden außerdem Patienten aus einer weiteren Induktionsstudie, der Studie U-ACCOMPLISH, eingeschlossen. Die Studie U-ACCOMPLISH (M14-675) ist eine randomisierte, doppelblinde, multizentrische, placebokontrollierte Phase-III-Induktionsstudie, die Sicherheit und Wirksamkeit von Upadacitinib in der Induktionsphase bei Patienten mit mittelschwerer bis schwerer aktiver CU untersucht. Sie entspricht in Design und Methodik der Substudie 2 der U-ACHIEVE.

Ustekinumab

UNIFI

Bei der Studie UNIFI handelt es sich um eine randomisierte, doppelblinde, multizentrische, placebokontrollierte Phase-III-Studie, die die Sicherheit und Wirksamkeit von Ustekinumab in der Induktions- und Erhaltungsphase bei Patienten mit mittelschwerer bis schwerer aktiver CU untersucht, die ein unzureichendes Ansprechen auf eine konventionelle oder biologische Therapie gezeigt haben oder eine entsprechende Intoleranz aufweisen. Die Aufteilung in eine Induktions- und eine Erhaltungsstudie entspricht zwei Einzelstudien unter einem gemeinsamen Studienprotokoll. Nach Abschluss der Erhaltungsphase über 44 Wochen schließt eine Verlängerungsphase von drei Jahren an.

Induktionsstudien

Upadacitinib

U-ACHIEVE Substudie 1

In der U-ACHIEVE Substudie 1 wurde das Dosis-Wirkungs-Verhältnis sowie die Verträglichkeit von Upadacitinib gegenüber Placebo in der Induktionsphase untersucht. In die Studie wurden erwachsene Patienten zwischen 18 und 75 Jahren mit mittelschwerer bis schwerer aktiver CU eingeschlossen, welche auf eine konventionelle oder biologische Therapie nicht oder nicht mehr adäquat ansprechen oder eine entsprechende Intoleranz aufweisen. Die Patienten der U-ACHIEVE Substudie 1 wurden in einer 8-wöchigen Induktionsphase mit Upadacitinib oder Placebo behandelt. Haupteinschlusskriterien der Studie waren hierbei:

- Diagnose der CU mindestens 90 Tage vor Baseline, bestätigt durch Koloskopie während der Screening-Periode, mit Ausschluss einer aktuellen Infektion, Kolon-Dysplasie und/oder Malignität,

- Diagnose einer mittelschweren bis schweren aktiven CU per Adaptiertem Mayo Score mit einem Wert von 5 bis 9 (bestehend aus den Subscores Stuhlfrequenz, rektale Blutungen und endoskopischer Befund), wobei der Subscore des endoskopischen Befunds einen Wert 2 oder 3 aufweisen musste und

- Unzureichendes oder ausbleibendes Ansprechen oder Unverträglichkeit gegenüber mindestens einer vorherigen Behandlung mit oralen Aminosalicylaten, Kortikosteroiden, Immunsuppressiva und/oder biologischen Therapien.

Weitere Ein- und Ausschlusskriterien der Studie sind in Anhang 4-E aufgeführt.

Der Einschluss von Patienten mit unzureichendem Ansprechen auf eine vorherige biologische Therapie oder einer Intoleranz gegenüber einer biologischen Therapie war mit einem Anteil von bis zu 75 % vorgesehen.

Insgesamt wurden 250 Patienten im Verhältnis 1:1:1:1:1 in fünf Behandlungsarme (Upadacitinib in einer Dosierung von 7,5 mg, 15 mg, 30 mg oder 45 mg oder Placebo) randomisiert. Die Randomisierung erfolgte stratifiziert nach vorheriger biologischer Therapie (ja, nein), Behandlung mit Kortikosteroiden zu Baseline (ja, nein) und Adaptiertem Mayo Score zu Baseline (≤ 7, > 7). Pro Behandlungsarm waren jeweils 50 Patienten vorgesehen, randomisiert wurden in den Upadacitinib-Arm mit einer Dosierung von 7,5 mg 47 Patienten, 15 mg mit 49 Patienten, 30 mg mit 52 Patienten, 45 mg mit 56 Patienten und in den Placebo-Arm 46 Patienten. Basierend auf den Ergebnissen der Substudie 1 wurde eine Dosis von 45 mg Upadacitinib für die Substudie 2 der U-ACHIEVE und die U-ACCOMPLISH als Phase-III-Induktionsstudien festgelegt. Zusätzlich wurden parallel zur primären Analyse auf Basis der randomisierten 250 Patienten zur Dosisfindung weitere 132 Patienten in zwei Upadacitinib-Behandlungsarme mit der Dosierung von 30 mg und 45 mg Upadacitinib randomisiert, die aufgrund separater Auswertungen im vorliegenden Dossier nicht dargestellt werden. Diese zusätzliche Rekrutierung wurde vorgenommen, um eine ausreichende Anzahl an Patienten mit klinischem Ansprechen für den Einschluss in die Erhaltungsstudie der U-ACHIEVE (Substudie 3) sicherzustellen, ohne dabei die laufenden Studienaktivitäten und Analysen der Substudie 1 zu unterbrechen. Da die Zulassung der Behandlung in der Induktionsphase nur die Dosierungen von 45mg Upadacitinib umfasst, werden die weiteren Dosierungen von Upadacitinib im Folgenden nicht dargestellt. Die Analyse der Wirksamkeit erfolgte anhand aller randomisierter Patienten der zugelassenen Dosierung, die mindestens eine Dosis der Studienmedikation erhalten haben, entsprechend der Randomisierung und umfasst insgesamt 102 Patienten (Upadacitinib: 56 Patienten, Placebo 46 Patienten). Zur Untersuchung der Sicherheit erfolgte die Analyse anhand aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben, entsprechend der Behandlung, die sie erhalten haben und umfasst ebenfalls 102 Patienten.

Der primäre Endpunkt der Substudie 1 ist der Anteil der Patienten mit klinischer Remission per Adaptiertem Mayo Score zu Woche 8, definiert über das Erreichen eines Adaptierten Mayo Score ≤ 2 mit einem Subscore der Stuhlfrequenz ≤ 1, einem Subscore für rektale Blutungen = 0 und einem endoskopischen Subscore ≤ 1.

U-ACHIEVE Substudie 2

In der Studie U-ACHIEVE Substudie 2 wurde die Wirksamkeit und Sicherheit von Upadacitinib in der Dosierung 45 mg gegenüber Placebo untersucht. In die Studie wurden erwachsene Patienten zwischen 18 und 75 Jahren mit mittelschwerer bis schwerer aktiver CU, welche auf eine konventionelle oder biologische Therapie nicht oder nicht mehr adäquat ansprechen oder eine entsprechende Intoleranz aufweisen, eingeschlossen. Jugendliche Patienten im Alter von 16 bis 17 Jahren konnten, sofern von der jeweiligen lokalen Regulierungs- bzw. Gesundheitsbehörde zugelassen, ebenfalls eingeschlossen werden, sofern sie ein Körpergewicht von ≥ 40 kg und ein Tanner-Stadium von 5 aufweisen. In Deutschland wurden nur Patienten im Alter von ≥ 18 Jahren eingeschlossen. Die Patienten der Studie U-ACHIEVE Substudie 2 wurden in einer 8-wöchigen Induktionsphase mit Upadacitinib in einer Dosierung von 45 mg oder Placebo behandelt. Die 8-wöchige Induktionsphase entspricht Teil 1 der Substudie 2 und wurde mit einer Patientenzahl von insgesamt 462 Patienten geplant. Haupteinschlusskriterien der Studie waren analog zu Substudie 1.

Die Ein- und Ausschlusskriterien der Studie sind in Anhang 4-E aufgeführt.

Der Einschluss von Patienten ohne vorheriges unzureichendes Ansprechen auf eine biologische Therapie oder eine Intoleranz gegenüber einer biologischen Therapie war auf einen Anteil von mindestens 25 % bis zu maximal 50 % festgelegt.

Insgesamt wurden 474 Patienten im Verhältnis 2:1 in die beiden Behandlungsarme randomisiert. In den Upadacitinib-Arm in der Dosierung von 45 mg wurden 319 Patienten und in den Placebo-Arm 155 Patienten eingeschlossen. Die Randomisierung erfolgte stratifiziert nach Ansprechen auf eine vorherige biologische Therapie (Bio-IR-Status: Bio-IR, Non-Bio-IR), Behandlung mit Kortikosteroiden zu Baseline (ja, nein) und Adaptiertem Mayo Score zu Baseline (≤ 7, > 7). Patienten mit unzureichendem Ansprechen auf eine vorherige biologische Therapie (Bio-IR-Patienten) wurden in einem weiteren Schritt nach Anzahl der vorherigen biologischen Therapien (≤ 1, > 1) stratifiziert. Patienten, die kein unzureichendes Ansprechen auf eine vorherige biologische Therapie hatten (Non-Bio-IR-Patienten), wurden ebenfalls erneut stratifiziert nach Patienten, die bereits eine biologische Therapie erhalten haben bzw. Patienten, die noch keine biologische Therapie erhalten haben (bio-naive Patienten). In Teil 2 der Substudie 2 wurden Patienten ohne klinisches Ansprechen zu Woche 8 eingeschlossen und erhielten in einer offenen Behandlung Upadacitinib in der Dosierung von 45 mg für weitere 8 Wochen unabhängig von der Behandlung in Teil 1 der Substudie 2. Die Auswertung der Wirksamkeit und der Sicherheit erfolgt entsprechend der primären Analyse anhand der Analysepopulation des doppelt-verblindeten Teil 1 der Studie. In die Wirksamkeitsanalyse wurden alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben, eingeschlossen und entsprechend der Randomisierung ausgewertet. Zur Untersuchung der Sicherheit erfolgte die Analyse anhand aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben und wurde entsprechend der Behandlung, die sie erhalten haben, ausgewertet. Beide Analysepopulationen umfassen die 474 randomisierten Patienten.

Der primäre Endpunkt der Substudie 2 ist der Anteil der Patienten mit klinischer Remission per Adaptiertem Mayo Score zu Woche 8, definiert über das Erreichen eines Adaptierten Mayo Score ≤ 2 mit einem Subscore der Stuhlfrequenz ≤ 1 und nicht größer als Baseline, einem Subscore für rektale Blutungen = 0 und einem endoskopischem Subscore ≤ 1.

U-ACCOMPLISH

In der Studie U-ACCOMPLISH wurde die Wirksamkeit und Sicherheit von Upadacitinib in der Dosierung 45 mg gegenüber Placebo untersucht. In die Studie wurden erwachsene Patienten zwischen 18 und 75 Jahren mit mittelschwerer bis schwerer aktiver CU, welche auf eine konventionelle oder biologische Therapie nicht oder nicht mehr adäquat ansprechen oder eine entsprechende Intoleranz aufweisen, eingeschlossen. Jugendliche Patienten im Alter von 16 bis 17 Jahren konnten, sofern von der jeweiligen lokalen Regulierungs- bzw. Gesundheitsbehörde zugelassen, ebenfalls eingeschlossen werden, sofern sie ein Körpergewicht von ≥ 40 kg und ein Tanner-Stadium von 5 aufweisen. In Deutschland wurden nur Patienten im Alter von ≥ 18 Jahren eingeschlossen. Die Patienten der Studie U-ACCOMPLISH wurden in einer 8-wöchigen Induktionsphase mit Upadacitinib in einer Dosierung von 45 mg oder Placebo behandelt. Die 8-wöchige Induktionsphase entspricht Teil 1 der Studie und wurde mit einer Patientenzahl von insgesamt 462 Patienten geplant. Haupteinschlusskriterien der Studie waren analog zu Substudie 1 und 2 der U-ACHIEVE Studie.

Die Ein- und Ausschlusskriterien der Studie sind in Anhang 4-E aufgeführt.

Der Einschluss von Patienten ohne vorheriges unzureichendes Ansprechen auf eine biologische Therapie oder eine Intoleranz gegenüber einer biologischen Therapie war mit einem Anteil von mindestens 25 % bis zu maximal 50 % vorgesehen.

Insgesamt wurden 522 Patienten im Verhältnis 2:1 in die beiden Behandlungsarme der Studie randomisiert. In den Upadacitinib-Arm in der Dosierung von 45 mg wurden 345 Patienten und in den Placebo-Arm 177 Patienten eingeschlossen. Die Randomisierung erfolgte stratifiziert nach Ansprechen auf eine vorherige biologische Therapie (Bio-IR-Status: Bio-IR, Non-Bio-IR), Behandlung mit Kortikosteroiden zu Baseline (ja, nein) und Adaptiertem Mayo Score zu Baseline (≤ 7, > 7) im Verhältnis 2:1. Patienten mit unzureichendem Ansprechen auf eine vorherige biologische Therapie (Bio-IR-Patienten) wurden in einem weiteren Schritt nach Anzahl der vorherigen biologischen Therapien (≤ 1, > 1) stratifiziert. Patienten, die kein unzureichendes Ansprechen auf eine vorherige biologische Therapie hatten (Non-Bio-IR-Patienten), wurden ebenfalls erneut stratifiziert nach Patienten, die bereits eine biologische Therapie erhalten haben bzw. Patienten, die noch keine biologische Therapie erhalten haben (bio-naive Patienten). In Teil 2 der Studie wurden Patienten ohne klinisches Ansprechen zu Woche 8 eingeschlossen und erhielten in einer offenen Behandlung Upadacitinib in der Dosierung von 45 mg für weitere 8 Wochen unabhängig von der Behandlung in Teil 1 der Studie. Die Auswertung der Wirksamkeit und der Sicherheit erfolgt entsprechend der primären Analyse anhand der Analysepopulation des doppelt-verblindeten Teil 1 der Studie. In der Wirksamkeitsanalyse wurden alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben, berücksichtigt und entsprechend der Randomisierung

ausgewertet. Die Population umfasst 515 Patienten (Upadacitinib: 341 Patienten, Placebo: 174 Patienten). Zur Untersuchung der Sicherheit erfolgte die Analyse anhand aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben und sie wurde entsprechend der Behandlung, die sie erhalten haben, ausgewertet. Die Population umfasst 521 Patienten (Upadacitinib: 344 Patienten, Placebo: 177 Patienten).

Der primäre Endpunkt der Studie U-ACCOMPLISH ist der Anteil der Patienten mit klinischer Remission per Adaptiertem Mayo Score zu Woche 8, definiert über das Erreichen eines Adaptierten Mayo Score ≤ 2 mit einem Subscore der Stuhlfrequenz ≤ 1 und nicht größer als Baseline, einem Subscore für rektale = 0 und einem endoskopischem Subscore ≤ 1.

Ustekinumab

UNIFI Induktionsstudie

In der Induktionsstudie der Studie UNIFI wurde die Wirksamkeit und Sicherheit von Ustekinumab gegenüber Placebo in der Induktionsphase bei Patienten mit mittelschwerer bis schwerer aktiver CU untersucht. In die Studie wurden erwachsene Patienten ab 18 Jahren mit mittelschwerer bis schwerer aktiver CU eingeschlossen und die Haupteinschlusskriterien waren hierbei:

- Mayo Gesamtscore zu Baseline der Induktionsphase von ≥ 6 bis ≤ 12 mit einem Endoskopie Subscore von ≥ 2 durch einen zentralen Prüfer,
- Unzureichendes oder ausbleibendes Ansprechen oder Verlust des Ansprechens auf eine biologische Therapie mit einem oder mehreren TNF-Inhibitor oder Vedolizumab in einer für die Behandlung der CU zugelassenen Dosierung oder bei Intoleranz gegenüber dieser Wirkstoffe oder
- Patienten, die noch keine biologische Therapie erhalten haben oder kein unzureichendes Ansprechen auf eine biologische Therapie oder Intoleranz gegenüber einer biologischen Therapie zeigten und bereits mindestens eine konventionelle Therapie mit CU-bezogenen Medikationen wie oralen Kortikosteroiden, Aminosalicylaten, 6-Mercaptopurin (6-MP), Azathioprin (AZA) oder Methotrexat (MTX) erhalten haben.

Der Anteil von Patienten mit unzureichendem Ansprechen auf eine vorherige biologische Therapie wurde auf mindestens 40 % und maximal 50 % der gesamten Studienpopulation festgelegt.

Insgesamt wurden 961 Patienten in einem Verhältnis von 1:1:1 in drei Behandlungsarme (Ustekinumab 130 mg: 320 Patienten, Ustekinumab ~ 6 mg/kg: 322 Patienten, Placebo: 319 Patienten) randomisiert. Die Randomisierung erfolgte stratifiziert nach unzureichendem Ansprechen auf eine vorherige biologische Therapie (ja, nein) und geografischer Region (Osteuropa, Asien, Rest der Welt). Patienten mit klinischem Ansprechen nach Mayo Gesamtscore auf eine 8-wöchige Induktionsphase wurden in die Erhaltungsstudie randomisiert.

Patienten ohne klinisches Ansprechen wurden weitere 8 Wochen mit ~6 mg/kg Ustekinumab behandelt, unabhängig von der randomisierten Behandlung. Da die Zulassung der Behandlung in der Induktionsphase nur die Dosierungen von ~6 mg/kg Ustekinumab umfasst, wird die Dosierung von 130 mg Ustekinumab im Folgenden nicht dargestellt. Die Analyse der Wirksamkeit erfolgt anhand aller randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben entsprechend der Randomisierung und umfasst 641 Patienten (Ustekinumab: 322 Patienten, Placebo: 319 Patienten). Zur Untersuchung der Sicherheit erfolgt die Analyse anhand aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben entsprechend der Behandlung, die sie erhalten haben, und umfasst 639 Patienten (Ustekinumab: 320 Patienten, Placebo: 319 Patienten).

Der primäre Endpunkt der UNIFI Induktionsstudie ist das Erreichen der klinischen Remission per Mayo Gesamtscore zu Woche 8, definiert als Gesamtscore ≤ 2 und keinem Subscore > 1.

Erhaltungsstudien

Upadacitinib

U-ACHIEVE Substudie 3

In der Studie U-ACHIEVE Substudie 3 wurde die Wirksamkeit und Sicherheit von Upadacitinib in der Dosierung 15 mg und 30 mg gegenüber Placebo in der Erhaltungsphase untersucht. In die Studie wurden Patienten der drei Induktionsstudien Substudie 1 und Substudie 2 der Studie U-ACHIEVE sowie der Studie U-ACCOMPLISH nach klinischem Ansprechen zu Woche 8 bzw. Woche 16 eingeschlossen. Die Substudie 3 der U-ACHIEVE umfasst insgesamt 4 Kohorten, die sich aus der Behandlung in der jeweiligen Induktionsstudie sowie dem Ansprechen auf die Behandlung in der Induktionsphase ergeben und wurde mit insgesamt 750 Patienten geplant. Die Ein- und Ausschlusskriterien entsprechen den Ein- und Ausschlusskriterien der Induktionsstudien. Patienten mit fehlendem Endoskopiebefund zu Woche 8 bzw. 16 konnten nicht in die Substudie 3 aufgenommen werden. Patienten, deren Induktionsphase im Rahmen der Studie U-ACCOMPLISH erfolgte, mussten eine zusätzliche Einwilligungserklärung für die Studie U-ACHIEVE Substudie 3 unterzeichnen. Die Abschlussvisite der Induktionsstudie zu Woche 8 bzw. Woche 16 entsprach der Baseline-Visite der Erhaltungsstudie.

Die Patienten der Substudie 3 der U-ACHIEVE wurden nach der Behandlung in der Induktionsphase für weitere 52 Wochen behandelt. Insgesamt wurden 1.046 Patienten unterschiedlichen Kohorten zugeordnet und anschließend randomisiert, sofern in der jeweiligen Kohorte erforderlich. In Kohorte 1 wurden Patienten mit klinischem Ansprechen auf eine 8-wöchige Induktionsphase mit Upadacitinib in den Dosierungen 15 mg, 30 mg oder 45 mg eingeschlossen. Patienten, die in Teil 1 der U-ACHIEVE Substudie 2 oder der Studie U-ACCOMPLISH mit Placebo behandelt wurden und in Teil 2 der jeweiligen Studie (8-wöchige, offene Verlängerung unter Behandlung mit Upadacitinib 45 mg) ein klinisches Ansprechen hatten, wurden ebenfalls in Kohorte 1 aufgenommen. Kohorte 2 umfasst Patienten, die ein klinisches Ansprechen auf eine Behandlung mit Placebo zeigten. Kohorte 3 umfasst Patienten, die ein klinisches Ansprechen auf eine verlängerte Induktionsbehandlung von

16 Wochen (Teil 2 der U-ACHIEVE Substudie 2 oder Teil 2 der Studie U-ACCOMPLISH) mit Upadacitinib zeigten. In Kohorte 4 wurden Patienten aus Substudie 1 eingeschlossen, die mit einer Dosierung von 7,5 mg Upadacitinib behandelt wurden. Patienten ohne klinisches Ansprechen zu Woche 8 in Substudie 1 bzw. ohne klinisches Ansprechen zu Woche 16 in Teil 2 der Substudie 2 oder in Teil 2 der Studie U-ACCOMPLISH konnten in eine weitere, offene Verlängerungsstudie (M14-533) eingeschlossen werden und mit Upadacitinib in einer Erhaltungsphase behandelt werden. Für das vorliegende Dossier und die Nutzenbewertung ist Kohorte 1 der U-ACHIEVE Substudie 3 relevant, daher werden im Folgenden die Kohorten 2 bis 4 nicht weiter beschrieben oder dargestellt.

In Kohorte 1 wurden Patienten aus den beschriebenen Induktionsstudien U-ACHIEVE Substudie 1 (Upadacitinib 30 mg oder 45 mg), Substudie 2 oder U-ACCOMPLISH eingeschlossen. Die Randomisierung erfolgte im Verhältnis 1:1:1 in drei Behandlungsarme: Upadacitinib 15 mg oder Upadacitinib 30 mg oder Placebo. Patienten, die ein klinisches Ansprechen auf Upadacitinib 15 mg in der U-ACHIEVE Substudie 1 zeigten, wurden ebenfalls in Kohorte 1 randomisiert, konnten jedoch nicht in den Behandlungsarm mit 30 mg randomisiert werden, um eine höhere Dosierung in der Erhaltung als in der Induktion zu vermeiden. Stratifizierungsfaktoren waren für Patienten aus der U-ACHIEVE Substudie 2 und der Studie U-ACCOMPLISH das Ansprechen auf eine vorherige biologische Therapie vor der Induktionsphase (Bio-IR-Status: Bio-IR, Non-Bio-IR), das Vorliegen einer klinischen Remission zu Baseline (ja, nein) und die Behandlung mit Kortikosteroiden zu Baseline (ja, nein). Für Patienten aus der U-ACHIEVE Substudie 1 erfolgte die Randomisierung stratifiziert nach Dosierung von Upadacitinib in der Induktionsphase (15 mg, 30 mg, 45 mg) und dem Ansprechen auf eine vorherige biologische Therapie vor der Induktionsphase (Bio-IR-Status: Bio-IR, Non-Bio-IR). Insgesamt wurden 847 Patienten in Kohorte 1 randomisiert (Upadacitinib 15 mg: 289 Patienten; Upadacitinib 30 mg: 276 Patienten, Placebo: 282 Patienten). Im Amendment 3 des Studienprotokolls wurde die Dauer der Erhaltungsphase von 44 Wochen auf 52 Wochen verlängert. Die Auswertung zur Wirksamkeit und Verträglichkeit erfolgte nach dem ITT-Prinzip und die Analysepopulationen wurden entsprechend dem SAP unter Berücksichtigung der verlängerten Behandlung in der Erhaltungsphase ab Amendment 3 des Studienprotokolls festgelegt. Für die Wirksamkeitsanalyse wurden die ersten 451 randomisierten Patienten der Kohorte 1, die ab Amendment 3 in die Studie aufgenommen wurden und mindestens eine Dosis der Studienmedikation erhalten haben, entsprechend der randomisierten Behandlung ausgewertet (Upadacitinib 15 mg: 148 Patienten; Upadacitinib 30 mg: 154 Patienten, Placebo: 149 Patienten). Die Analysepopulation entspricht der primären Analysepopulation für die Endpunkte zur Wirksamkeit des Studienprotokolls und wird im vorliegenden Dossier dargestellt. Für die Auswertung zur Sicherheit und Verträglichkeit wurden alle Patienten der Kohorte 1, die mindestens eine Dosis der Studienmedikation erhalten haben, berücksichtigt, unabhängig unter welchem Studienprotokoll sie in die Studie eingeschlossen wurden und umfasst insgesamt 746 Patienten (Upadacitinib 15 mg: 250 Patienten; Upadacitinib 30 mg: 251 Patienten, Placebo: 245 Patienten). Die Safety Population wird im SAP der Studie als Safety Population C bezeichnet und wird im vorliegenden Dossier dargestellt.

Der primäre Endpunkt der U-ACHIEVE Substudie 3 ist der Anteil der Patienten mit klinischer Remission per Adaptiertem Mayo Score zu Woche 52, definiert über das Erreichen eines Adaptierten Mayo Score ≤ 2 mit einem Subscore der Stuhlfrequenz ≤ 1 und nicht größer als Baseline, einem Subscore für rektale Blutungen = 0 und einem endoskopischem Subscore ≤ 1.

Ab Woche 4 der Erhaltungsphase können Patienten, die nicht mehr ansprechen und nach Ermessen des Prüfarztes eine medizinische Behandlung benötigen, eine Rescue-Therapie erhalten. Diese wird in Form der Einleitung oder Erhöhung der Dosierung zugelassener CU-bezogener Medikamente nach Ermessen des Prüfarztes gegeben, um neue oder sich verschlimmernde CU-Symptome zu behandeln. Die zulässigen CU-bezogenen Medikamente sind lokal wirkende, orale oder intravenöse Kortikosteroide, Aminosalicylate, MTX oder CU-bezogene Antibiotika.

Ustekinumab

UNIFI Erhaltungsstudie

In der Erhaltungsstudie der Studie UNIFI wurde die Wirksamkeit und Sicherheit von Ustekinumab gegenüber Placebo in der Erhaltungsphase bei Patienten mit mittelschwerer bis schwerer aktiver CU nach klinischem Ansprechen auf eine Induktionsphase untersucht. Die Randomisierung der Patienten mit klinischem Ansprechen auf die Induktionsphase erfolgte stratifiziert im Verhältnis 1:1:1 in die drei Behandlungsarme Ustekinumab 90 mg alle 12 Wochen, Ustekinumab 90 mg alle 8 Wochen und Placebo. Stratifizierungsfaktoren waren die Dosis von Ustekinumab in der Induktionsphase (130 mg, ~6 mg/kg), die klinische Remission zu Baseline der Erhaltungsphase (ja, nein) und die Einnahme oraler Kortikosteroide zu Baseline (ja, nein). Insgesamt 523 Patienten wurden in die Erhaltungsstudie randomisiert (Ustekinumab 90 mg alle 12 Wochen: 172 Patienten; Ustekinumab 90 mg alle 8 Wochen: 176 Patienten; Placebo: 175 Patienten), die nach klinischem Ansprechen per Mayo Gesamtscore in der Induktionsphase eingeschlossen wurden. Die Analyse der Wirksamkeit erfolgte anhand aller randomisierten Patienten entsprechend der Randomisierung, die mindestens eine Dosis der Studienmedikation erhalten hatten und umfasst die 523 randomisierten Patienten. Zur Untersuchung der Sicherheit erfolgt die Analyse anhand aller Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben und wurde entsprechend der Behandlung, die sie erhalten haben, ausgewertet und umfasst ebenfalls die 523 randomisierten Patienten.

Der primäre Endpunkt der UNIFI Erhaltungsstudie ist das Erreichen der klinischen Remission per Mayo Gesamtscore zu Woche 44 der Erhaltungsphase, definiert als Gesamtscore ≤ 2 und keinem Subscore > 1.

Charakterisierung der Studienpopulationen

Induktionsstudien

U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH

Über alle berichteten Merkmale bezüglich der demografischen Charakteristika, der krankheitsspezifischen Charakteristika sowie der medizinischen Vorgeschichte wurden

zwischen den Behandlungsarmen in den Substudien 1 und 2 der U-ACHIEVE sowie der Studie U-ACCOMPLISH keine klinisch relevanten Unterschiede festgestellt und auch zwischen den Studien zeigten sich keine relevanten Unterschiede.

UNIFI Induktionsstudie

Über alle berichteten Merkmale bezüglich der demografischen Charakteristika, der krankheitsspezifischen Charakteristika sowie der medizinischen Vorgeschichte wurden zwischen den Behandlungsarmen in der UNIFI Induktionsstudie keine klinisch relevanten Unterschiede festgestellt.

Erhaltungsstudien

U-ACHIEVE Substudie 3

Über alle berichteten Merkmale bezüglich der demografischen Charakteristika, der krankheitsspezifischen Charakteristika sowie der medizinischen Vorgeschichte wurden zwischen den Behandlungsarmen in der Substudie 3 der U-ACHIEVE keine klinisch relevanten Unterschiede festgestellt.

UNIFI Erhaltungsstudie

Über alle berichteten Merkmale bezüglich der demografischen Charakteristika, der krankheitsspezifischen Charakteristika sowie der medizinischen Vorgeschichte wurden zwischen den Behandlungsarmen in der UNIFI Erhaltungsstudie keine klinisch relevanten Unterschiede festgestellt.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Studien U-ACCOMPLISH, U-ACHIEVE und UNIFI sind aus mehreren Gründen auf den deutschen Versorgungskontext übertragbar. Die gemäß Zulassung vorausgesetzten Vortherapien (Bio-IR / Non-Bio-IR) wurden in den Ein- und Ausschlusskriterien der Studien berücksichtigt. Die Diagnose- und Behandlungskriterien entsprechen in den Studien den Empfehlungen der deutschen Leitlinie. Die Dosierung und Verabreichung von Upadacitinib und Ustekinumab entspricht in den Studien ebenfalls den Vorgaben der jeweiligen Fachinformation. Des Weiteren wurden beide Studien vorwiegend in Industrieländern durchgeführt. Es konnten daher für die Studien keine Hinweise identifiziert werden, die einer Übertragbarkeit widersprechen.

4.3.2.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

Tabelle 4-53: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
Induktionsstudien							
U-ACHIEVE Substudie 1	ja	ja	ja	ja	ja	ja	niedrig
U-ACHIEVE Substudie 2	ja	ja	ja	ja	ja	ja	niedrig
U-ACCOMPLISH	ja	ja	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	ja	ja	ja	ja	ja	ja	niedrig
Erhaltungsstudien							
U-ACHIEVE Substudie 3	ja	ja	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	ja	ja	ja	ja	ja	ja	niedrig
RCT: Randomisierte kontrollierte Studie							

Induktionsstudien

U-ACHIEVE Substudie 1 und Substudie 2, U-ACCOMPLISH

Bei der Studie U-ACHIEVE Substudie 1 handelt es sich um eine randomisierte, multizentrische, doppelblinde Phase-IIb-Studie. Bei den Studien U-ACHIEVE Substudie 2 und U-ACCOMPLISH handelt es sich um randomisierte, multizentrische, doppelblinde Phase-III-Studien. Die Randomisierung der Studienteilnehmer in die Behandlungsarme erfolgte mittels einer computergenerierten Randomisierungssequenz auf die Behandlungsarme Upadacitinib und Placebo. Die Zuteilung erfolgte jeweils durch ein Interaktives Response Technologie System (Interactive Response Technology, IRT), wodurch die Verdeckung der Gruppenzuteilung gegeben war. Die Randomisierung in der Studie U-ACHIEVE Substudie 1 erfolgte stratifiziert nach vorheriger biologischer Therapie (ja, nein), Behandlung mit Kortikosteroiden zu Baseline (ja, nein) und Adaptiertem Mayo Score zu Baseline (≤ 7, > 7). In der Studie U-ACHIEVE Substudie 2 und der Studie U-ACCOMPLISH erfolgte die Randomisierung stratifiziert nach Ansprechen auf eine vorherige biologische Therapie (Bio-IR-Status: Bio-IR, Non-Bio-IR), Behandlung mit Kortikosteroiden zu Baseline (ja, nein) und

Adaptiertem Mayo Score zu Baseline (≤ 7, > 7). Patienten mit unzureichendem Ansprechen auf eine vorherige biologische Therapie (Bio-IR-Patienten) wurden in einem weiteren Schritt nach Anzahl der vorherigen biologischen Therapien (≤ 1, > 1) stratifiziert. Patienten, die kein unzureichendes Ansprechen auf eine vorherige biologische Therapie hatten (Non-Bio-IR-Patienten), wurden ebenfalls erneut stratifiziert nach Patienten, die bereits eine biologische Therapie erhalten haben bzw. Patienten, die noch keine biologische Therapie erhalten haben (bio-naive Patienten). Sowohl die Patienten als auch die behandelnden Personen waren über die gesamte Behandlungsdauer der Induktionsphase von 8 Wochen verblindet. Endpunkte wurden, soweit im Rahmen des indirekten Vergleichs möglich, gemäß dem SAP berichtet. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet und entsprechend in der Operationalisierung beschrieben und somit kann von einer ergebnisunabhängigen Berichterstattung ausgegangen werden. Es lagen auch keine sonstigen Aspekte vor, die auf ein Verzerrungspotenzial auf Studienebene hindeuten. Daher wird das Verzerrungspotenzial für die Studie U-ACHIEVE Substudie 3 als niedrig bewertet.

UNIFI Induktionsstudie

Bei der UNIFI Induktionsstudie handelt es sich um eine randomisierte, doppelblinde, multizentrische, placebokontrollierte Phase-III-Studie. Die Randomisierung der Studienteilnehmer in die Behandlungsarme erfolgte mittels einer computergenerierten Randomisierungssequenz auf die Behandlungsarme. Die Zuteilung erfolgte jeweils durch ein IRT, wodurch die Verdeckung der Gruppenzuteilung gegeben war. Die Randomisierung erfolgte stratifiziert nach unzureichendem Ansprechen auf eine vorherige biologische Therapie (ja, nein) und geografischer Region (Osteuropa, Asien, Rest der Welt). Sowohl die Patienten als auch die behandelnden Personen waren über die gesamte Behandlungsdauer verblindet. Endpunkte wurden, soweit im Rahmen des indirekten Vergleichs möglich, gemäß dem SAP berichtet. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet und entsprechend in der Operationalisierung beschrieben und somit kann von einer ergebnisunabhängigen Berichterstattung ausgegangen werden. Es lagen auch keine sonstigen Aspekte vor, die auf ein Verzerrungspotenzial auf Studienebene hindeuten. Daher wird das Verzerrungspotenzial für die Studie UNIFI Induktionsstudie als niedrig bewertet.

Erhaltungsstudien

U-ACHIEVE Substudie 3

Bei der Studie U-ACHIEVE Substudie 3 handelt es sich um eine randomisierte, multizentrische, doppelblinde Phase-III-Studie. Die Randomisierung der Studienteilnehmer in die Behandlungsarme erfolgte mittels einer computergenerierten Randomisierungssequenz im Verhältnis 1:1:1 auf die Behandlungsarme Upadacitinib und Placebo. Die Zuteilung erfolgte jeweils durch ein IRT, wodurch die Verdeckung der Gruppenzuteilung gegeben war. Die Randomisierung erfolgte für Patienten, die in der U-ACHIEVE Substudie 2 oder der Studie U-ACCOMPLISH behandelt wurden, stratifiziert nach Ansprechen auf eine vorherige biologische Therapie vor der Induktionsphase (Bio-IR-Status: Bio-IR, Non-Bio-IR), Vorliegen

einer klinischen Remission zu Baseline (ja, nein) und Behandlung mit Kortikosteroiden zu Baseline (ja, nein). Für Patienten aus der U-ACHIEVE Substudie 1 erfolgte die Randomisierung stratifiziert nach Dosierung in der Induktionsphase (15 mg, 30 mg, 45 mg) und dem Ansprechen auf eine vorherige biologische Therapie vor der Induktionsphase (Bio-IR-Status: Bio-IR, Non-Bio-IR). Sowohl die Patienten als auch die behandelnden Personen waren über die gesamte Behandlungsdauer verblindet. Endpunkte wurden, soweit im Rahmen des indirekten Vergleichs möglich, gemäß dem SAP berichtet. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet und entsprechend in der Operationalisierung beschrieben und somit kann von einer ergebnisunabhängigen Berichterstattung ausgegangen werden. Es lagen auch keine sonstigen Aspekte vor, die auf ein Verzerrungspotenzial auf Studienebene hindeuten. Daher wird das Verzerrungspotenzial für die Studie U-ACHIEVE Substudie 3 als niedrig bewertet.

UNIFI Erhaltungsstudie

Bei der Studie UNIFI Erhaltungsstudie handelt es sich um eine randomisierte, multizentrische, doppelblinde Phase-III-Studie. Die Randomisierung der Studienteilnehmer in die Behandlungsarme erfolgte mittels einer computergenerierten Randomisierungssequenz im Verhältnis 1:1:1 auf die Behandlungsarme. Die Zuteilung erfolgte jeweils durch ein IRT, wodurch die Verdeckung der Gruppenzuteilung gegeben war. Die Randomisierung erfolgte stratifiziert nach der Dosis von Ustekinumab in der Induktionsphase (130 mg, ~6 mg/kg), klinischer Remission zu Baseline der Erhaltungsphase (ja, nein) und der Einnahme oraler Kortikosteroide (ja, nein). Sowohl die Patienten als auch die behandelnden Personen waren über die gesamte Behandlungsdauer verblindet. Endpunkte wurden, soweit im Rahmen des indirekten Vergleichs möglich, gemäß dem SAP berichtet. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet und entsprechend in der Operationalisierung beschrieben und somit kann von einer ergebnisunabhängigen Berichterstattung ausgegangen werden. Es lagen auch keine sonstigen Aspekte vor, die auf ein Verzerrungspotenzial auf Studienebene hindeuten. Daher wird das Verzerrungspotenzial für die Studie UNIFI Erhaltungsstudie als niedrig bewertet.

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 4-54: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	Mortalität[a]	Morbidität[b]	Gesundheits-bezogene Lebensqualität[c]	Verträglich-keit[d]
Induktionsstudien				
U-ACHIEVE Substudie 1	ja	ja	ja	ja
U-ACHIEVE Substudie 2	ja	ja	ja	ja
U-ACCOMPLISH	ja	ja	ja	ja
UNIFI Induktionsstudie	ja	ja	ja	ja
Erhaltungsstudien				
U-ACHIEVE Substudie 3	ja	ja	ja	ja
UNIFI Erhaltungsstudie	ja	ja	ja	ja

a: Die Erhebung und Darstellung der Mortalität erfolgte im Rahmen der Erhebung der UE.

b:Umfasst die Endpunkte Klinische Remission, Klinisches Ansprechen, Mukosaheilung, Stuhlfrequenz, rektale Blutungsrate, EQ-5D VAS.

c: Umfasst die Endpunkte IBDQ und SF-36

d: Für die Verträglichkeit wurden im Rahmen des indirekten Vergleiches Alle UE, SUE und UE, die zum Therapieabbruch führten herangezogen. Für die Erhaltungsphase werden für Upadacitinib die schweren UE und die UE, die zum Tod führten sowie die UESI angeführt.

EQ-5D: EuroQol-5 dimensions; IBDQ: Inflammatory Bowel Disease Questionnaire; RCT: Randomisierte kontrollierte Studie; SF-36: Short Form – 36; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; UESI: Unerwünschtes Ereignis von speziellem Interesse; VAS: Visuelle Analogskala

4.3.2.1.3.1 Mortalität – 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 drei 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 4-55: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	•	•	
1	UNIFI Induktionsstudie		•	•
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	•	•	
1	UNIFI Erhaltungsstudie		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

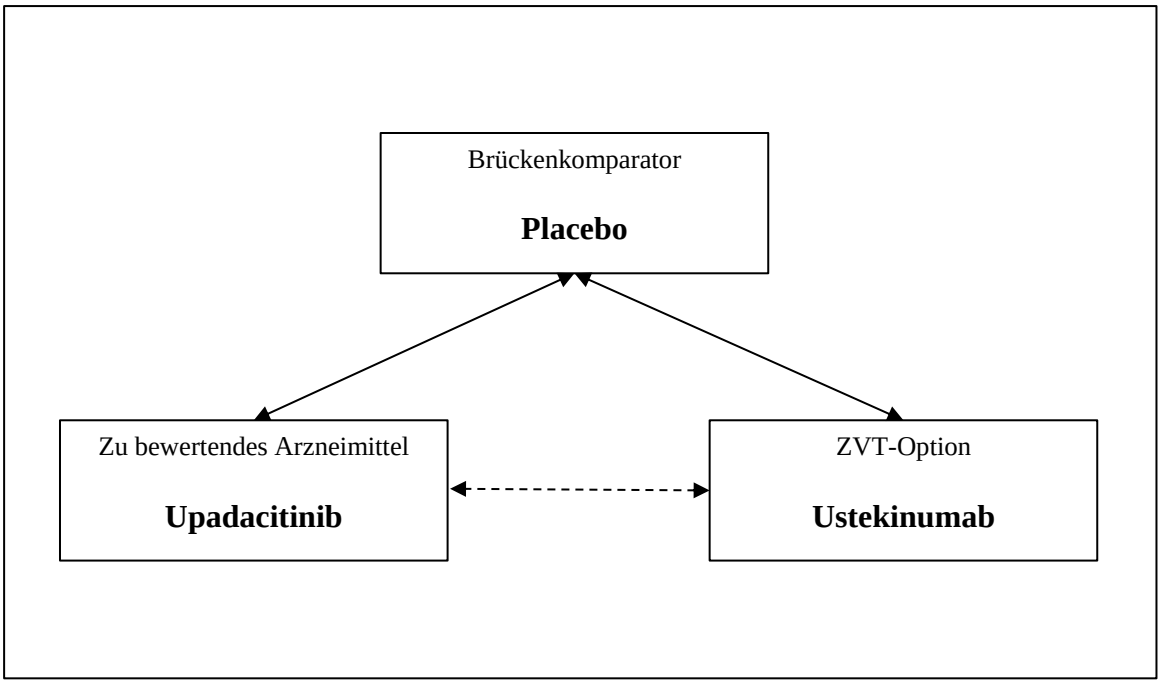

Abbildung 7: Netzwerkstruktur des indirekten Vergleichs

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 der Mortalität

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Die Erhebung der Mortalität erfolgte im Rahmen der UE und wird daher in Abschnitt 4.3.2.1.3.10 dargestellt.
UNIFI Induktionsstudie	Die Erhebung der Mortalität erfolgte im Rahmen der UE und wird daher in Abschnitt 4.3.2.1.3.10 dargestellt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Die Erhebung der Mortalität erfolgte im Rahmen der UE und wird daher in Abschnitt 4.3.2.1.3.10 dargestellt.
UNIFI Erhaltungsstudie	Die Erhebung der Mortalität erfolgte im Rahmen der UE und wird daher in Abschnitt 4.3.2.1.3.10 dargestellt.
UE: Unerwünschtes Ereignis	

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 Mortalität in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Nicht zutreffend, Mortalität wird im Rahmen der UE im Abschnitt 4.3.2.1.3.10 dargestellt					

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
UNIFI Induktionsstudie	Nicht zutreffend, Mortalität wird im Rahmen der UE im Abschnitt 4.3.2.1.3.10 dargestellt					
U-ACHIEVE Substudie 3	Nicht zutreffend, Mortalität wird im Rahmen der UE im Abschnitt 4.3.2.1.3.10 dargestellt					
UNIFI Erhaltungsstudie	Nicht zutreffend, Mortalität wird im Rahmen der UE im Abschnitt 4.3.2.1.3.10 dargestellt					
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie; UE: Unerwünschtes Ereignis						

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend, Mortalität wird im Rahmen der UE im Abschnitt 4.3.2.1.3.10 dargestellt.

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

Nicht zutreffend, Mortalität wird im Rahmen der UE im Abschnitt 4.3.2.1.3.10 dargestellt.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend, Mortalität wird im Rahmen der UE im Abschnitt 4.3.2.1.3.10 dargestellt.

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 Klinische Remission – indirekte Vergleiche aus RCT

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

Tabelle 4-58: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	•	•	
1	UNIFI Induktionsstudie		•	•
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	•	•	
1	UNIFI Erhaltungsstudie		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

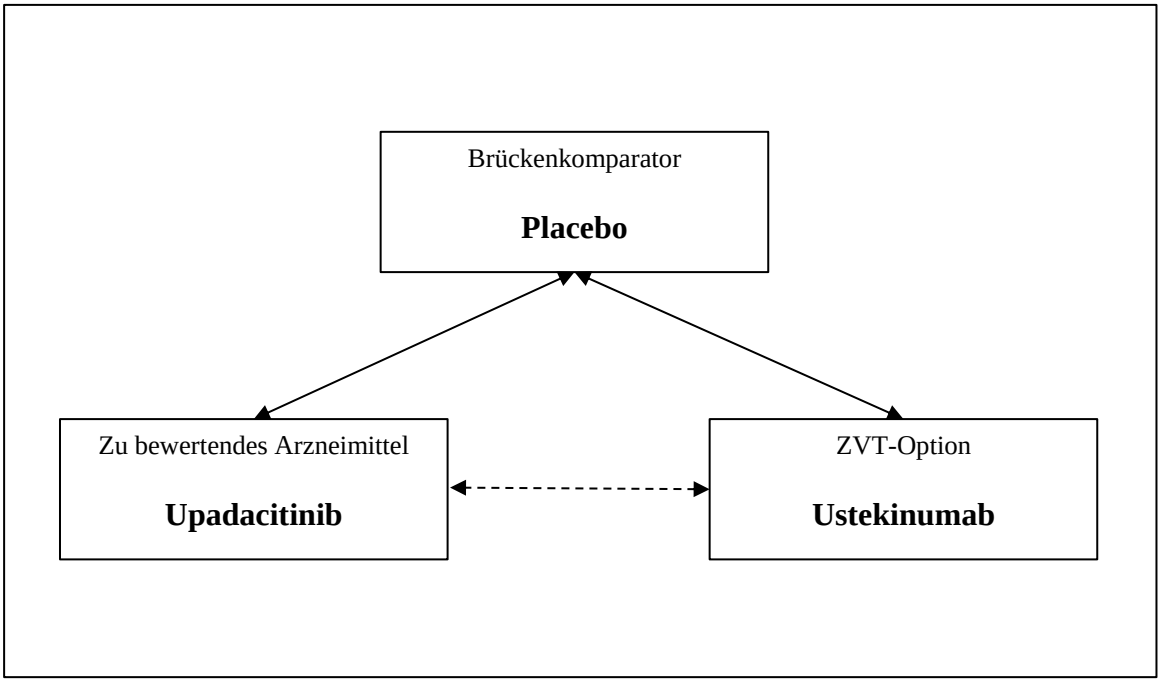

Abbildung 8: Netzwerkstruktur des indirekten Vergleichs

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

Tabelle 4-59: Operationalisierung von klinische Remission

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Die klinische Remission wird anhand des Mayo Gesamtscores beziehungsweise anhand des Adaptierten Mayo Scores beschrieben. Der Mayo Gesamtscore setzt sich aus den Subscores der Stuhlfrequenz, der rektalen Blutung, der Beurteilung durch den Prüfarzt und der Endoskopie zusammen und hat einen Wertebereich von null bis zwölf. Der Adaptierte Mayo Score beinhaltet die gleichen Subscores mit Außnahme der Prüfarztbeurteilung und hat einen Wertebereich von null bis neun. Sowohl für den Mayo Gesamtscore als auch den Adaptierten Mayo Score bilden höhere Werte einen schwereren Krankheitsverlauf ab. Die klinische Remission per Mayo Gesamtscore ist definiert als ein Mayo Gesamtscore ≤ 2 und keinem Subscore > 1. Die klinische Remission per Adaptiertem Mayo Score ist definiert als ein Adaptierter Mayo Score ≤ 2, mit einer Stuhlfrequenz ≤ 1 und die nicht höher als ihr Ausgangswert ist, einer rektalen Blutung von null und einer Endoskopie von ≤ 1. Der Mayo Gesamtscore und der Adaptierte Mayo Score wurden zu Baseline und zu Woche 8 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer klinischen Remission per Mayo Gesamtscore beziehungsweise per Adaptiertem Mayo Score dargestellt. Der Anteil der Patienten beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden mittels NRI- beziehungsweise NRI-C-Analyse ersetzt. Für die Beschreibung der Induktionsphase, wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH meta-analytisch zusammengefasst.

Studie	Operationalisierung
UNIFI Induktionsstudie	Die klinische Remission wird anhand des Mayo Gesamtscores beziehungsweise anhand des Adaptierten Mayo Scores beschrieben. Der Mayo Gesamtscore setzt sich aus den Subscores der Stuhlfrequenz, der rektalen Blutung, der Beurteilung durch den Prüfarzt und der Endoskopie zusammen und hat einen Wertebereich von null bis zwölf. Der Adaptierte Mayo Score beinhaltet die gleichen Subscores mit Außnahme der Prüfarztbeurteilung und hat einen Wertebereich von null bis neun. Sowohl für den Mayo Gesamtscore als auch den Adaptierten Mayo Score bilden höhere Werte einen schwereren Krankheitsverlauf ab. Die klinische Remission per Mayo Gesamtscore ist definiert als ein Mayo Gesamtscore ≤ 2 und keinem Subscore > 1. Die klinische Remission per Adaptiertem Mayo Score ist definiert als eine absolute Stuhlhäufigkeit ≤ 3, einer rektalen Blutung von null und einer Endoskopie von ≤ 1. Der Mayo Gesamtscore und der Adaptierte Mayo Score wurden zu Baseline und zu Woche 8 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer klinischen Remission per Mayo Gesamtscore beziehungsweise per Adaptiertem Mayo Score dargestellt. Die Auswertung zur klinischen Remission beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Die klinische Remission wird anhand des Mayo Gesamtscores beziehungsweise anhand des Adaptierten Mayo Scores beschrieben. Der Mayo Gesamtscore setzt sich aus den Subscores der Stuhlfrequenz, der rektalen Blutung, der Beurteilung durch den Prüfarzt und der Endoskopie zusammen und hat einen Wertebereich von null bis zwölf. Der Adaptierte Mayo Score beinhaltet die gleichen Subscores mit Außnahme der Prüfarztbeurteilung und hat einen Wertebereich von null bis neun. Sowohl für den Mayo Gesamtscore als auch den Adaptierten Mayo Score bilden höhere Werte einen schwereren Krankheitsverlauf ab. Die klinische Remission per Mayo Gesamtscore ist definiert als ein Mayo Gesamtscore ≤ 2 und keinem Subscore > 1. Die klinische Remission per Adaptiertem Mayo Score ist definiert als ein Adaptierter Mayo Score ≤ 2, mit einer Stuhlfrequenz ≤ 1 und die nicht höher als ihr Ausgangswert ist, einer rektalen Blutung von null und einer Endoskopie von ≤ 1. Der Mayo Gesamtscore und der Adaptierte Mayo Score wurden zu Baseline und zu Woche 52 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer klinischen Remission per Mayo Gesamtscore beziehungsweise per Adaptiertem Mayo Score dargestellt. Der Anteil der Patienten beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden mittels NRI-C ersetzt.
UNIFI Erhaltungsstudie	Die klinische Remission wird anhand des Mayo Gesamtscores beziehungsweise anhand des Adaptierten Mayo Scores beschrieben. Der Mayo Gesamtscore setzt sich aus den Subscores der Stuhlfrequenz, der rektalen Blutung, der Beurteilung durch den Prüfarzt und der Endoskopie zusammen und hat einen Wertebereich von null bis zwölf. Der Adaptierte Mayo Score beinhaltet die gleichen Subscores mit Außnahme der Prüfarztbeurteilung und hat einen Wertebereich von null bis neun. Sowohl für den Mayo Gesamtscore als auch den Adaptierten Mayo Score bilden höhere Werte einen schwereren Krankheitsverlauf ab. Die klinische Remission per Mayo Gesamtscore ist definiert als ein Mayo Gesamtscore ≤ 2 und keinem Subscore > 1. Die klinische Remission per Adaptiertem Mayo Score ist definiert als eine absolute Stuhlhäufigkeit ≤ 3, einer rektalen Blutung von null und einer Endoskopie von ≤ 1. Der Mayo Gesamtscore und der Adaptierte Mayo Score wurden zu Baseline und zu Woche 44 erhoben.

Studie	Operationalisierung
	Für das vorliegende Dossier wird der Anteil an Patienten mit einer klinischen Remission per Mayo Gesamtscore beziehungsweise per Adaptiertem Mayo Score dargestellt. Die Auswertung zur klinischen Remission beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; LOCF: Last Observation Carried Forward; MI: Multiple Imputation; NRI: Non-Responder-Imputation; NRI-C: NRI incorporating MI to handle missing data due to COVID-19; SAP: Statistischer Analyseplan	

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 4-60: Bewertung des Verzerrungspotenzials für klinische Remission in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung der klinischen Remission erfolgte zu Woche 8 und zu Woche 44 bzw. 52. Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore)

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

Tabelle 4-61: Ergebnisse für Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore) – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	11/56 (19,6)	0/46 (0,0)	18,96 [1,15; 313,42] 0,0398	23,51 [1,35; 410,76] 0,0305	0,20 [0,09; 0,30] 0,0004
U-ACHIEVE Substudie 2 [45mg]	73/319 (22,9)	4/154 (2,6)	8,81 [3,28; 23,66] <0,0001	11,13 [3,99; 31,07] <0,0001	0,20 [0,15; 0,26] <0,0001
U-ACCOMPLISH [45mg]	93/341 (27,3)	4/174 (2,3)	11,86 [4,43; 31,74] <0,0001	15,94 [5,75; 44,18] <0,0001	0,25 [0,20; 0,30] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	177/716 (24,7)	8/374 (2,1)	10,74 [5,47; 21,09] <0,0001	13,96 [6,93; 28,11] <0,0001	0,22 [0,19; 0,26] <0,0001
Heterogenitätstest (p-Wert)			0,8388		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	60/148 (40,5)	16/149 (10,7)	3,78 [2,28; 6,24] <0,0001	5,67 [3,07; 10,47] <0,0001	0,30 [0,20; 0,39] <0,0001

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
U-ACHIEVE Substudie 3 [30mg]	73/154 (47,4)	16/149 (10,7)	4,41 [2,70; 7,22] <0,0001	7,49 [4,08; 13,75] <0,0001	0,37 [0,27; 0,46] <0,0001
Analysepopulation: ITT Imputation: NRI (U-ACHIEVE Substudie 1) und NRI-C (U-ACHIEVE Substudie 2, U-ACHIEVE Substudie 3, U-ACCOMPLISH) RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen. Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Upadacitinib dar. ITT: Intention-to-Treat; KI: Konfidenzintervall; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; NRI-C: Non-Responder-Imputation incorporating multiple imputation to handle missing data due to COVID-19; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko					

Hinsichtlich der klinischen Remission (Mayo Gesamtscore) zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts klinische Remission (Mayo Gesamtscore) für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

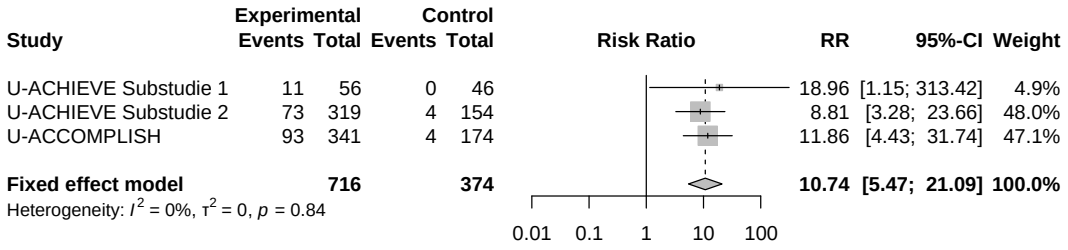

Abbildung 9: Forest-Plot für klinische Remission (Mayo Gesamtscore) zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-62: Ergebnisse für Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore) – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
UNIFI [6mg/kg]	50/322 (15,5)	17/319 (5,3)	2,91 [1,72; 4,94] <0,0001	3,27 [1,84; 5,80] <0,0001	0,10 [0,06; 0,15] <0,0001
Erhaltungsphase					
UNIFI [90mg/12w]	66/172 (38,4)	42/175 (24,0)	1,60 [1,16; 2,21] 0,0046	1,97 [1,24; 3,13] 0,0041	0,14 [0,05; 0,24] 0,0035
UNIFI [90mg/8w]	77/176 (43,8)	42/175 (24,0)	1,82 [1,33; 2,49] 0,0002	2,46 [1,56; 3,89] 0,0001	0,20 [0,10; 0,29] <0,0001

Analysepopulation: ITT

Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5.

Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Ustekinumab dar.

ITT: Intention-to-Treat; KI: Konfidenzintervall; LOCF: Last observation carried forward; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko

Hinsichtlich der klinischen Remission (Mayo Gesamtscore) zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-63: Ergebnisse für Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore)	10,74 [5,47; 21,09] <0,0001	2,91 [1,72; 4,94] <0,0001	3,69 [1,57; 8,68] 0,0028

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko

Bezüglich der klinischen Remission mittels Mayo Gesamtscore zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (RR [95 %-KI]: 3,69 [1,57; 8,68]; p-Wert: = 0,0028). Damit ist die Wahrscheinlichkeit für Patienten, eine klinische Remission zu erreichen, unter einer Behandlung mit Upadacitinib ungefähr 3,7-mal höher als unter einer Behandlung mit Ustekinumab.

Tabelle 4-64: Ergebnisse für Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo RR [95%-KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit klinischer Remission (Mayo Gesamtscore)	3,78 [2,28; 6,24] <0,0001	1,60 [1,16; 2,21] 0,0046	2,36 [1,30; 4,29] 0,0049	4,41 [2,70; 7,22] <0,0001	1,82 [1,33; 2,49] 0,0002	2,42 [1,35; 4,34] 0,0029

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko

Bezüglich der klinischen Remission mittels Mayo Gesamtscore zeigt sich für die Erhaltungsphase ein statistisch signifikanter Vorteil sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W (RR [95 %-KI]: 2,36 [1,30; 4,29]; p-Wert: = 0,0049), als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W (RR [95 %-KI]: 2,42 [1,35; 4,34]; p-Wert: = 0,0029). Damit ist die Wahrscheinlichkeit für eine klinische Remission unter einer Behandlung mit Upadacitinib ungefähr 2,4-mal höher als unter einer Behandlung mit Ustekinumab.

Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score)

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

Tabelle 4-65: Ergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	12/56 (21,4)	0/46 (0,0)	20,61 [1,25; 339,06] 0,0342	26,12 [1,50; 454,53] 0,0252	0,21 [0,10; 0,33] 0,0002
U-ACHIEVE Substudie 2 [45mg]	83/319 (26,0)	7/154 (4,5)	5,72 [2,71; 12,08] <0,0001	7,39 [3,32; 16,41] <0,0001	0,21 [0,16; 0,27] <0,0001
U-ACCOMPLISH [45mg]	114/341 (33,4)	7/174 (4,0)	8,31 [3,96; 17,44] <0,0001	11,98 [5,44; 26,37] <0,0001	0,29 [0,24; 0,35] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	209/716 (29,2)	14/374 (3,7)	7,39 [4,42; 12,36] <0,0001	10,06 [5,82; 17,41] <0,0001	0,25 [0,21; 0,29] <0,0001
Heterogenitätstest (p-Wert)			0,5882		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	63/148 (42,6)	18/149 (12,1)	3,52 [2,20; 5,65] <0,0001	5,39 [2,99; 9,74] <0,0001	0,30 [0,21; 0,40] <0,0001
U-ACHIEVE Substudie 3 [30mg]	80/154 (51,9)	18/149 (12,1)	4,30 [2,72; 6,81] <0,0001	7,87 [4,38; 14,13] <0,0001	0,40 [0,30; 0,49] <0,0001

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Analysepopulation: ITT Imputation: NRI (U-ACHIEVE Substudie 1) und NRI-C (U-ACHIEVE Substudie 2, U-ACHIEVE Substudie 3, U-ACCOMPLISH) RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen. Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Upadacitinib dar. ITT: Intention-to-Treat; KI: Konfidenzintervall; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; NRI-C: Non-Responder-Imputation incorporating multiple imputation to handle missing data due to COVID-19; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko					

Hinsichtlich der klinischen Remission (Adaptierter Mayo Score) zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts klinische Remission (Adaptierter Mayo Score) für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

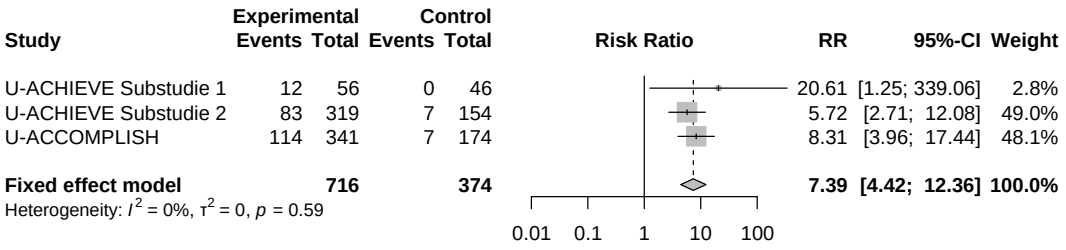

Abbildung 10: Forest-Plot für klinische Remission (Adaptierter Mayo Score) zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-66: Ergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
UNIFI [6mg/kg]	61/322 (18,9)	20/319 (6,3)	3,02 [1,87; 4,89] <0,0001	3,49 [2,05; 5,95] <0,0001	0,13 [0,08; 0,18] <0,0001
Erhaltungsphase					
UNIFI [90mg/12w]	68/172 (39,5)	43/175 (24,6)	1,61 [1,17; 2,21] 0,0034	2,01 [1,27; 3,18] 0,0030	0,15 [0,05; 0,25] 0,0025
UNIFI [90mg/8w]	75/176 (42,6)	43/175 (24,6)	1,73 [1,27; 2,37] 0,0005	2,28 [1,45; 3,60] 0,0004	0,18 [0,08; 0,28] 0,0003

Analysepopulation: ITT

Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5.

Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Ustekinumab dar.

ITT: Intention-to-Treat; KI: Konfidenzintervall; LOCF: Last observation carried forward; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko

Hinsichtlich der klinischen Remission (Adaptierter Mayo Score) zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-67: Ergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score)	7,39 [4,42; 12,36] <0,0001	3,02 [1,87; 4,89] <0,0001	2,45 [1,21; 4,95] 0,0127

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko

Bezüglich der klinischen Remission mittels Adaptiertem Mayo Score zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (RR [95 %-KI]: 2,45 [1,21; 4,95]; p-Wert: = 0,0127). Damit ist die Wahrscheinlichkeit für eine klinische Remission mittels Adaptiertem Mayo Score unter einer Behandlung mit Upadacitinib ungefähr 2,5-mal höher als unter einer Behandlung mit Ustekinumab.

Tabelle 4-68: Ergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo RR [95%-KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score)	3,52 [2,20; 5,65] <0,0001	1,61 [1,17; 2,21] 0,0034	2,19 [1,24; 3,87] 0,0070	4,30 [2,72; 6,81] <0,0001	1,73 [1,27; 2,37] 0,0005	2,48 [1,42; 4,32] 0,0013

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko

Bezüglich der klinischen Remission mittels Adaptiertem Mayo Score zeigt sich für die Erhaltungsphase ein statistisch signifikanter Vorteil sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W (RR [95 %-KI]: 2,19 [1,24; 3,87]; p-Wert: = 0,0070), als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W (RR [95 %-KI]: 2,48 [1,42; 4,32]; p-Wert: = 0,0013). Damit ist die Wahrscheinlichkeit für eine klinische Remission mittels Adaptiertem Mayo Score unter einer Behandlung mit Upadacitinib ungefähr 2,2-mal höher, beziehungsweise ungefähr 2,5-mal höher als unter einer Behandlung mit Ustekinumab.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.3.3 Klinisches Ansprechen – indirekte Vergleiche aus RCT

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

Tabelle 4-69: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	•	•	
1	UNIFI Induktionsstudie		•	•
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	•	•	
1	UNIFI Erhaltungsstudie		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

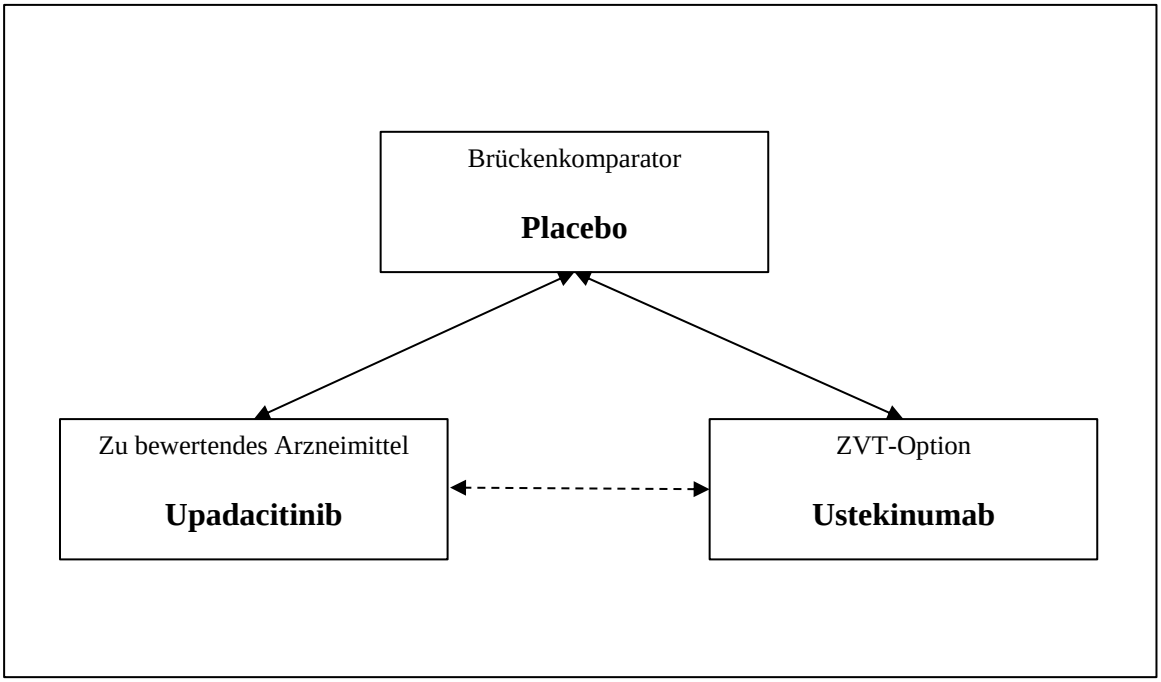

Abbildung 11: Netzwerkstruktur des indirekten Vergleichs

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

Tabelle 4-70: Operationalisierung von klinisches Ansprechen

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Das klinische Ansprechen wird anhand des Mayo Gesamtscores beschrieben. Der Mayo Gesamtscore hat einen Wertebereich von null bis zwölf wobei höhere Werte einen schwereren Krankheitsverlauf abbilden. Das klinische Ansprechen per Mayo Gesamtscore ist definiert als eine Verbesserung des Mayo Gesamtscores ≥ 3 und einer Verbesserung von ≥ 30 % gegenüber dem Ausgangswert und zusätzlich einer Verringerung der rektalen Blutung um ≥ 1 oder einer absoluten rektalen Blutung von ≤ 1. Der Mayo Gesamtscore wurde zu Baseline und zu Woche 8 erhoben und das klinische Ansprechen zu Woche 8 ermittelt. Für das vorliegende Dossier wird der Anteil an Patienten mit einem klinischen Ansprechen per Mayo Gesamtscore dargestellt. Die Auswertung zum klinischen Ansprechen beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden mittels NRI- beziehungsweise NRI-C-Analyse ersetzt. Für die Beschreibung der Induktionsphase, wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH meta-analytisch zusammengefasst.
UNIFI Induktionsstudie	Das klinische Ansprechen wird anhand des Mayo Gesamtscores beschrieben. Der Mayo Gesamtscore hat einen Wertebereich von null bis zwölf, wobei höhere Werte einen schwereren Krankheitsverlauf abbilden. Das klinische Ansprechen per Mayo Gesamtscore ist definiert als eine Verbesserung des Mayo Gesamtscores ≥ 3 und einer

Studie	Operationalisierung
Induktionsstudien	
	Verbesserung von ≥ 30 % gegenüber dem Ausgangswert und zusätzlich einer Verringerung der rektalen Blutung um ≥ 1 oder einer absoluten rektalen Blutung von ≤ 1. Der Mayo Gesamtscore wurde zu Baseline und zu Woche 8 erhoben und das klinische Ansprechen zu Woche 8 ermittelt. Für das vorliegende Dossier wird der Anteil an Patienten mit einem klinischen Ansprechen per Mayo Gesamtscore dargestellt. Die Auswertung zum klinischen Ansprechen beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Das klinische Ansprechen wird anhand des Mayo Gesamtscores beschrieben. Der Mayo Gesamtscore hat einen Wertebereich von null bis zwölf, wobei höhere Werte einen schwereren Krankheitsverlauf abbilden. Das klinische Ansprechen per Mayo Gesamtscore ist definiert als eine Verbesserung des Mayo Gesamtscores ≥ 3 und einer Verbesserung von ≥ 30 % gegenüber dem Ausgangswert und zusätzlich einer Verringerung der rektalen Blutung um ≥ 1 oder einer absoluten rektalen Blutung von ≤ 1. Der Mayo Gesamtscore wurde zu Baseline und zu Woche 52 erhoben und das klinische Ansprechen zu Woche 52 ermittelt. Für das vorliegende Dossier wird der Anteil an Patienten mit einem klinischen Ansprechen per Mayo Gesamtscore dargestellt. Die Auswertung zum klinischen Ansprechen beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden per NRI-C-Analyse ersetzt.
UNIFI Erhaltungsstudie	Das klinische Ansprechen wird anhand des Mayo Gesamtscores beschrieben. Der Mayo Gesamtscore hat einen Wertebereich von null bis zwölf, wobei höhere Werte einen schwereren Krankheitsverlauf abbilden. Das klinische Ansprechen per Mayo Gesamtscore ist definiert als eine Verbesserung des Mayo Gesamtscores ≥ 3 und einer Verbesserung von ≥ 30 % gegenüber dem Ausgangswert und zusätzlich einer Verringerung der rektalen Blutung um ≥ 1 oder einer absoluten rektalen Blutung von ≤ 1. Der Mayo Gesamtscore wurde zu Baseline und zu Woche 44 erhoben und das klinische Ansprechen zu Woche 44 ermittelt. Für das vorliegende Dossier wird der Anteil an Patienten mit einem klinischen Ansprechen per Mayo Gesamtscore dargestellt. Die Auswertung zum klinischen Ansprechen beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; LOCF: Last Observation Carried Forward; MI: Multiple Imputation; NRI: Non-Responder-Imputation; NRI-C: NRI incorporating MI to handle missing data due to COVID-19; SAP: Statistischer Analyseplan	

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 4-71: Bewertung des Verzerrungspotenzials für klinisches Ansprechen in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des klinischen Ansprechens erfolgte zu Woche 8 und zu Woche 44 bzw. 52. Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

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

Tabelle 4-72: Ergebnisse für Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore) – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	32/56 (57,1)	8/46 (17,4)	3,29 [1,68; 6,42] 0,0005	6,33 [2,50; 16,02] <0,0001	0,40 [0,23; 0,57] <0,0001
U-ACHIEVE Substudie 2 [45mg]	235/319 (73,7)	37/154 (24,0)	3,07 [2,30; 4,09] <0,0001	8,85 [5,66; 13,82] <0,0001	0,50 [0,41; 0,58] <0,0001
U-ACCOMPLISH [45mg]	254/341 (74,5)	43/174 (24,7)	3,01 [2,31; 3,94] <0,0001	8,89 [5,83; 13,56] <0,0001	0,50 [0,42; 0,58] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	521/716 (72,8)	88/374 (23,5)	3,06 [2,53; 3,69] <0,0001	8,57 [6,40; 11,46] <0,0001	0,49 [0,43; 0,54] <0,0001
Heterogenitätstest (p-Wert)			0,9725		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	92/148 (62,2)	33/149 (22,1)	2,81 [2,03; 3,89] <0,0001	5,77 [3,47; 9,61] <0,0001	0,40 [0,30; 0,50] <0,0001
U-ACHIEVE Substudie 3 [30mg]	116/154 (75,3)	33/149 (22,1)	3,40 [2,48; 4,66] <0,0001	10,73 [6,30; 18,28] <0,0001	0,53 [0,44; 0,63] <0,0001

Analysepopulation: ITT

Imputation: NRI (U-ACHIEVE Substudie 1) und NRI-C (U-ACHIEVE Substudie 2, U-ACHIEVE Substudie 3, U-ACCOMPLISH)

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen.

Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Upadacitinib dar.

ITT: Intention-to-Treat; KI: Konfidenzintervall; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; NRI-C: Non-Responder-Imputation incorporating multiple imputation to handle missing data due to COVID-19; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko

Hinsichtlich des klinischen Ansprechens (Mayo Gesamtscore) zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts klinisches Ansprechen (Mayo Gesamtscore) für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

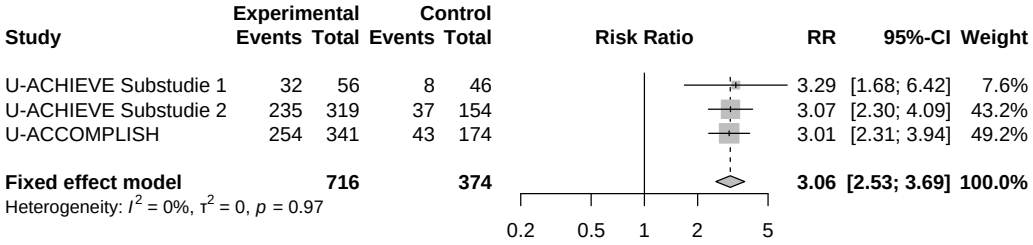

Abbildung 12: Forest-Plot für klinisches Ansprechen (Mayo Gesamtscore) zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-73: Ergebnisse für Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore) – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
UNIFI [6mg/kg]	199/322 (61,8)	100/319 (31,3)	1,97 [1,64; 2,37] <0,0001	3,54 [2,56; 4,91] <0,0001	0,30 [0,23; 0,38] <0,0001
Erhaltungsphase					
UNIFI [90mg/12w]	117/172 (68,0)	78/175 (44,6)	1,53 [1,26; 1,85] <0,0001	2,65 [1,71; 4,10] <0,0001	0,23 [0,13; 0,34] <0,0001
UNIFI [90mg/8w]	125/176 (71,0)	78/175 (44,6)	1,59 [1,32; 1,93] <0,0001	3,05 [1,96; 4,74] <0,0001	0,26 [0,16; 0,36] <0,0001

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Analysepopulation: ITT Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Ustekinumab dar. ITT: Intention-to-Treat; KI: Konfidenzintervall; LOCF: Last observation carried forward; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko					

Hinsichtlich des klinischen Ansprechens (Mayo Gesamtscore) zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-74: Ergebnisse für Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore)	3,06 [2,53; 3,69] <0,0001	1,97 [1,64; 2,37] <0,0001	1,55 [1,19; 2,02] 0,0011

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko

Bezüglich des klinischen Ansprechens mittels Mayo Gesamtscore zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (RR [95 %-KI]: 1,55 [1,19; 2,02]; p-Wert: = 0,0011). Damit ist die Wahrscheinlichkeit für ein klinisches Ansprechen mittels Mayo Gesamtscore unter einer Behandlung mit Upadacitinib ungefähr 1,5-mal höher als unter einer Behandlung mit Ustekinumab.

Tabelle 4-75: Ergebnisse für Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore) – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo RR [95%-KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit klinischem Ansprechen (Mayo Gesamtscore)	2,81 [2,03; 3,89] <0,0001	1,53 [1,26; 1,85] <0,0001	1,84 [1,26; 2,69] 0,0017	3,40 [2,48; 4,66] <0,0001	1,59 [1,32; 1,93] <0,0001	2,13 [1,48; 3,08] <0,0001

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko

Bezüglich des klinischen Ansprechens mittels Mayo Gesamtscore zeigt sich für die Erhaltungsphase ein statistisch signifikanter Vorteil sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W (RR [95 %-KI]: 1,84 [1,26; 2,69]; p-Wert: = 0,0017), als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W (RR [95 %-KI]: 2,13 [1,48; 3,08]; p-Wert: < 0,0001). Damit ist die Wahrscheinlichkeit für ein klinisches Ansprechen mittels Mayo Gesamtscore unter einer Behandlung mit Upadacitinib ungefähr 1,8-mal höher, beziehungsweise ungefähr 2,1-mal höher als unter einer Behandlung mit Ustekinumab.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.3.4 Mukosaheilung – indirekte Vergleiche aus RCT

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

Tabelle 4-76: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
2	U-ACHIEVE Substudie 2 U-ACCOMPLISH	•	•	
1	UNIFI Induktionsstudie		•	•
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	•	•	
1	UNIFI Erhaltungsstudie		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

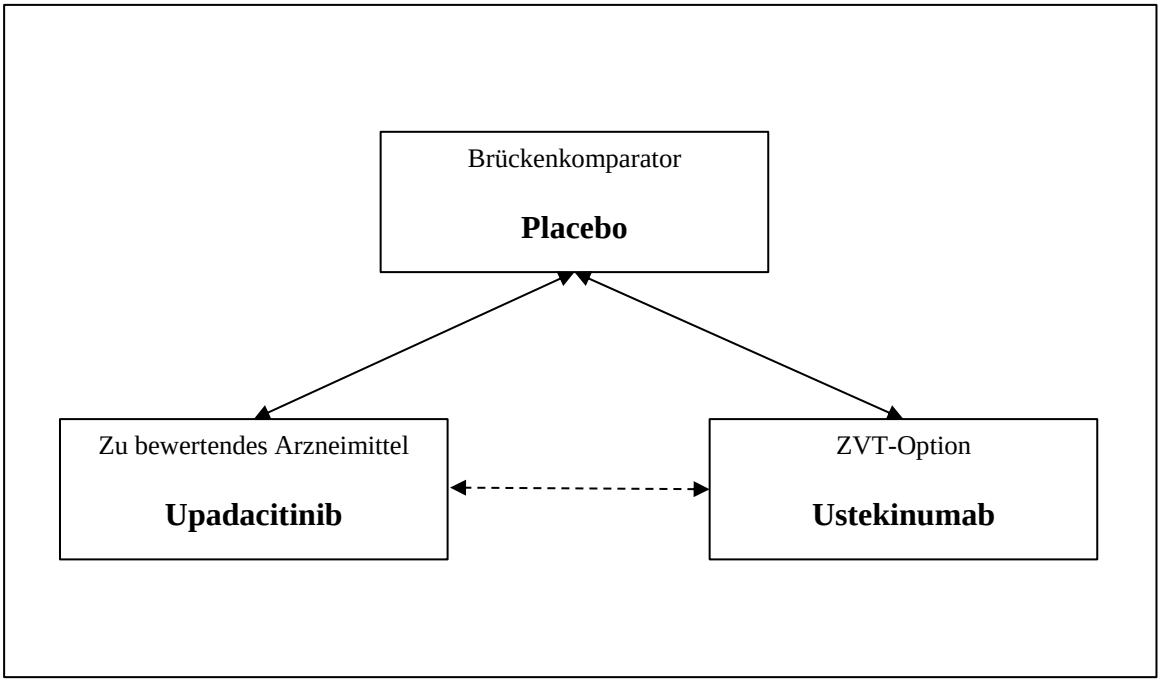

Abbildung 13: Netzwerkstruktur des indirekten Vergleichs

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

Tabelle 4-77: Operationalisierung von Mukosaheilung

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Für die vorliegende Studie wird eine histologische, endoskopische Verbesserung der Mukosa als Mukosaheilung definiert. Ausschlaggebend ist dafür ein endoskopischer Subscore des Full Mayo Gesamtscores von null oder eins und ein Geboes Score ≤ 3,1. Die Mukosaheilung wurde zu Baseline und zu Woche 8 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer Mukosaheilung dargestellt. Der Anteil der Patienten beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden mittels NRI- beziehungsweise NRI-C-Analyse ersetzt. Für die Beschreibung der Induktionsphase, wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH meta-analytisch zusammengefasst.
UNIFI Induktionsstudie	Eine Mukosaheilung ist definiert durch sowohl eine histologische Verbesserung in Form einer neutrophilen Infiltration in < 5 % der Krypten, keine Zerstörung der Krypten und keine Erosionen, Ulzerationen oder Granulationsgewebe als auch einer endoskopischen Verbesserung in Form eines Endoskopie Subscore des Mayo Gesamtscores von ≤ 1. Die Mukosaheilung wurde zu Baseline und zu Woche 8 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit Mukosaheilung dargestellt. Die Auswertung zur Mukosaheilung beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests.

Studie	Operationalisierung
Induktionsstudien	
	Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Für die vorliegende Studie wird eine histologische, endoskopische Verbesserung der Mukosa als Mukosaheilung definiert. Ausschlaggebend ist dafür ein endoskopischer Subscore des Mayo Gesamtscores von null oder eins und ein Geboes Score ≤ 3,1. Die Mukosaheilung wurde zu Baseline und zu Woche 52 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer Mukosaheilung dargestellt. Der Anteil der Patienten beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden per NRI-C-Analyse ersetzt.
UNIFI Erhaltungsstudie	Eine Mukosaheilung ist definiert durch sowohl eine histologische Verbesserung in Form einer neutrophilen Infiltration in < 5 % der Krypten, keine Zerstörung der Krypte und keine Erosionen, Ulzerationen oder Granulationsgewebe als auch einer endoskopischen Verbesserung in Form eines Endoskopie Subscore des Mayo Gesamtscores von ≤ 1. Die Mukosaheilung wurde zu Baseline und zu Woche 44 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit Mukosaheilung dargestellt. Die Auswertung zur Mukosaheilung beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; LOCF: Last Observation Carried Forward; MI: Multiple Imputation; NRI: Non-Responder-Imputation; NRI-C: NRI incorporating MI to handle missing data due to COVID-19; SAP: Statistischer Analyseplan	

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 4-78: Bewertung des Verzerrungspotenzials für Mukosaheilung in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung der Mukosaheilung erfolgte zu Woche 8 und zu Woche 44 bzw. 52. Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

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

Tabelle 4-79: Ergebnisse für Anteil der Patienten mit Mukosaheilung – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
U-ACHIEVE Substudie 2 [45mg]	96/319 (30,1)	10/154 (6,5)	4,63 [2,49; 8,63] <0,0001	6,20 [3,13; 12,29] <0,0001	0,24 [0,17; 0,30] <0,0001
U-ACCOMPLISH [45mg]	125/341 (36,7)	10/174 (5,7)	6,38 [3,44; 11,83] <0,0001	9,49 [4,83; 18,65] <0,0001	0,31 [0,25; 0,37] <0,0001
Meta-Analyse U-ACHIEVE Substudie 2 + U-ACCOMPLISH	221/660 (33,5)	20/328 (6,1)	5,50 [3,55; 8,52] <0,0001	7,75 [4,80; 12,52] <0,0001	0,27 [0,23; 0,32] <0,0001
Heterogenitätstest (p-Wert)			0,4744		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	52/148 (35,1)	18/149 (12,1)	2,91 [1,79; 4,73] <0,0001	3,94 [2,17; 7,16] <0,0001	0,23 [0,14; 0,32] <0,0001
U-ACHIEVE Substudie 3 [30mg]	77/154 (50,0)	18/149 (12,1)	4,14 [2,61; 6,56] <0,0001	7,28 [4,05; 13,07] <0,0001	0,38 [0,28; 0,47] <0,0001

Analysepopulation: ITT

Imputation: NRI-C (U-ACHIEVE Substudie 2, U-ACHIEVE Substudie 3, U-ACCOMPLISH)

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen.

Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Upadacitinib dar.

ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; NRI-C: Non-Responder-Imputation incorporating multiple imputation to handle missing data due to COVID-19; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko

Hinsichtlich der Mukosaheilung zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts Mukosaheilung für die Studien U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

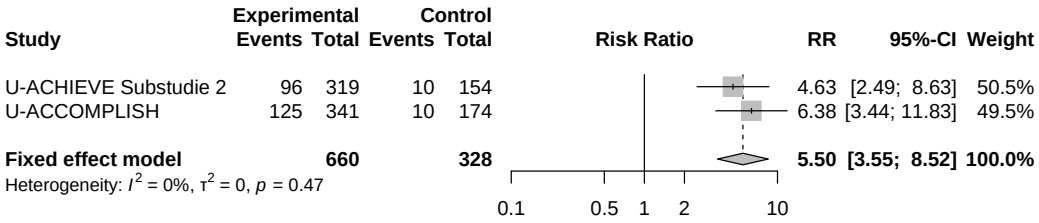

Abbildung 14: Forest-Plot für Mukosaheilung zu Woche 8 aus RCT (U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-80: Ergebnisse für Anteil der Patienten mit Mukosaheilung – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
UNIFI [6mg/kg]	58/315 (18,4)	28/316 (8,9)	2,08 [1,36; 3,17] 0,0007	2,32 [1,43; 3,76] 0,0006	0,10 [0,04; 0,15] 0,0004
Erhaltungsphase					
UNIFI [90mg/12w]	66/170 (38,8)	41/170 (24,1)	1,61 [1,16; 2,23] 0,0043	2,00 [1,25; 3,19] 0,0037	0,15 [0,05; 0,24] 0,0031
UNIFI [90mg/8w]	79/172 (45,9)	41/170 (24,1)	1,90 [1,39; 2,60] <0,0001	2,67 [1,68; 4,24] <0,0001	0,22 [0,12; 0,32] <0,0001

Analysepopulation: ITT

Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5.

Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Ustekinumab dar.

ITT: Intention-to-Treat; KI: Konfidenzintervall; LOCF: Last observation carried forward; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko

Hinsichtlich der Mukosaheilung zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-81: Ergebnisse für Anteil der Patienten mit Mukosaheilung – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit Mukosaheilung	5,50 [3,55; 8,52] <0,0001	2,08 [1,36; 3,17] 0,0007	2,65 [1,44; 4,86] 0,0017
Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo. Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar. KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko			

Bezüglich der Mukosaheilung zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (RR [95 %-KI]: 2,65 [1,44; 4,86]; p-Wert: = 0,0017). Damit ist die Wahrscheinlichkeit für eine Mukosaheilung unter einer Behandlung mit Upadacitinib ungefähr 2,7-mal höher als unter einer Behandlung mit Ustekinumab.

Tabelle 4-82: Ergebnisse für Anteil der Patienten mit Mukosaheilung – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo RR [95%-KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit Mukosaheilung	2,91 [1,79; 4,73] <0,0001	1,61 [1,16; 2,23] 0,0043	1,81 [1,01; 3,24] 0,0475	4,14 [2,61; 6,56] <0,0001	1,90 [1,39; 2,60] <0,0001	2,17 [1,25; 3,79] 0,0063

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko

Bezüglich der Mukosaheilung zeigt sich für die Erhaltungsphase ein statistisch signifikanter Vorteil sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W (RR [95 %-KI]: 1,81 [1,01; 3,24]; p-Wert: = 0,0475), als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W (RR [95 %-KI]: 2,17 [1,25; 3,79]; p-Wert: = 0,0063). Damit ist die Wahrscheinlichkeit für eine Mukosaheilung unter einer Behandlung mit Upadacitinib ungefähr 1,8-mal höher, beziehungsweise ungefähr 2,2-mal höher als unter einer Behandlung mit Ustekinumab.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.3.5 Stuhlfrequenz – indirekte Vergleiche aus RCT

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

Tabelle 4-83: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	●	●	
1	UNIFI Induktionsstudie		●	●
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	●	●	
1	UNIFI Erhaltungsstudie		●	●

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

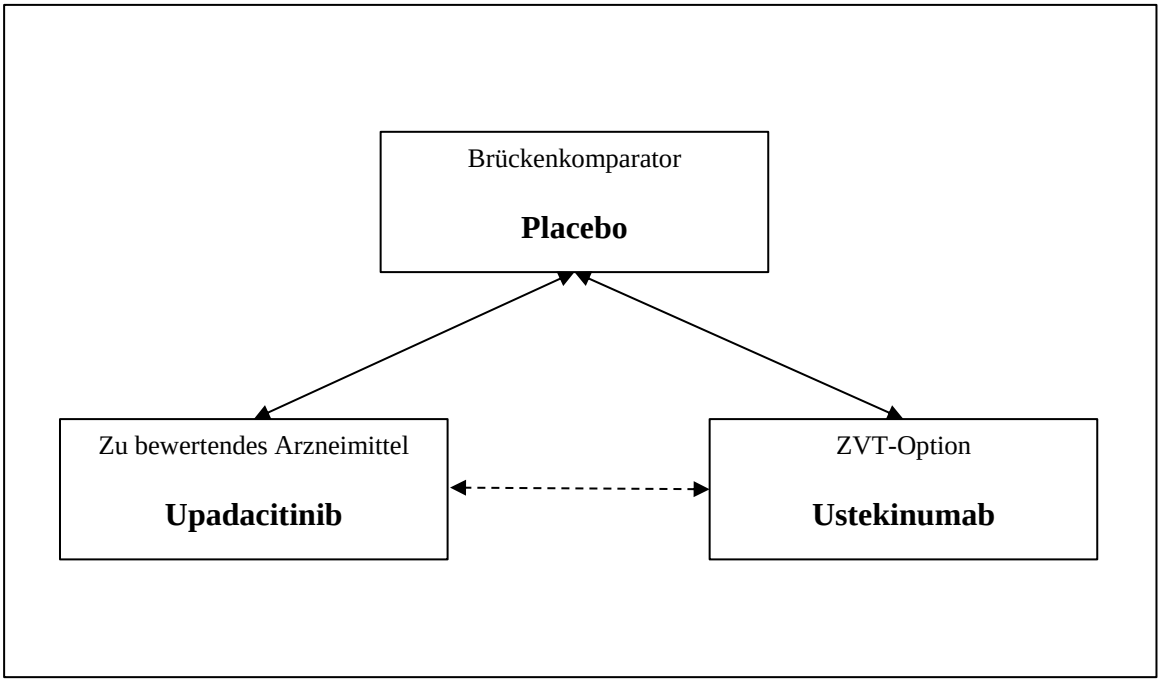

Abbildung 15: Netzwerkstruktur des indirekten Vergleichs

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

Tabelle 4-84: Operationalisierung von Stuhlfrequenz

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Der Schweregrad der Stuhlfrequenz wird durch den Subscore Stuhlfrequenz des Mayo Gesamtscores auf einer Skala von null („normale Stuhlfrequenz“) bis drei („5 Stuhlgänge oder mehr als normal/Tag“) Punkten kategorisiert. Die Stuhlfrequenz wurde zu Baseline, Woche 2, Woche 4, Woche 6 und zu Woche 8 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer Stuhlfrequenz ≤ 1 dargestellt. Der Anteil der Patienten beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden mittels NRI- beziehungsweise NRI-C-Analyse ersetzt. Für die Beschreibung der Induktionsphase wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH meta-analytisch zusammengefasst.
UNIFI Induktionsstudie	Der Schweregrad der Stuhlfrequenz wird durch den Subscore Stuhlfrequenz des Mayo Gesamtscores auf einer Skala von null („normale Stuhlfrequenz“) bis drei („5 Stuhlgänge oder mehr als normal/Tag“) Punkten kategorisiert. Die Stuhlfrequenz wurde zu Baseline, Woche 2, Woche 4 und zu Woche 8 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer Stuhlfrequenz ≤ 1 dargestellt. Die Auswertung beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests.

Studie	Operationalisierung
Induktionsstudien	
	Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Der Schweregrad der Stuhlfrequenz wird durch den Subscore Stuhlfrequenz des Mayo Gesamtscores auf einer Skala von null („normale Stuhlfrequenz“) bis drei („5 Stuhlgänge oder mehr als normal/Tag“) Punkten kategorisiert. Die Stuhlfrequenz wurde zu Baseline, Woche 12, Woche 20, Woche 28, Woche 36, Woche 44 und zu Woche 52 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer Stuhlfrequenz ≤ 1 zu Woche 44 dargestellt. Der Anteil der Patienten beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden per NRI-C-Analyse ersetzt.
UNIFI Erhaltungsstudie	Der Schweregrad der Stuhlfrequenz wird durch den Subscore Stuhlfrequenz des Mayo Gesamtscores auf einer Skala von null („normale Stuhlfrequenz“) bis drei („5 Stuhlgänge oder mehr als normal/Tag“) Punkten kategorisiert. Die Stuhlfrequenz wurde zu Baseline, Woche 2, Woche 4, Woche 8, Woche 12, Woche 16, Woche 20, Woche 24, Woche 28, Woche 32, Woche 36, Woche 40 und zu Woche 44 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer Stuhlfrequenz ≤ 1 dargestellt. Die Auswertung beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; LOCF: Last Observation Carried Forward; MI: Multiple Imputation; NRI: Non-Responder-Imputation; NRI-C: NRI incorporating MI to handle missing data due to COVID-19; SAP: Statistischer Analyseplan	

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 4-85: Bewertung des Verzerrungspotenzials für Stuhlfrequenz in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung der Stuhlfrequenz erfolgte zu Woche 8 und zu Woche 44. Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

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

Tabelle 4-86: Ergebnisse für Anteil der Patienten mit Stuhlfrequenz ≤ 1 – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	27/56 (48,2)	6/46 (13,0)	3,70 [1,67; 8,18] 0,0012	6,21 [2,27; 16,97] 0,0004	0,35 [0,19; 0,51] <0,0001
U-ACHIEVE Substudie 2 [45mg]	191/319 (59,9)	40/154 (26,0)	2,31 [1,74; 3,05] <0,0001	4,25 [2,78; 6,50] <0,0001	0,34 [0,25; 0,43] <0,0001
U-ACCOMPLISH [45mg]	211/341 (61,9)	34/174 (19,5)	3,17 [2,32; 4,33] <0,0001	6,68 [4,33; 10,31] <0,0001	0,42 [0,35; 0,50] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	429/716 (59,9)	80/374 (21,4)	2,76 [2,25; 3,38] <0,0001	5,40 [4,04; 7,21] <0,0001	0,38 [0,32; 0,44] <0,0001
Heterogenitätstest (p-Wert)			0,2424		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	86/148 (58,1)	33/149 (22,1)	2,62 [1,88; 3,65] <0,0001	4,88 [2,94; 8,09] <0,0001	0,36 [0,26; 0,46] <0,0001
U-ACHIEVE Substudie 3 [30mg]	106/154 (68,8)	33/149 (22,1)	3,11 [2,26; 4,28] <0,0001	7,76 [4,64; 13,00] <0,0001	0,47 [0,37; 0,57] <0,0001

Analysepopulation: ITT

Imputation: NRI (U-ACHIEVE Substudie 1) und NRI-C (U-ACHIEVE Substudie 2, U-ACHIEVE Substudie 3, U-ACCOMPLISH)

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen.

Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Upadacitinib dar.

ITT: Intention-to-Treat; KI: Konfidenzintervall; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; NRI-C: Non-Responder-Imputation incorporating multiple imputation to handle missing data due to COVID-19; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko

Hinsichtlich der Stuhlfrequenz ≤ 1 zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts Stuhlfrequenz ≤ 1 für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

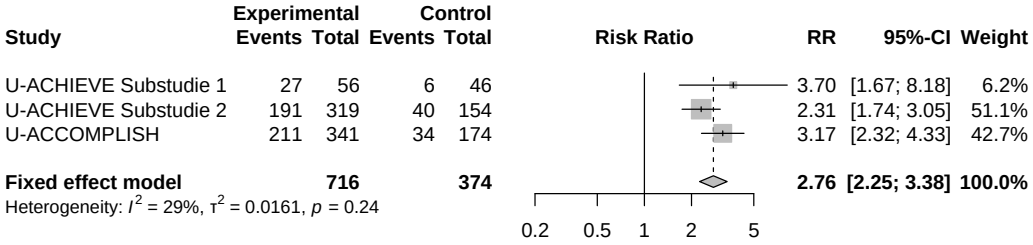

Abbildung 16: Forest-Plot für Stuhlfrequenz ≤ 1 zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-87: Ergebnisse für Anteil der Patienten mit Stuhlfrequenz ≤ 1 – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
UNIFI [6mg/kg]	181/321 (56,4)	121/319 (37,9)	1,49 [1,25; 1,76] <0,0001	2,12 [1,54; 2,90] <0,0001	0,18 [0,11; 0,26] <0,0001
Erhaltungsphase					
UNIFI [90mg/12w]	119/172 (69,2)	99/175 (56,6)	1,22 [1,04; 1,44] 0,0160	1,72 [1,11; 2,68] 0,0154	0,13 [0,03; 0,23] 0,0141
UNIFI [90mg/8w]	143/176 (81,3)	99/175 (56,6)	1,44 [1,24; 1,67] <0,0001	3,33 [2,05; 5,39] <0,0001	0,25 [0,15; 0,34] <0,0001

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Analysepopulation: ITT Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Ustekinumab dar. ITT: Intention-to-Treat; KI: Konfidenzintervall; LOCF: Last observation carried forward; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko					

Hinsichtlich der Stuhlfrequenz ≤ 1 zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-88: Ergebnisse für Anteil der Patienten mit Stuhlfrequenz ≤ 1 – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit Stuhlfrequenz ≤ 1	2,76 [2,25; 3,38] <0,0001	1,49 [1,25; 1,76] <0,0001	1,86 [1,42; 2,42] <0,0001
Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo. Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar. KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial) RR: Relatives Risiko			

Bezüglich der Stuhlfrequenz ≤ 1 zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (RR [95 %-KI]: 1,86 [1,42; 2,42]; p-Wert: < 0,0001). Damit ist die Wahrscheinlichkeit für eine Stuhlfrequenz ≤ 1 unter einer Behandlung mit Upadacitinib ungefähr 1,9-mal höher als unter einer Behandlung mit Ustekinumab.

Tabelle 4-89: Ergebnisse für Anteil der Patienten mit Stuhlfrequenz ≤ 1 – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo RR [95%-KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit Stuhlfrequenz ≤ 1	2,62 [1,88; 3,65] <0,0001	1,22 [1,04; 1,44] 0,0160	2,15 [1,48; 3,10] <0,0001	3,11 [2,26; 4,28] <0,0001	1,44 [1,24; 1,67] <0,0001	2,16 [1,52; 3,08] <0,0001

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko

Bezüglich der Stuhlfrequenz ≤ 1 zeigt sich für die Erhaltungsphase ein statistisch signifikanter Vorteil sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W (RR [95 %-KI]: 2,15 [1,48; 3,10]; p-Wert: < 0,0001), als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W (RR [95 %-KI]: 2,16 [1,52; 3,08]; p-Wert: < 0,0001). Damit ist die Wahrscheinlichkeit für eine Stuhlfrequenz ≤ 1 unter einer Behandlung mit Upadacitinib ungefähr 2,2-mal höher als unter einer Behandlung mit Ustekinumab.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.3.6 Rektale Blutungen – indirekte Vergleiche aus RCT

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

Tabelle 4-90: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	•	•	
1	UNIFI Induktionsstudie		•	•
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	•	•	
1	UNIFI Erhaltungsstudie		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

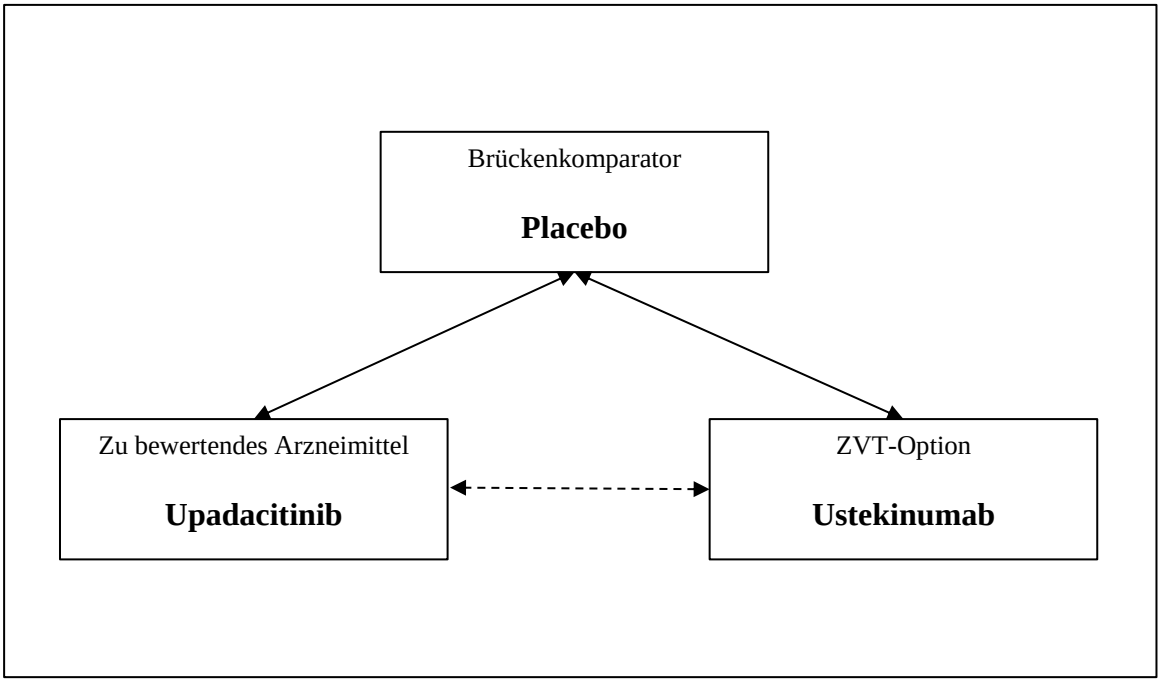

Abbildung 17: Netzwerkstruktur des indirekten Vergleichs

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 rektale Blutung

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Die Schwere der Blutung wird durch den Subscore rektale Blutung des Mayo Gesamtscores auf einer Skala von null („keine Blutung“) bis drei („hauptsächlich Blut“) Punkten kategorisiert. Die rektale Blutung wurde zu Baseline, Woche 2, Woche 4, Woche 6 und zu Woche 8 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer rektalen Blutung = 0 dargestellt. Der Anteil der Patienten beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden mittels NRI- beziehungsweise NRI-C-Analyse ersetzt. Für die Beschreibung der Induktionsphase, wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH meta-analytisch zusammengefasst.
UNIFI Induktionsstudie	Die Schwere der Blutung wird durch den Subscore rektale Blutung des Mayo Gesamtscores auf einer Skala von null („keine Blutung“) bis drei („hauptsächlich Blut“) Punkten kategorisiert. Die rektale Blutung wurde zu Baseline, Woche 2, Woche 4 und zu Woche 8 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer rektalen Blutung = 0 dargestellt. Die Auswertung beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests.

Studie	Operationalisierung
Induktionsstudien	
	Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Die Schwere der Blutung wird durch den Subscore rektale Blutung des Mayo Gesamtscores auf einer Skala von null („keine Blutung“) bis drei („hauptsächlich Blut“) Punkten kategorisiert. Die rektale Blutung wurde zu Baseline, Woche 12, Woche 20, Woche 28, Woche 36, Woche 44 und zu Woche 52 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer rektalen Blutung = 0 dargestellt. Der Anteil der Patienten beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Fehlende Werte wurden per NRI-C-Analyse ersetzt.
UNIFI Erhaltungsstudie	Die Schwere der Blutung wird durch den Subscore rektale Blutung des Mayo Gesamtscores auf einer Skala von null („keine Blutung“) bis drei („hauptsächlich Blut“) Punkten kategorisiert. Die rektale Blutung wurde zu Baseline, Woche 2, Woche 4, Woche 8, Woche 12, Woche 16, Woche 20, Woche 24, Woche 28, Woche 32, Woche 36, Woche 40 und zu Woche 44 erhoben. Für das vorliegende Dossier wird der Anteil an Patienten mit einer rektalen Blutung = 0 dargestellt. Die Auswertung beruht auf der ITT-Population. Die statistische Analyse erfolgte mithilfe eines unadjustierten CMH-Tests. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
CMH: Cochran-Mantel-Haenszel; ITT: Intention-to-Treat; LOCF: Last Observation Carried Forward; MI: Multiple Imputation; NRI: Non-Responder-Imputation; NRI-C: NRI incorporating MI to handle missing data due to COVID-19; SAP: Statistischer Analyseplan	

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 rektale Blutung in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung der rektalen Blutungen erfolgte zu Woche 8 und zu Woche 44. Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

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

Tabelle 4-93: Ergebnisse für Anteil der Patienten mit rektaler Blutung = 0 – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	34/56 (60,7)	12/46 (26,1)	2,33 [1,37; 3,95] 0,0018	4,38 [1,87; 10,23] 0,0007	0,35 [0,17; 0,53] 0,0002
U-ACHIEVE Substudie 2 [45mg]	234/319 (73,4)	43/154 (27,9)	2,63 [2,02; 3,41] <0,0001	7,11 [4,62; 10,93] <0,0001	0,45 [0,37; 0,54] <0,0001
U-ACCOMPLISH [45mg]	235/341 (68,9)	57/174 (32,8)	2,10 [1,68; 2,63] <0,0001	4,55 [3,08; 6,73] <0,0001	0,36 [0,28; 0,45] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	503/716 (70,3)	112/374 (29,9)	2,33 [1,98; 2,74] <0,0001	5,43 [4,13; 7,13] <0,0001	0,40 [0,34; 0,46] <0,0001
Heterogenitätstest (p-Wert)			0,4491		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	89/148 (60,1)	34/149 (22,8)	2,64 [1,91; 3,64] <0,0001	5,10 [3,08; 8,45] <0,0001	0,37 [0,27; 0,48] <0,0001
U-ACHIEVE Substudie 3 [30mg]	114/154 (74,0)	34/149 (22,8)	3,24 [2,38; 4,42] <0,0001	9,64 [5,70; 16,30] <0,0001	0,51 [0,42; 0,61] <0,0001

Analysepopulation: ITT

Imputation: NRI (U-ACHIEVE Substudie 1) und NRI-C (U-ACHIEVE Substudie 2, U-ACHIEVE Substudie 3, U-ACCOMPLISH)

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen.

Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Upadacitinib dar.

ITT: Intention-to-Treat; KI: Konfidenzintervall; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; NRI-C: Non-Responder-Imputation incorporating multiple imputation to handle missing data due to COVID-19; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko

Hinsichtlich der rektalen Blutung = 0 zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts rektale Blutung = 0 für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

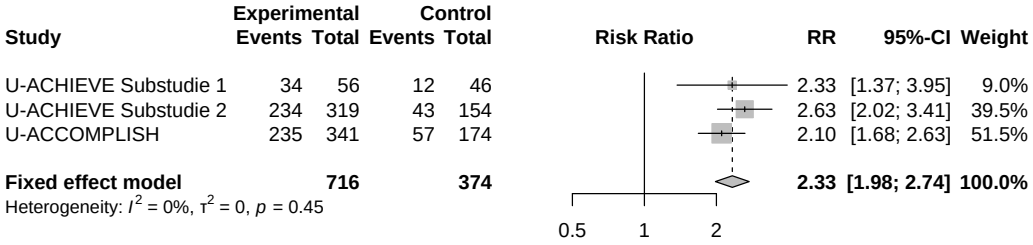

Abbildung 18: Forest-Plot für rektale Blutung = 0 zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-94: Ergebnisse für Anteil der Patienten mit rektaler Blutung = 0 – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Induktionsphase					
UNIFI [6mg/kg]	204/322 (63,4)	119/319 (37,3)	1,70 [1,44; 2,00] <0,0001	2,91 [2,11; 4,00] <0,0001	0,26 [0,19; 0,34] <0,0001
Erhaltungsphase					
UNIFI [90mg/12w]	137/172 (79,7)	101/175 (57,7)	1,38 [1,19; 1,60] <0,0001	2,87 [1,78; 4,62] <0,0001	0,22 [0,12; 0,31] <0,0001
UNIFI [90mg/8w]	139/176 (79,0)	101/175 (57,7)	1,37 [1,18; 1,59] <0,0001	2,75 [1,72; 4,41] <0,0001	0,21 [0,12; 0,31] <0,0001

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Analysepopulation: ITT Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Ein RR und OR größer Eins bzw. eine RD größer Null stellen einen Vorteil von Ustekinumab dar. ITT: Intention-to-Treat; KI: Konfidenzintervall; LOCF: Last observation carried forward; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko					

Hinsichtlich der rektalen Blutung = 0 zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase höhere Ansprechraten unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-95: Ergebnisse für Anteil der Patienten mit rektaler Blutung = 0 – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit rektaler Blutung = 0	2,33 [1,98; 2,74] <0,0001	1,70 [1,44; 2,00] <0,0001	1,37 [1,09; 1,73] 0,0073
Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo. Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar. KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko			

Bezüglich der rektalen Blutung = 0 zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (RR [95 %-KI]: 1,37 [1,09; 1,73]; p-Wert: = 0,0073). Damit ist die Wahrscheinlichkeit für eine rektale Blutung = 0 unter einer Behandlung mit Upadacitinib ungefähr 1,4-mal höher als unter einer Behandlung mit Ustekinumab.

Tabelle 4-96: Ergebnisse für Anteil der Patienten mit rektaler Blutung = 0 – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo RR [95%-KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Anteil der Patienten mit rektaler Blutung = 0	2,64 [1,91; 3,64] <0,0001	1,38 [1,19; 1,60] <0,0001	1,91 [1,34; 2,72] 0,0004	3,24 [2,38; 4,42] <0,0001	1,37 [1,18; 1,59] <0,0001	2,37 [1,68; 3,34] <0,0001
Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo. Ein RR größer Eins im indirekten Vergleich stellt einen Vorteil von Upadacitinib dar. KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko						

Bezüglich dem Anteil der Patienten mit rektaler Blutung = 0 zeigt sich für die Erhaltungsphase ein statistisch signifikanter Vorteil sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W (RR [95 %-KI]: 1,91 [1,34; 2,72]; p-Wert: = 0,0004), als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W (RR [95 %-KI]: 2,37 [1,68; 3,34]; p-Wert: < 0,0001). Damit ist die Wahrscheinlichkeit für eine rektale Blutung = 0 unter einer Behandlung mit Upadacitinib ungefähr 1,9-mal höher, beziehungsweise ungefähr 2,4-mal höher als unter einer Behandlung mit Ustekinumab.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.3.7 EQ-5D VAS – indirekte Vergleiche aus RCT

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

Tabelle 4-97: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	•	•	
1	UNIFI Induktionsstudie		•	•
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	•	•	
1	UNIFI Erhaltungsstudie		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

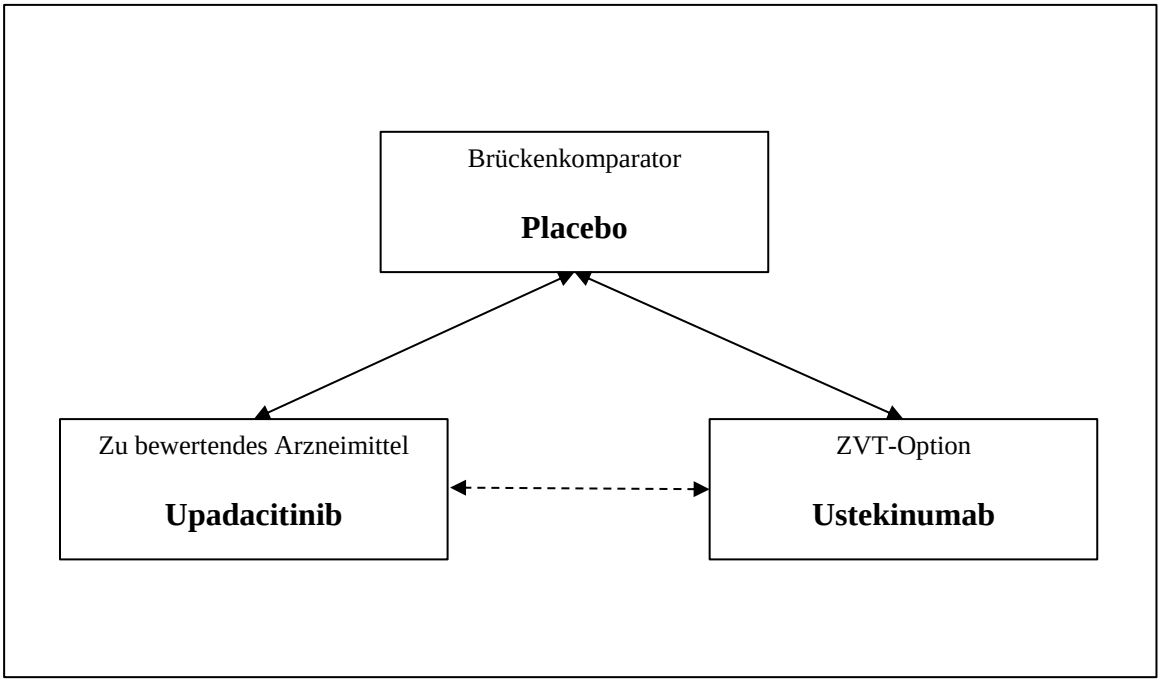

Abbildung 19: Netzwerkstruktur des indirekten Vergleichs

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

Tabelle 4-98: Operationalisierung von EQ-5D VAS

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Zur Beurteilung des allgemeinen Gesundheitszustands wurde die EQ-5D VAS eingesetzt, ein standardisiertes, generisches Erhebungsinstrument, durch welches der Patient seinen Gesundheitszustand auf einer Skala von 0 bis 100 (0: schlechtester vorstellbarer Gesundheitszustand; 100: bester vorstellbarer Gesundheitszustand) einschätzt. Die EQ-5D VAS wurde zu Baseline, Woche 2, Woche 4 und zu Woche 8 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung der EQ-5D VAS zwischen Baseline und Woche 8 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF (U-ACHIEVE Substudie 1) ersetzt, beziehungsweise über ein MMRM (U-ACHIEVE Substudie 2, U-ACCOMPLISH) adressiert. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Für die Beschreibung der Induktionsphase, wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die U-ACCOMPLISH meta-analytisch zusammengefasst.
UNIFI Induktionsstudie	Zur Beurteilung des allgemeinen Gesundheitszustands wurde die EQ-5D VAS eingesetzt, ein standardisiertes, generisches Erhebungsinstrument, durch welches der Patient seinen Gesundheitszustand auf einer Skala von 0 bis 100 (0: schlechtester

Studie	Operationalisierung
Induktionsstudien	
	vorstellbarer Gesundheitszustand; 100: bester vorstellbarer Gesundheitszustand) einschätzt. Die EQ-5D VAS wurde zu Baseline und zu Woche 8 erhoben. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF ersetzt. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Zur Beurteilung des allgemeinen Gesundheitszustands wurde die EQ-5D VAS eingesetzt, ein standardisiertes, generisches Erhebungsinstrument, durch welches der Patient seinen Gesundheitszustand auf einer Skala von 0 bis 100 (0: schlechtester vorstellbarer Gesundheitszustand; 100: bester vorstellbarer Gesundheitszustand) einschätzt. Die EQ-5D VAS wurde zu Baseline und zu Woche 52 erhoben. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Fehlende Werte wurden per RTB-MI-Analyse ersetzt.
UNIFI Erhaltungsstudie	Zur Beurteilung des allgemeinen Gesundheitszustands wurde die EQ-5D VAS eingesetzt, ein standardisiertes, generisches Erhebungsinstrument, durch welches der Patient seinen Gesundheitszustand auf einer Skala von 0 bis 100 (0: schlechtester vorstellbarer Gesundheitszustand; 100: bester vorstellbarer Gesundheitszustand) einschätzt. Die EQ-5D VAS wurde zu Baseline, Woche 8, Woche 20 und zu Woche 44 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung der EQ-5D VAS zwischen Baseline und Woche 44 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF ersetzt. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.

EQ-5D: EuroQol-5 dimensions; ITT: Intention-to-Treat; LOCF: Last Observation Carried Forward; MMRM: Mixed Model for Repeated Measures; N: Anzahl der Patienten; NRI: Non-Responder-Imputation; RTB-MI: Multiple Imputation Incorporating Return-to-Baseline; SAP: Statistischer Analyseplan; SD: Standardabweichung; SE: Standardfehler; VAS: Visuelle Analogskala

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 4-99: Bewertung des Verzerrungspotenzials für EQ-5D VAS in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig

EQ-5D: EuroQol-5 dimensions; ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie; VAS: Visuelle Analogskala

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des EQ-5D VAS erfolgte zu Woche 8 und zu Woche 44 bzw. 52. Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

Stellen Sie die Ergebnisse für den Endpunkt EQ-5D VAS für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-100: Ergebnisse für mittlere Veränderung der EQ-5D VAS – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Endpunkt	**Upadacitinib**				**Placebo**				**Upadacitinib vs. Placebo**
	N	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE*)**	**N**	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE*)**	**Hedges' g [95 %-KI] p-Wert**
Induktionsphase									
U-ACHIEVE Substudie 1 [45mg]	54	50,1 (k. A.)	71,6 (k. A.)	21,5 (23,2)	44	55,7 (k. A.)	54,4 (k. A.)	-1,3 (14,9)	1,13 [0,70; 1,56] <0,0001
U-ACHIEVE Substudie 2 [45mg]	291	55,1 (k. A.)	74,0 (k. A.)	19,3 (1,1)	125	54,6 (k. A.)	61,2 (k. A.)	6,1 (1,6)	0,73 [0,51; 0,94] <0,0001
U-ACCOMPLISH [45mg]	312	54,2 (k. A.)	72,3 (k. A.)	17,6 (1,0)	155	56,1 (k. A.)	60,7 (k. A.)	5,0 (1,4)	0,74 [0,54; 0,94] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	657	k. A. (k. A.)	k. A. (k. A.)	k. A. (k. A.)	324	k. A. (k. A.)	k. A. (k. A.)	k. A. (k. A.)	0,78 [0,64; 0,91] <0,0001
Heterogenitätstest (p-Wert)									0,2268
Erhaltungsphase									
U-ACHIEVE Substudie 3 [15mg]	148	54,5 (k. A.)	72,5 (k. A.)	17,4 (1,7)	149	57,7 (k. A.)	63,6 (k. A.)	7,7 (1,7)	0,48 [0,24; 0,71] <0,0001
U-ACHIEVE Substudie 3 [30mg]	154	53,5 (k. A.)	77,1 (k. A.)	20,4 (1,6)	149	57,7 (k. A.)	63,6 (k. A.)	7,7 (1,7)	0,63 [0,40; 0,86] <0,0001

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	Hedges‘ g [95 %-KI] p-Wert
*: Für U-ACHIEVE Substudie 1 wird statt dem SE die SD dargestellt. Analysepopulation: ITT Imputation: LOCF (U-ACHIEVE Substudie 1), MMRM (U-ACHIEVE Substudie 2, U-ACCOMPLISH), RTB-MI (U-ACHIEVE Substudie 3) Stratifizierungsvariablen: Bio-IR-Status, Einnahme von Kortikosteroiden zu Baseline, Adaptierter Mayo Score (≤ 7 oder > 7) zu Baseline Angaben zum Hedges‘ g basieren für die U-ACHIEVE Substudie 1 und die U-ACHIEVE Substudie 3 auf einem ANCOVA-Modell mit Behandlung, dem Baseline-Wert und den Stratifizierungsvariablen als Kovariablen. Für die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH basiert das Hedges‘ g auf einem MMRM-Modell mit Behandlung, dem Baseline-Wert, Visite, Behandlung-Visite-Interaktion und den Stratifizierungsvariablen als Kovariablen. Die Meta-Analyse basiert auf den Studien U-ACHIEVE Substudie 1 und U-ACHIEVE Substudie 2 und U-ACCOMPLISH. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen. Ein Hedges‘ g größer Null stellt einen Vorteil von Upadacitinib dar. ANCOVA: Kovarianzanalyse; EQ-5D-VAS: EuroQol-5 dimensions; ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; LOCF: Last observation carried forward; LS: Least-Squares; MI: Multiple Imputation; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl Patienten in der Analyse; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; RTB-MI: Multiple Imputation Incorporating Return-to-Baseline; SD: Standardabweichung (Standard Deviation); SE: Standardfehler; VAS: Visuelle Analogskala;								

Hinsichtlich der EQ-5D VAS zeigt sich sowohl für die Induktionsphase als auch für die Erhaltungsphase eine größere Verbesserung unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts mittlere Veränderung der EQ-5D VAS für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

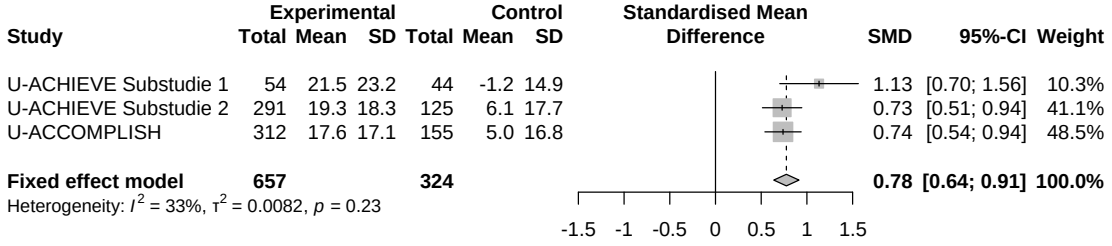

Abbildung 20: Forest-Plot für die mittlere Veränderung der EQ-5D VAS zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-101: Ergebnisse für mittlere Veränderung der EQ-5D VAS – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Endpunkt	**Ustekinumab**				**Placebo**				**Ustekinumab vs. Placebo**
	N	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE)**	**N**	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE)**	**Hedges‘ g [95 %-KI] p-Wert**
Induktionsphase									
UNIFI [6mg/kg]	322	k. A. (k. A.)	k. A. (k. A.)	13,5 (18,4)	317	k. A. (k. A.)	k. A. (k. A.)	5,7 (19,6)	0,41 [0,25; 0,57] <0,0001
Erhaltungsphase									
UNIFI [90mg/12w]	172	k. A. (k. A.)	k. A. (k. A.)	-2,2 (19,9)	175	k. A. (k. A.)	k. A. (k. A.)	-7,7 (18,8)	0,28 [0,07; 0,50] 0,0085
UNIFI [90mg/8w]	176	k. A. (k. A.)	k. A. (k. A.)	2,4 (17,3)	175	k. A. (k. A.)	k. A. (k. A.)	-7,7 (18,8)	0,56 [0,35; 0,77] <0,0001

Analysepopulation: ITT

Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte

Stratifizierungsvariablen: Bio-IR-Status, geografische Region

Angaben zum Hedges‘ g basieren auf einem ANCOVA-Modell mit dem Baseline-Wert und den Stratifizierungsvariablen als Kovariablen.

Ein Hedges‘ g größer Null stellt einen Vorteil von Ustekinumab dar.

ANCOVA: Kovarianzanalyse; EQ-5D-VAS: EuroQol-5 dimensions; ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; LOCF: Last observation carried forward; LS: Least-Squares; MW: Mittelwert; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; SD: Standardabweichung (Standard Deviation); SE: Standardfehler; VAS. Visuelle Analogskala

Hinsichtlich der EQ-5D VAS zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase Vorteile unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-102: Ergebnisse für mittlere Veränderung der EQ-5D VAS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo Hedges‘ g [95% KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**
mittlere Veränderung der EQ-5D VAS	0,78 [0,64; 0,91] <0,0001	0,41 [0,25; 0,57] <0,0001	0,37 [0,16; 0,57] 0,0006

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein Hedges‘ g größer Null im indirekten Vergleich, stellt einen Vorteil von Upadacitinib dar.

EQ-5D: EuroQol-5 dimensions; KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); VAS: Visuelle Analogskala

Bezüglich der mittleren Veränderung der EQ-5D VAS zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (Hedges’ g [95 %-KI]; 0,37 [0,16; 0,57]; p-Wert = 0,0006).

Tabelle 4-103: Ergebnisse für mittlere Veränderung der EQ-5D VAS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo Hedges‘ g [95% KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**
mittlere Veränderung der EQ-5D VAS	0,48 [0,24; 0,71] <0,0001	0,28 [0,07; 0,50] 0,0085	0,19 [-0,12; 0,50] 0,2310	0,63 [0,40; 0,86] <0,0001	0,56 [0,35; 0,77] <0,0001	0,07 [-0,24; 0,38] 0,6594

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein Hedges‘ g größer Null im indirekten Vergleich, stellt einen Vorteil von Upadacitinib dar.

EQ-5D: EuroQol-5 dimensions; KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); VAS: Visuelle Analogskala

Bezüglich der mittleren Veränderung der EQ-5D VAS zeigt sich für die Erhaltungsphase ein numerischer Vorteil von Upadacitinib gegenüber Ustekinumab, sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.3.8 Gesundheitsbezogene Lebensqualität: IBDQ – indirekte Vergleiche aus RCT

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

Tabelle 4-104: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	•	•	
1	UNIFI Induktionsstudie		•	•
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	•	•	
1	UNIFI Erhaltungsstudie		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

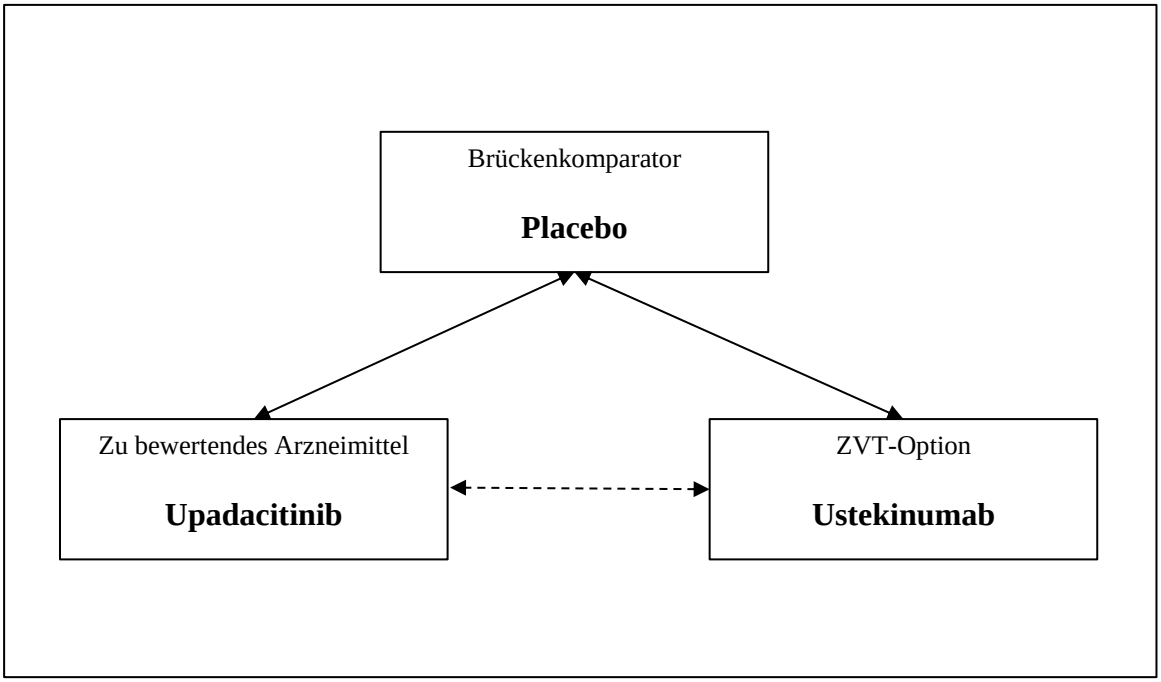

Abbildung 21: Netzwerkstruktur des indirekten Vergleichs

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

Tabelle 4-105: Operationalisierung von IBDQ

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Der IBDQ umfasst 32 Fragen, welche vier untergeordnete Komponenten beschreiben: Darmsymptome (10 Elemente), systemische Symptome (5 Elemente), emotionale Funktion (12 Elemente) und Sozialfunktion (5 Elemente). Jede Frage kann auf einer Skala von 1 (schlechteste Situation) bis 7 (beste Situation) bewertet werden. Die Gesamtpunktzahl reicht von 32 bis 224 Punkten, wobei eine höhere Gesamtpunktzahl eine bessere gesundheitsbezogene Lebensqualität indiziert. Der IBDQ wurde zu Baseline, Woche 2, Woche 4 und zu Woche 8 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung des IBDQ Gesamtscore zwischen Baseline und Woche 8 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF (U-ACHIEVE Substudie 1) ersetzt, beziehungsweise über ein MMRM (U-ACHIEVE Substudie 2, U-ACCOMPLISH) adressiert. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Für die Beschreibung der Induktionsphase, wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH meta-analytisch zusammengefasst.

Studie	Operationalisierung
UNIFI Induktionsstudie	Der IBDQ umfasst 32 Fragen, welche vier untergeordnete Komponenten beschreiben: Darmsymptome (10 Elemente), systemische Symptome (5 Elemente), emotionale Funktion (12 Elemente) und Sozialfunktion (5 Elemente). Jede Frage kann auf einer Skala von 1 (schlechteste Situation) bis 7 (beste Situation) bewertet werden. Die Gesamtpunktzahl reicht von 32 bis 224 Punkten, wobei eine höhere Gesamtpunktzahl eine bessere gesundheitsbezogene Lebensqualität indiziert. Der IBDQ wurde zu Baseline und zu Woche 8 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung des IBDQ Gesamtscore zwischen Baseline und Woche 8 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF ersetzt. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Der IBDQ umfasst 32 Fragen, welche vier untergeordnete Komponenten beschreiben: Darmsymptome (10 Elemente), systemische Symptome (5 Elemente), emotionale Funktion (12 Elemente) und Sozialfunktion (5 Elemente). Jede Frage kann auf einer Skala von 1 (schlechteste Situation) bis 7 (beste Situation) bewertet werden. Die Gesamtpunktzahl reicht von 32 bis 224 Punkten, wobei eine höhere Gesamtpunktzahl eine bessere gesundheitsbezogene Lebensqualität indiziert. Der IBDQ wurde zu Baseline und zu Woche 52 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung des IBDQ Gesamtscore zwischen Baseline und Woche 52 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Fehlende Werte wurden per RTB-MI Analyse ersetzt.
UNIFI Erhaltungsstudie	Der IBDQ umfasst 32 Fragen, welche vier untergeordnete Komponenten beschreiben: Darmsymptome (10 Elemente), systemische Symptome (5 Elemente), emotionale Funktion (12 Elemente) und Sozialfunktion (5 Elemente). Jede Frage kann auf einer Skala von 1 (schlechteste Situation) bis 7 (beste Situation) bewertet werden. Die Gesamtpunktzahl reicht von 32 bis 224 Punkten, wobei eine höhere Gesamtpunktzahl eine bessere gesundheitsbezogene Lebensqualität indiziert. Der IBDQ wurde zu Baseline, Woche 8, Woche 20 und zu Woche 44 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung des IBDQ Gesamtscore zwischen Baseline und Woche 44 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF ersetzt. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt.

Studie	Operationalisierung
	Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
IBDQ: Inflammatory Bowel Disease Questionnaire; ITT: Intention-to-Treat; LOCF: Last Observation Carried Forward; MMRM: Mixed Model for Repeated Measures; N: Anzahl der Patienten; NRI: Non-Responder-Imputation; RTB-MI: Multiple Imputation Incorporating Return-to-Baseline; SAP: Statistischer Analyseplan; SD: Standardabweichung; SE: Standardfehler	

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 4-106: Bewertung des Verzerrungspotenzials für IBDQ in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig
IBDQ: Inflammatory Bowel Disease Questionnaire; ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des IBDQ erfolgte zu Woche 8 und zu Woche 44 bzw. 52. Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population

verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

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

Tabelle 4-107: Ergebnisse für mittlere Veränderung des IBDQ – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Endpunkt	**Upadacitinib**				**Placebo**				**Upadacitinib vs. Placebo**
	N	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE*)**	**N**	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE*)**	**Hedges' g [95 %-KI] p-Wert**
Induktionsphase									
U-ACHIEVE Substudie 1 [45mg]	55	128,4 (k. A.)	177,3 (k. A.)	48,9 (32,1)	44	129,1 (k. A.)	141,0 (k. A.)	12,0 (26,5)	1,23 [0,80; 1,67] <0,0001
U-ACHIEVE Substudie 2 [45mg]	292	123,5 (k. A.)	177,7 (k. A.)	55,3 (1,9)	125	121,8 (k. A.)	144,7 (k. A.)	21,7 (2,9)	1,02 [0,80; 1,24] <0,0001
U-ACCOMPLISH [45mg]	315	123,3 (k. A.)	176,2 (k. A.)	52,2 (1,9)	156	125,4 (k. A.)	146,4 (k. A.)	21,1 (2,6)	0,94 [0,74; 1,14] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	662	k. A. (k. A.)	k. A. (k. A.)	k. A. (k. A.)	325	k. A. (k. A.)	k. A. (k. A.)	k. A. (k. A.)	1,01 [0,87; 1,15] <0,0001
Heterogenitätstest (p-Wert)									0,4862
Erhaltungsphase									
U-ACHIEVE Substudie 3 [15mg]	148	125,8 (k. A.)	176,0 (k. A.)	49,2 (3,4)	149	122,6 (k. A.)	144,3 (k. A.)	17,9 (3,6)	0,73 [0,50; 0,97] <0,0001
U-ACHIEVE Substudie 3 [30mg]	154	121,3 (k. A.)	184,2 (k. A.)	58,9 (3,4)	149	122,6 (k. A.)	144,3 (k. A.)	17,9 (3,6)	0,94 [0,71; 1,18] <0,0001

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	Hedges‘ g [95 %-KI] p-Wert
*: Für U-ACHIEVE Substudie 1 wird statt dem SE die SD dargestellt. Analysepopulation: ITT Imputation: LOCF (U-ACHIEVE Substudie 1), MMRM (U-ACHIEVE Substudie 2, U-ACCOMPLISH), RTB-MI (U-ACHIEVE Substudie 3) Stratifizierungsvariablen: Bio-IR-Status, Einnahme von Kortikosteroiden zu Baseline, Adaptierter Mayo Score (≤ 7 oder > 7) zu Baseline Angaben zum Hedges‘ g basieren für die U-ACHIEVE Substudie 1 und die U-ACHIEVE Substudie 3 auf einem ANCOVA-Modell mit Behandlung, dem Baseline-Wert und den Stratifizierungsvariablen als Kovariablen. Für die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH basiert das Hedges‘ g auf einem MMRM-Modell mit Behandlung, dem Baseline-Wert, Visite, Behandlung-Visite-Interaktion und den Stratifizierungsvariablen als Kovariablen. Die Meta-Analyse basiert auf den Studien U-ACHIEVE Substudie 1 und U-ACHIEVE Substudie 2 und U-ACCOMPLISH. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen. Ein Hedges‘ g größer Null stellt einen Vorteil von Upadacitinib dar. ANCOVA: Kovarianzanalyse; ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; LOCF: Last observation carried forward; LS: Least-Squares; MI: Multiple Imputation; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl Patienten in der Analyse; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RTB-MI: Multiple Imputation Incorporating Return-to-Baseline; SD: Standardabweichung (Standard Deviation); SE: Standardfehler								

Hinsichtlich des IBDQ zeigt sich sowohl für die Induktionsphase als auch für die Erhaltungsphase eine größere Verbesserung unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts mittlere Veränderung des IBDQ für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

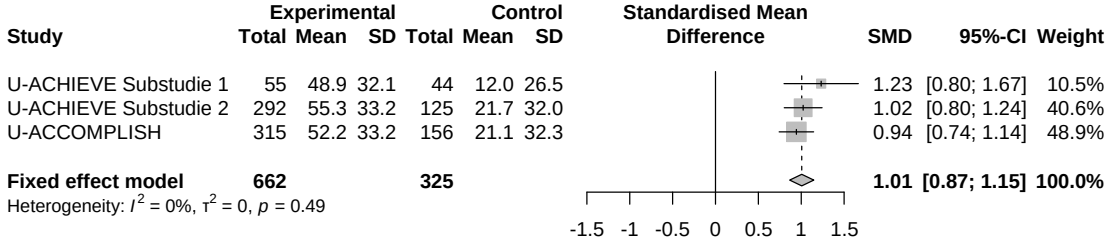

Abbildung 22: Forest-Plot für die mittlere Veränderung des IBDQ zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-108: Ergebnisse für mittlere Veränderung des IBDQ – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Endpunkt	**Ustekinumab**				**Placebo**				**Ustekinumab vs. Placebo**
	N	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE)**	**N**	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE)**	**Hedges‘ g [95 %-KI] p-Wert**
Induktionsphase									
UNIFI [6mg/kg]	322	127,0 (k. A.)	161,9 (k. A.)	35,0 (31,9)	319	127,4 (k. A.)	143,5 (k. A.)	16,1 (31,4)	0,60 [0,44; 0,76] <0,0001
Erhaltungsphase									
UNIFI [90mg/12w]	172	175,4 (k. A.)	172,3 (k. A.)	-3,0 (32,9)	175	174,3 (k. A.)	159,3 (k. A.)	-15,1 (35,4)	0,35 [0,14; 0,57] 0,0011
UNIFI [90mg/8w]	176	174,1 (k. A.)	178,2 (k. A.)	3,9 (31,5)	175	174,3 (k. A.)	159,3 (k. A.)	-15,1 (35,4)	0,57 [0,35; 0,78] <0,0001

Analysepopulation: ITT

Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte

Stratifizierungsvariablen: Bio-IR-Status, geografische Region

Angaben zum Hedges‘ g basieren auf einem ANCOVA-Modell mit dem Baseline-Wert und den Stratifizierungsvariablen als Kovariablen.

Ein Hedges‘ g größer Null stellt einen Vorteil von Ustekinumab dar.

ANCOVA: Kovarianzanalyse; ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; LOCF: Last observation carried forward; LS: Least-Squares; MW: Mittelwert; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation); SE: Standardfehler

Hinsichtlich des IBDQ zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase Vorteile unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-109: Ergebnisse für mittlere Veränderung des IBDQ – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo Hedges‘ g [95% KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**
mittlere Veränderung des IBDQ	1,01 [0,87; 1,15] <0,0001	0,60 [0,44; 0,76] <0,0001	0,41 [0,20; 0,62] 0,0002

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein Hedges‘ g größer Null im indirekten Vergleich, stellt einen Vorteil von Upadacitinib dar.

IBDQ: Inflammatory Bowel Disease Questionnaire; KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial)

Bezüglich der mittleren Veränderung des IBDQ zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (Hedges’ g [95 %-KI]; 0,41 [0,20; 0,62]; p-Wert = 0,0002).

Tabelle 4-110: Ergebnisse für mittlere Veränderung des IBDQ – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo Hedges‘ g [95% KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**
mittlere Veränderung des IBDQ	0,73 [0,50; 0,97] <0,0001	0,35 [0,14; 0,57] 0,0011	0,38 [0,06; 0,69] 0,0195	0,94 [0,71; 1,18] <0,0001	0,57 [0,35; 0,78] <0,0001	0,38 [0,06; 0,70] 0,0205

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein Hedges‘ g größer Null im indirekten Vergleich, stellt einen Vorteil von Upadacitinib dar.

IBDQ: Inflammatory Bowel Disease Questionnaire; KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial)

Bezüglich der mittleren Veränderung des IBDQ zeigt sich für die Erhaltungsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab, sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W (Hedges’ g [95 %-KI]; 0,38 [0,06; 0,69]; p-Wert = 0,0195) als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W (Hedges’ g [95 %-KI]; 0,38 [0,06; 0,70]; p-Wert = 0,0205).

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.3.9 Gesundheitsbezogene Lebensqualität: SF-36 – indirekte Vergleiche aus RCT

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

Tabelle 4-111: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	●	●	
1	UNIFI Induktionsstudie		●	●
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	●	●	
1	UNIFI Erhaltungsstudie		●	●

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

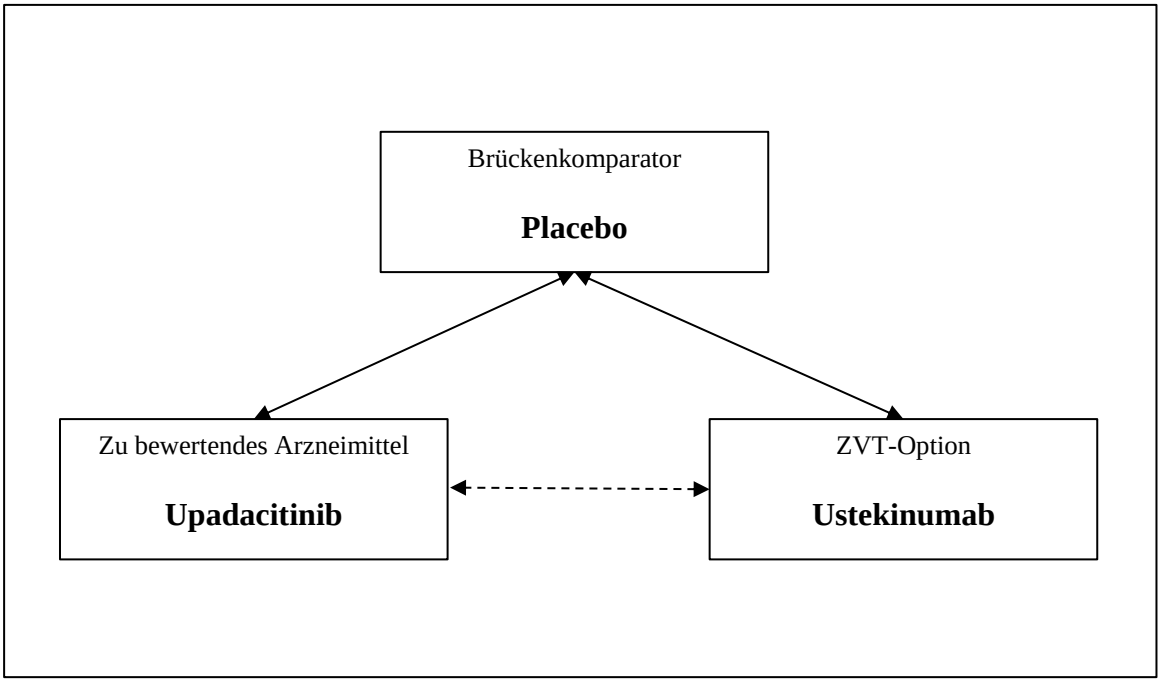

Abbildung 23: Netzwerkstruktur des indirekten Vergleichs

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

Tabelle 4-112: Operationalisierung von SF-36 PCS/MCS

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Der SF-36 erfasst acht Dimensionen der gesundheitsbezogenen Lebensqualität mittels 36 Items. Aus diesen Dimensionen können die Summenscores PCS und MCS errechnet werden. Die Berechnung des SF-36 PCS und des SF-36 MCS ist in Abschnitt 4.2.5.2 beschrieben. Der SF-36 wurde zu Baseline, Woche 2, Woche 4 und zu Woche 8 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung des SF-36 PCS/MCS zwischen Baseline und Woche 8 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF (U-ACHIEVE Substudie 1) ersetzt, beziehungsweise über ein MMRM (U-ACHIEVE Substudie 2, U-ACCOMPLISH) adressiert. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Für die Beschreibung der Induktionsphase, wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH meta-analytisch zusammengefasst.
UNIFI Induktionsstudie	Der SF-36 erfasst acht Dimensionen der gesundheitsbezogenen Lebensqualität mittels 36 Items. Aus diesen Dimensionen können die Summenscores PCS und MCS errechnet

Studie	Operationalisierung
	werden. Die Berechnung des SF-36 PCS und des SF-36 MCS ist in Abschnitt 4.2.5.2 beschrieben. Der SF-36 wurde zu Baseline und zu Woche 8 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung des SF-36 PCS/MCS zwischen Baseline und Woche 8 für die Induktionsphase basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF ersetzt. Für den Fall dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Der SF-36 erfasst acht Dimensionen der gesundheitsbezogenen Lebensqualität mittels 36 Items. Aus diesen Dimensionen können die Summenscores PCS und MCS errechnet werden. Die Berechnung des SF-36 PCS und des SF-36 MCS ist in Abschnitt 4.2.5.2 beschrieben. Der SF-36 wurde zu Baseline und zu Woche 52 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung des SF-36 PCS/MCS zwischen Baseline und Woche 52 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Für den Fall, dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Fehlende Werte wurden per RTB-MI Analyse ersetzt.
UNIFI Erhaltungsstudie	Der SF-36 erfasst acht Dimensionen der gesundheitsbezogenen Lebensqualität mittels 36 Items. Aus diesen Dimensionen können die Summenscores PCS und MCS errechnet werden. Die Berechnung des SF-36 PCS und des SF-36 MCS ist in Abschnitt 4.2.5.2 beschrieben. Der SF-36 wurde zu Baseline, Woche 8, Woche 20 und zu Woche 44 erhoben. Für das vorliegende Dossier wird die mittlere Veränderung des SF-36 PCS/MCS zwischen Baseline und Woche 44 basierend auf der ITT-Population dargestellt. Die statistische Analyse resultiert aus einer Berechnung der standardisierten Mittelwertdifferenz per inversen Varianz-Methode mittels der Populationsgröße, der mittleren Veränderung und der zugehörigen Standardabweichung. Fehlende Werte wurden per LOCF ersetzt. Für den Fall dass die Standardabweichung der mittleren Veränderung nicht gegeben ist, erfolgte eine Schätzung über den Standardfehler mittels der Formel $SD = SE * \sqrt{N}$. Eine detaillierte Angabe für die statistischen Methoden ist im Abschnitt 4.2.5.2 aufgeführt. Gemäß SAP wurden fehlende Werte für stetige Endpunkte per LOCF und fehlende Werte für dichotome Endpunkte per NRI ersetzt.

Studie	Operationalisierung
ITT: Intention-to-Treat; LOCF: Last Observation Carried Forward; MCS: Mental Health Component Summary Score; MMRM: Mixed Model for Repeated Measures; N: Anzahl der Patienten; NRI: Non-Responder-Imputation; PCS: Physical Health Component Summary Score; RTB-MI: Multiple Imputation Incorporating Return-to-Baseline; SAP: Statistischer Analyseplan; SD: Standardabweichung; SE: Standardfehler; SF-36: Short Form – 36	

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 4-113: Bewertung des Verzerrungspotenzials für SF-36 PCS/MCS in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; MCS: Mental Health Component Summary Score; PCS: Physical Health Component Summary Score; RCT: Randomisierte kontrollierte Studie; SF-36: Short Form – 36						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des SF-36 PCS/MCS erfolgte zu Woche 8 und zu Woche 44 bzw. 52. Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend

der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

SF-36 PCS

Stellen Sie die Ergebnisse für den Endpunkt SF-36 PCS/MCS für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-114: Ergebnisse für mittlere Veränderung des SF-36 PCS – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	Hedges‘ g [95 %-KI] p-Wert
Induktionsphase									
U-ACHIEVE Substudie 1 [45mg]	54	42,9 (k. A.)	50,8 (k. A.)	7,9 (6,7)	44	43,6 (k. A.)	44,1 (k. A.)	0,5 (6,6)	1,10 [0,67; 1,52] <0,0001
U-ACHIEVE Substudie 2 [45mg]	291	43,7 (k. A.)	50,9 (k. A.)	7,4 (0,4)	125	42,8 (k. A.)	45,3 (k. A.)	2,3 (0,5)	0,83 [0,62; 1,05] <0,0001
U-ACCOMPLISH [45mg]	313	42,4 (k. A.)	50,2 (k. A.)	7,6 (0,4)	156	43,9 (k. A.)	46,5 (k. A.)	2,9 (0,5)	0,75 [0,55; 0,95] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	658	k. A. (k. A.)	k. A. (k. A.)	k. A. (k. A.)	325	k. A. (k. A.)	k. A. (k. A.)	k. A. (k. A.)	0,82 [0,68; 0,96] <0,0001
Heterogenitätstest (p-Wert)									0,3482
Erhaltungsphase									
U-ACHIEVE Substudie 3 [15mg]	148	43,2 (k. A.)	51,6 (k. A.)	8,0 (0,6)	149	43,2 (k. A.)	46,0 (k. A.)	2,6 (0,6)	0,69 [0,45; 0,92] <0,0001

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	Hedges‘ g [95 %-KI] p-Wert
U-ACHIEVE Substudie 3 [30mg]	154	42,6 (k. A.)	52,5 (k. A.)	9,0 (0,6)	149	43,2 (k. A.)	46,0 (k. A.)	2,6 (0,6)	0,82 [0,58; 1,05] <0,0001

*: Für U-ACHIEVE Substudie 1 wird statt dem SE die SD dargestellt.

Analysepopulation: ITT

Imputation: LOCF (U-ACHIEVE Substudie 1), MMRM (U-ACHIEVE Substudie 2, U-ACCOMPLISH), RTB-MI (U-ACHIEVE Substudie 3)

Stratifizierungsvariablen: Bio-IR-Status, Einnahme von Kortikosteroiden zu Baseline, Adaptierter Mayo Score (≤ 7 oder > 7) zu Baseline

Angaben zum Hedges‘ g basieren für die U-ACHIEVE Substudie 1 und die U-ACHIEVE Substudie 3 auf einem ANCOVA-Modell mit Behandlung, dem Baseline-Wert und den Stratifizierungsvariablen als Kovariablen. Für die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH basiert das Hedges‘ g auf einem MMRM-Modell mit Behandlung, dem Baseline-Wert, Visite, Behandlung-Visite-Interaktion und den Stratifizierungsvariablen als Kovariablen. Die Meta-Analyse basiert auf den Studien U-ACHIEVE Substudie 1 und U-ACHIEVE Substudie 2 und U-ACCOMPLISH. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen.

Ein Hedges‘ g größer Null stellt einen Vorteil von Upadacitinib dar.

ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; LOCF: Last observation carried forward; LS: Least-Squares; MI: Multiple Imputation; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl Patienten in der Analyse; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; RTB-MI: Multiple Imputation Incorporating Return-to-Baseline; SD: Standardabweichung (Standard Deviation); SE: Standardfehler; SF-36: Short Form-36

Hinsichtlich des SF-36 PCS zeigt sich sowohl für die Induktionsphase als auch für die Erhaltungsphase eine größere Verbesserung unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts mittlere Veränderung des SF-36 PCS für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

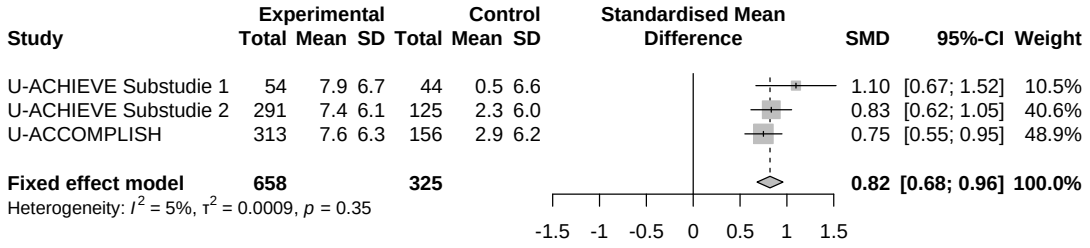

Abbildung 24: Forest-Plot für die mittlere Veränderung des SF-36 PCS zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-115: Ergebnisse für mittlere Veränderung des SF-36 PCS – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Endpunkt	Ustekinumab				Placebo				Ustekinumab vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE)	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE)	Hedges‘ g [95 %-KI] p-Wert
Induktionsphase									
UNIFI [6mg/kg]	322	k. A. (k. A.)	k. A. (k. A.)	5,2 (6,2)	319	k. A. (k. A.)	k. A. (k. A.)	2,1 (6,4)	0,49 [0,34; 0,65] <0,0001
Erhaltungsphase									
UNIFI [90mg/12w]	172	k. A. (k. A.)	k. A. (k. A.)	-0,4 (7,1)	175	k. A. (k. A.)	k. A. (k. A.)	-1,7 (6,5)	0,19 [-0,02; 0,40] 0,0763
UNIFI [90mg/8w]	176	k. A. (k. A.)	k. A. (k. A.)	1,3 (5,7)	175	k. A. (k. A.)	k. A. (k. A.)	-1,7 (6,5)	0,49 [0,28; 0,71] <0,0001

Analysepopulation: ITT

Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte

Stratifizierungsvariablen: Bio-IR-Status, geografische Region

Angaben zum Hedges‘ g basieren auf einem ANCOVA-Modell mit dem Baseline-Wert und den Stratifizierungsvariablen als Kovariablen.

Ein Hedges‘ g größer Null stellt einen Vorteil von Ustekinumab dar.

ANCOVA: Kovarianzanalyse; ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; LOCF: Last observation carried forward; LS: Least-Squares; MW: Mittelwert; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; SD: Standardabweichung (Standard Deviation); SE: Standardfehler; SF-36: Short Form-36

Hinsichtlich des SF-36 PCS zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase Vorteile unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-116: Ergebnisse für mittlere Veränderung des SF-36 PCS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo Hedges‘ g [95% KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**
mittlere Veränderung des SF-36 PCS	0,82 [0,68; 0,96] <0,0001	0,49 [0,34; 0,65] <0,0001	0,33 [0,12; 0,54] 0,0023

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein Hedges‘ g größer Null im indirekten Vergleich, stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; PCS: Physical Health Component Summary Score; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); SF-36: Short Form – 36

Bezüglich der mittleren Veränderung des SF-36 PCS zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (Hedges’ g [95 %-KI]; 0,33 [0,12; 0,54]; p-Wert = 0,0023).

Tabelle 4-117: Ergebnisse für mittlere Veränderung des SF-36 PCS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo Hedges‘ g [95% KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**
mittlere Veränderung des SF-36 PCS	0,69 [0,45; 0,92] <0,0001	0,19 [-0,02; 0,40] 0,0763	0,49 [0,18; 0,81] 0,0021	0,82 [0,58; 1,05] <0,0001	0,49 [0,28; 0,71] <0,0001	0,33 [0,01; 0,64] 0,0432

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein Hedges‘ g größer Null im indirekten Vergleich, stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; PCS: Physical Health Component Summary Score; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); SF-36: Short Form – 36

Bezüglich der mittleren Veränderung des SF-36 PCS zeigt sich für die Erhaltungsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab, sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W (Hedges’ g [95 %-KI]; 0,49 [0,18; 0,81]; p-Wert = 0,0021) als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W (Hedges’ g [95 %-KI]; 0,33 [0,01; 0,64]; p-Wert = 0,0432).

SF-36 MCS

Stellen Sie die Ergebnisse für den Endpunkt SF-36 PCS/MCS für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-118: Ergebnisse für mittlere Veränderung des SF-36 MCS – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	Hedges' g [95 %-KI] p-Wert
Induktionsphase									
U-ACHIEVE Substudie 1 [45mg]	54	41,5 (k. A.)	48,8 (k. A.)	7,3 (9,7)	44	41,2 (k. A.)	42,8 (k. A.)	1,6 (7,5)	0,64 [0,23; 1,05] 0,0020
U-ACHIEVE Substudie 2 [45mg]	291	40,4 (k. A.)	48,6 (k. A.)	8,3 (0,5)	125	40,9 (k. A.)	43,5 (k. A.)	2,4 (0,8)	0,69 [0,47; 0,90] <0,0001
U-ACCOMPLISH [45mg]	313	40,9 (k. A.)	48,4 (k. A.)	7,3 (0,5)	156	40,7 (k. A.)	43,7 (k. A.)	2,7 (0,7)	0,53 [0,34; 0,73] <0,0001
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	658	k. A. (k. A.)	k. A. (k. A.)	k. A. (k. A.)	325	k. A. (k. A.)	k. A. (k. A.)	k. A. (k. A.)	0,61 [0,47; 0,74] <0,0001
Heterogenitätstest (p-Wert)									0,5708
Erhaltungsphase									
U-ACHIEVE Substudie 3 [15mg]	148	42,2 (k. A.)	48,1 (k. A.)	6,4 (0,9)	149	40,7 (k. A.)	43,6 (k. A.)	2,4 (0,9)	0,38 [0,15; 0,61] 0,0013
U-ACHIEVE Substudie 3 [30mg]	154	40,7 (k. A.)	49,1 (k. A.)	7,7 (0,9)	149	40,7 (k. A.)	43,6 (k. A.)	2,4 (0,9)	0,49 [0,26; 0,72] <0,0001

Endpunkt	Upadacitinib				Placebo				Upadacitinib vs. Placebo
	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	N	Baseline MW (SD)	Visite MW (SD)	Δ LS MW (SE*)	Hedges‘ g [95 %-KI] p-Wert
*: Für U-ACHIEVE Substudie 1 wird statt dem SE die SD dargestellt. Analysepopulation: ITT Imputation: LOCF (U-ACHIEVE Substudie 1), MMRM (U-ACHIEVE Substudie 2, U-ACCOMPLISH), RTB-MI (U-ACHIEVE Substudie 3) Stratifizierungsvariablen: Bio-IR-Status, Einnahme von Kortikosteroiden zu Baseline, Adaptierter Mayo Score (≤ 7 oder > 7) zu Baseline Angaben zum Hedges‘ g basieren für die U-ACHIEVE Substudie 1 und die U-ACHIEVE Substudie 3 auf einem ANCOVA-Modell mit Behandlung, dem Baseline-Wert und den Stratifizierungsvariablen als Kovariablen. Für die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH basiert das Hedges‘ g auf einem MMRM-Modell mit Behandlung, dem Baseline-Wert, Visite, Behandlung-Visite-Interaktion und den Stratifizierungsvariablen als Kovariablen. Die Meta-Analyse basiert auf den Studien U-ACHIEVE Substudie 1 und U-ACHIEVE Substudie 2 und U-ACCOMPLISH. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen. Ein Hedges‘ g größer Null stellt einen Vorteil von Upadacitinib dar. ANCOVA: Kovarianzanalyse; ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; LOCF: Last observation carried forward; LS: Least-Squares; MI: Multiple Imputation; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl Patienten in der Analyse; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; RTB-MI: Multiple Imputation Incorporating Return-to-Baseline; SD: Standardabweichung (Standard Deviation); SE: Standardfehler; SF-36: Short Form-36									

Hinsichtlich des SF-36 MCS zeigt sich sowohl für die Induktionsphase als auch für die Erhaltungsphase eine größere Verbesserung unter einer Behandlung mit Upadacitinib als unter einer Behandlung mit Placebo.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts mittlere Veränderung des SF-36 MCS für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

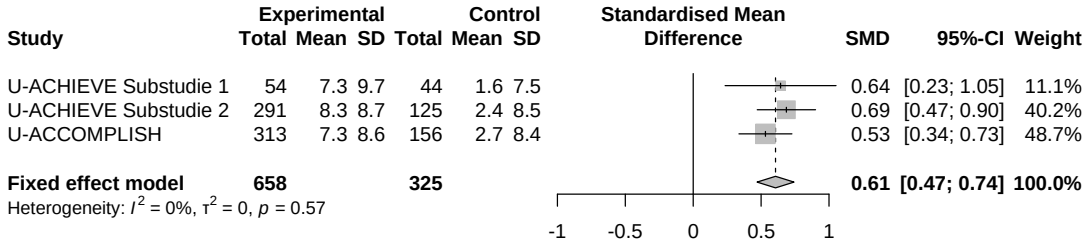

Abbildung 25: Forest-Plot für die mittlere Veränderung des SF-36 MCS zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-119: Ergebnisse für mittlere Veränderung des SF-36 MCS – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Endpunkt	**Ustekinumab**				**Placebo**				**Ustekinumab vs. Placebo**
	N	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE)**	**N**	**Baseline MW (SD)**	**Visite MW (SD)**	**Δ LS MW (SE)**	**Hedges‘ g [95 %-KI] p-Wert**
Induktionsphase									
UNIFI [6mg/kg]	322	k. A. (k. A.)	k. A. (k. A.)	5,1 (9,7)	319	k. A. (k. A.)	k. A. (k. A.)	2,2 (10,2)	0,29 [0,14; 0,45] 0,0003
Erhaltungsphase									
UNIFI [90mg/12w]	172	k. A. (k. A.)	k. A. (k. A.)	0,3 (8,4)	175	k. A. (k. A.)	k. A. (k. A.)	-2,4 (9,9)	0,29 [0,08; 0,50] 0,0066
UNIFI [90mg/8w]	176	k. A. (k. A.)	k. A. (k. A.)	0,3 (9,5)	175	k. A. (k. A.)	k. A. (k. A.)	-2,4 (9,9)	0,28 [0,07; 0,49] 0,0096

Analysepopulation: ITT

Imputation: LOCF für stetige Endpunkte, NRI für dichotome Endpunkte

Stratifizierungsvariablen: Bio-IR-Status, geografische Region

Angaben zum Hedges‘ g basieren auf einem ANCOVA-Modell mit dem Baseline-Wert und den Stratifizierungsvariablen als Kovariablen.

Ein Hedges‘ g größer Null stellt einen Vorteil von Ustekinumab dar.

ANCOVA: Kovarianzanalyse; ITT: Intention-to-Treat; k. A.: Keine Angabe; KI: Konfidenzintervall; LOCF: Last observation carried forward; LS: Least-Squares; MW: Mittelwert; N: Anzahl Patienten in der Analyse; NRI: Non-Responder-Imputation; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation); SE: Standardfehler; SF-36: Short Form-36

Hinsichtlich des SF-36 MCS zeigen sich sowohl für die Induktionsphase als auch für die Erhaltungsphase Vorteile unter einer Behandlung mit Ustekinumab als unter einer Behandlung mit Placebo.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-120: Ergebnisse für mittlere Veränderung des SF-36 MCS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo Hedges‘ g [95% KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**
mittlere Veränderung des SF-36 MCS	0,61 [0,47; 0,74] <0,0001	0,29 [0,14; 0,45] 0,0003	0,32 [0,11; 0,52] 0,0028

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein Hedges‘ g größer Null im indirekten Vergleich, stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; MCS: Mental Health Component Summary Score; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); SF-36: Short Form – 36

Bezüglich der mittleren Veränderung des SF-36 MCS zeigt sich für die Induktionsphase ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab (Hedges’ g [95 %-KI]; 0,32 [0,11; 0,52]; p-Wert = 0,0028).

Tabelle 4-121: Ergebnisse für mittlere Veränderung des SF-36 MCS – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo Hedges‘ g [95% KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo Hedges‘ g [95% KI] p-Wert**	**Upadacitinib vs. Ustekinumab Hedges‘ g [95% KI] p-Wert**
mittlere Veränderung des SF-36 MCS	0,38 [0,15; 0,61] 0,0013	0,29 [0,08; 0,50] 0,0066	0,08 [-0,23; 0,39] 0,6037	0,49 [0,26; 0,72] <0,0001	0,28 [0,07; 0,49] 0,0096	0,21 [-0,10; 0,52] 0,1797

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein Hedges‘ g größer Null im indirekten Vergleich, stellt einen Vorteil von Upadacitinib dar.

KI: Konfidenzintervall; MCS: Mental Health Component Summary Score; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); SF-36: Short Form – 36

Bezüglich der mittleren Veränderung des SF-36 MCS zeigt sich für die Erhaltungsphase ein numerischer Vorteil von Upadacitinib gegenüber Ustekinumab, sowohl für die Dosierung Upadacitinib 15 mg gegenüber Ustekinumab 90 mg/Q12W, als auch für die Dosierung Upadacitinib 30 mg gegenüber Ustekinumab 90 mg/Q8W.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.3.10 Verträglichkeit – indirekte Vergleiche aus RCT

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

Tabelle 4-122: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Upadacitinib	Placebo	Ustekinumab
Induktionsphase				
3	U-ACHIEVE Substudie 1 U-ACHIEVE Substudie 2 U-ACCOMPLISH	•	•	
1	UNIFI Induktionsstudie		•	•
Erhaltungsphase				
1	U-ACHIEVE Substudie 3	•	•	
1	UNIFI Erhaltungsstudie		•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

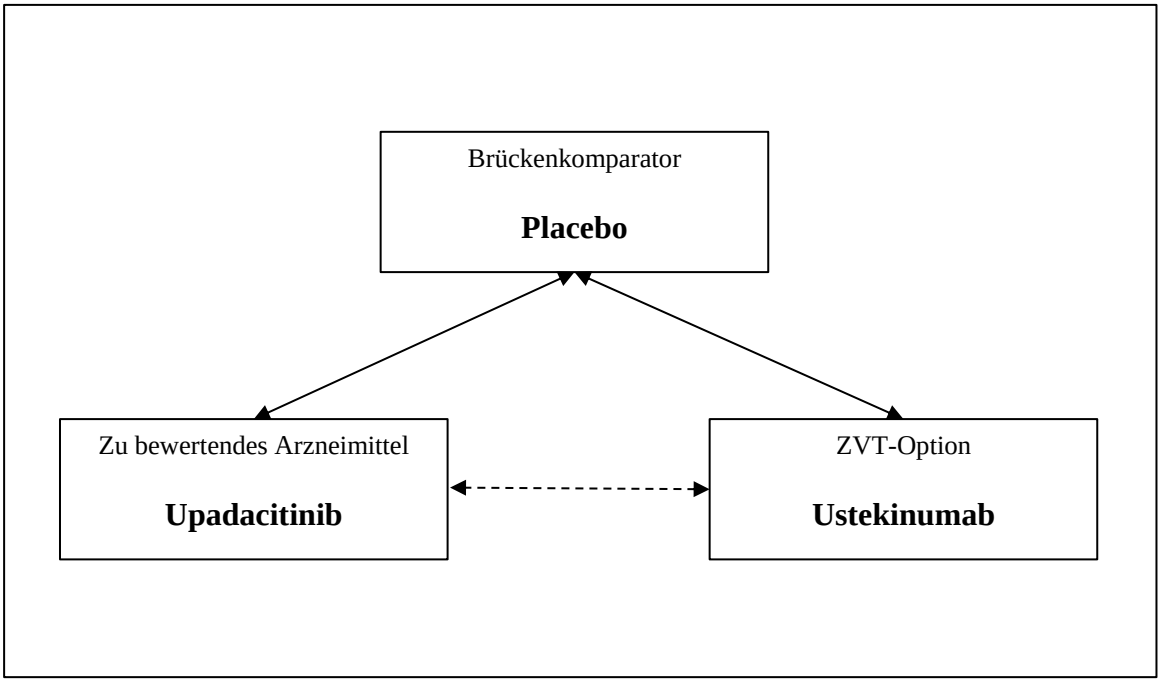

Abbildung 26: Netzwerkstruktur des indirekten Vergleichs

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

Tabelle 4-123: Operationalisierung von Verträglichkeit

Studie	Operationalisierung
Induktionsstudien	
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	Zur Beurteilung der Verträglichkeit werden für den indirekten Vergleich die Gesamtraten für folgende UE dargestellt: • Alle UE • Schwerwiegende UE (SUE) • UE, die zum Therapieabbruch führten • Todesfälle Die Kodierung der UE erfolgte gemäß MedDRA Version 23.0. UE wurden von Behandlungsbeginn bis Behandlungsende zuzüglich einer Nachbeobachtungsdauer von 30 Tagen erfasst. Für das vorliegende Dossier wird der Anteil der Patienten mit mindestens einem UE zu Woche 8 dargestellt. Die Analyse der UE basierte auf der Safety Population. Dieses umfasste alle Patienten, die mindestens eine Dosis der Studienmedikation erhielten gemäß der tatsächlich erhaltenen Studienmedikation. Für die Beschreibung der Induktionsphase wurden die U-ACHIEVE Substudie 1, die U-ACHIEVE Substudie 2 und die Studie U-ACCOMPLISH meta-analytisch zusammengefasst.
UNIFI Induktionsstudie	Zur Beurteilung der Verträglichkeit werden für den indirekten Vergleich die Gesamtraten für folgende UE dargestellt:

Studie	Operationalisierung
Induktionsstudien	
	• Alle UE • Schwerwiegende UE (SUE) • UE, die zum Therapieabbruch führten • Todesfälle Die Kodierung der UE erfolgte gemäß MedDRA Version 21.0. UE wurden von Behandlungsbeginn bis Behandlungsende zuzüglich einer Nachbeobachtungsdauer von bis zu 20 Wochen erfasst. Für das vorliegende Dossier wird der Anteil der Patienten mit mindestens einem UE zu Woche 8 dargestellt. Die Analyse der UE basierte auf der Safety Population. Dieses umfasste alle Patienten, die mindestens eine Dosis der Studienmedikation erhielten gemäß der tatsächlich erhaltenen Studienmedikation.
Erhaltungsstudien	
U-ACHIEVE Substudie 3	Zur Beurteilung der Verträglichkeit werden für den indirekten Vergleich die Gesamtraten für folgende UE dargestellt: • Alle UE • Schwerwiegende UE (SUE) • UE, die zum Therapieabbruch führten • Todesfälle Zusätzlich wurden ergänzend die Gesamtraten der schweren UE und der UE, die zum Tod führten, deskriptiv dargestellt. Ebenfalls ergänzend werden UE von speziellem Interesse (UESI) dargestellt. Folgende UESI waren im Prüfplan präspezifiziert: • Schwerwiegende Infektionen • Opportunistische Infektion (ohne Herpes Zoster und Tuberkulose) • Aktive Tuberkulose • Herpes Zoster • Neutropenie • Erhöhte Kreatinphosphokinase (CPK) • Malignität • Malignität außer NMSC • Nicht melanozytärer Hautkrebs (NMSC) • Lymphome • Nierendysfunktion • Hepatische Störung • Anämie • Lymphopenie • Gastrointestinale Perforation • Adjudizierte kardiovaskuläre Ereignisse • Adjudizierte Venöse thromboembolische Ereignisse (VTE) Die Kodierung der UE erfolgte gemäß MedDRA Version 23.0.

Studie	Operationalisierung
Induktionsstudien	
	UE wurden von Behandlungsbeginn bis Behandlungsende zuzüglich einer Nachbeobachtungsdauer von 30 Tagen erfasst. Für das vorliegende Dossier wird der Anteil der Patienten mit mindestens einem UE zu Woche 52 dargestellt. Die Analyse der UE basierte auf der Safety Population. Dieses umfasste alle Patienten, die mindestens eine Dosis der Studienmedikation erhielten gemäß der tatsächlich erhaltenen Studienmedikation und zum Ende der Induktionsphase ein Ansprechen auf Upadacitinib 45 mg hatten.
UNIFI Erhaltungsstudie	Zur Beurteilung der Verträglichkeit werden für den indirekten Vergleich die Gesamtraten für folgende UE dargestellt: • Alle UE • Schwerwiegende UE (SUE) • UE, die zum Therapieabbruch führten • Todesfälle Die Kodierung der UE erfolgte gemäß MedDRA Version 21.0. UE wurden von Behandlungsbeginn bis Behandlungsende zuzüglich einer Nachbeobachtungsdauer von bis zu 20 Wochen erfasst. Für das vorliegende Dossier wird der Anteil der Patienten mit mindestens einem UE zu Woche 44 dargestellt. Die Analyse der UE basierte auf der Safety Population. Dieses umfasste alle Patienten, die mindestens eine Dosis der Studienmedikation erhielten gemäß der tatsächlich erhaltenen Studienmedikation.
CPK: Kreatinphosphokinase; MedDRA: Medical Dictionary for Regulatory Activities; NMSC: Nicht melanozytärer Hautkrebs; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; UESI: UE von speziellem Interesse; VTE: Venöse thromboembolische Ereignisse	

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

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

Tabelle 4-124: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Induktionsstudien						
U-ACHIEVE Substudie 1 & U-ACHIEVE Substudie 2 & U-ACCOMPLISH	niedrig	ja	ja	ja	ja	niedrig
UNIFI Induktionsstudie	niedrig	ja	ja	ja	ja	niedrig
Erhaltungsstudien						
U-ACHIEVE Substudie 3	niedrig	ja	ja	ja	ja	niedrig
UNIFI Erhaltungsstudie	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung der Verträglichkeit erfolgte zu Woche 8 und zu Woche 44 bzw. 52. Aufgrund der einheitlichen Erfassung, Dokumentation und Berichterstattung von UE gemäß good clinical practice (GCP)-Standards erfolgt eine zusammenfassende Einstufung des Verzerrungspotenzials für alle UE.

Die Erhebung des Endpunkts erfolgte verblindet und für die Analysen wurde die präspezifizierte ITT-Population verwendet. Die Zuordnung der Patienten erfolgte für die Analyse entsprechend der Randomisierung. Das ITT-Prinzip wurde adäquat umgesetzt und es liegen keine Anhaltspunkte für eine ergebnisabhängige Berichterstattung vor. Da auch keine sonstigen Aspekte identifiziert wurden, die auf eine Verzerrung hindeuten, wird das Verzerrungspotenzial für den Endpunkt für die U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2, U-ACCOMPLISH, UNIFI Induktionsstudie, U-ACHIEVE Substudie 3 als auch für die UNIFI Erhaltungsstudie als niedrig eingestuft.

Gesamtraten

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

Tabelle 4-125: Ergebnisse für UE Gesamtraten – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Alle UE					
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	36/56 (64,3)	33/46 (71,7)	0,90 [0,69; 1,17] 0,4197	0,71 [0,31; 1,65] 0,4241	-0,07 [-0,26; 0,11] 0,4190
U-ACHIEVE Substudie 2 [45mg]	180/319 (56,4)	96/155 (61,9)	0,91 [0,78; 1,07] 0,2437	0,80 [0,54; 1,18] 0,2543	-0,06 [-0,15; 0,04] 0,2498
U-ACCOMPLISH [45mg]	182/344 (52,9)	70/177 (39,5)	1,34 [1,09; 1,65] 0,0060	1,72 [1,19; 2,48] 0,0040	0,13 [0,04; 0,22] 0,0034
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	398/719 (55,4)	199/378 (52,6)	1,06 [0,95; 1,19] 0,3099	1,14 [0,89; 1,47] 0,3038	0,03 [-0,03; 0,09] 0,2974
Heterogenitätstest (p-Wert)			0,0068		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	193/250 (77,2)	186/245 (75,9)	1,02 [0,92; 1,12] 0,7365	1,07 [0,71; 1,63] 0,7365	0,01 [-0,06; 0,09] 0,7365
U-ACHIEVE Substudie 3 [30mg]	198/251 (78,9)	186/245 (75,9)	1,04 [0,94; 1,14] 0,4303	1,19 [0,78; 1,81] 0,4299	0,03 [-0,04; 0,10] 0,4296
Schwerwiegende UE					
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	3/56 (5,4)	5/46 (10,9)	0,49 [0,12; 1,95] 0,3140	0,46 [0,10; 2,06] 0,3121	-0,06 [-0,16; 0,05] 0,3151

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
U-ACHIEVE Substudie 2 [45mg]	8/319 (2,5)	9/155 (5,8)	0,43 [0,17; 1,10] 0,0777	0,42 [0,16; 1,10] 0,0782	-0,03 [-0,07; 0,01] 0,1115
U-ACCOMPLISH [45mg]	11/344 (3,2)	8/177 (4,5)	0,71 [0,29; 1,73] 0,4473	0,70 [0,28; 1,77] 0,4480	-0,01 [-0,05; 0,02] 0,4693
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	22/719 (3,1)	22/378 (5,8)	0,55 [0,31; 0,98] 0,0417	0,53 [0,29; 0,98] 0,0417	-0,03 [-0,05; 0,00] 0,0562
Heterogenitätstest (p-Wert)			0,7454		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	21/250 (8,4)	23/245 (9,4)	0,89 [0,51; 1,57] 0,6996	0,89 [0,48; 1,64] 0,6996	-0,01 [-0,06; 0,04] 0,6995
U-ACHIEVE Substudie 3 [30mg]	21/251 (8,4)	23/245 (9,4)	0,89 [0,51; 1,57] 0,6894	0,88 [0,47; 1,64] 0,6894	-0,01 [-0,06; 0,04] 0,6893
UE, die zum Therapieabbruch führten					
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	5/56 (8,9)	4/46 (8,7)	1,03 [0,29; 3,60] 0,9671	1,03 [0,26; 4,08] 0,9671	0,00 [-0,11; 0,11] 0,9670
U-ACHIEVE Substudie 2 [45mg]	6/319 (1,9)	14/155 (9,0)	0,21 [0,08; 0,53] 0,0010	0,19 [0,07; 0,51] 0,0010	-0,07 [-0,12; -0,02] 0,0032
U-ACCOMPLISH [45mg]	6/344 (1,7)	9/177 (5,1)	0,34 [0,12; 0,95] 0,0392	0,33 [0,12; 0,95] 0,0391	-0,03 [-0,07; 0,00] 0,0629
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	17/719 (2,4)	27/378 (7,1)	0,36 [0,20; 0,63] 0,0005	0,34 [0,18; 0,63] 0,0006	-0,05 [-0,07; -0,02] 0,0016
Heterogenitätstest (p-Wert)			0,1355		

Studie	Upadacitinib	Placebo	Upadacitinib vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	10/250 (4,0)	25/245 (10,2)	0,39 [0,19; 0,80] 0,0099	0,37 [0,17; 0,78] 0,0093	-0,06 [-0,11; -0,02] 0,0069
U-ACHIEVE Substudie 3 [30mg]	18/251 (7,2)	25/245 (10,2)	0,70 [0,39; 1,25] 0,2331	0,68 [0,36; 1,28] 0,2323	-0,03 [-0,08; 0,02] 0,2303
Tod					
Induktionsphase					
U-ACHIEVE Substudie 1 [45mg]	0/56 (0,0)	0/46 (0,0)	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	0,00 [-0,04; 0,04] 1,0000
U-ACHIEVE Substudie 2 [45mg]	0/319 (0,0)	0/155 (0,0)	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	0,00 [-0,01; 0,01] 1,0000
U-ACCOMPLISH [45mg]	0/344 (0,0)	0/177 (0,0)	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	0,00 [-0,01; 0,01] 1,0000
Meta-Analyse U-ACHIEVE Substudie 1 + U-ACHIEVE Substudie 2 + U-ACCOMPLISH	0/719 (0,0)	0/378 (0,0)	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	0,00 [-0,01; 0,01] 1,0000
Heterogenitätstest (p-Wert)			k. A.		
Erhaltungsphase					
U-ACHIEVE Substudie 3 [15mg]	0/250 (0,0)	0/245 (0,0)	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	0,00 [-0,01; 0,01] 1,0000
U-ACHIEVE Substudie 3 [30mg]	0/251 (0,0)	0/245 (0,0)	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	0,00 [-0,01; 0,01] 1,0000

Analysepopulation: Safety Analysis Set

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5. Für die Meta-Analyse wurde zusätzlich Studie als fester Effekt ins Modell eingeschlossen.

Ein RR und OR kleiner Eins bzw. eine RD kleiner Null stellen einen Vorteil von Upadacitinib dar.

k. A.: Keine Angabe; KI: Konfidenzintervall; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis

Hinsichtlich der Gesamtraten zeigen sich bei UE jeglichen Schweregrades, sowohl für die Induktionsphase als auch für die Erhaltungsphase, keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen. Bei den SUE zeigen sich in der Induktionsphase statistisch signifikante Vorteile unter einer Behandlung mit Upadacitinib gegenüber einer Behandlung mit Placebo. Bei den UE, die zum Therapieabbruch führten, zeigen sich statistisch signifikante Vorteile, sowohl für die Induktionsphase als auch für die Erhaltungsphase, unter einer Behandlung mit Upadacitinib gegenüber einer Behandlung mit Placebo. In beiden Behandlungsgruppen traten keine Todesfälle auf.

Wie in Abschnitt 4.2.5.3 beschrieben, werden zur Visualisierung die Daten des Endpunkts UE Gesamtraten für die Studien U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH in Form eines Forest-Plots dargestellt.

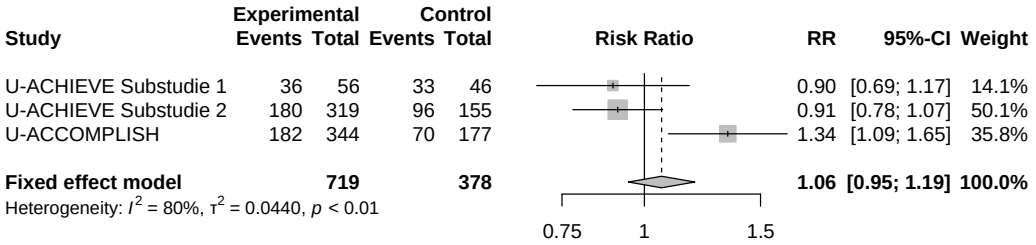

Abbildung 27: Forest-Plot für Alle UE zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

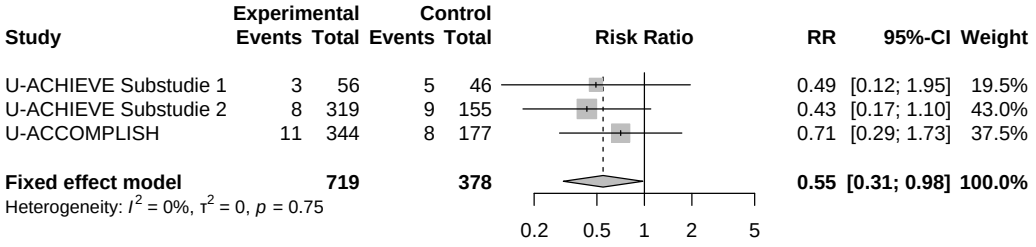

Abbildung 28: Forest-Plot für Schwerwiegende UE zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

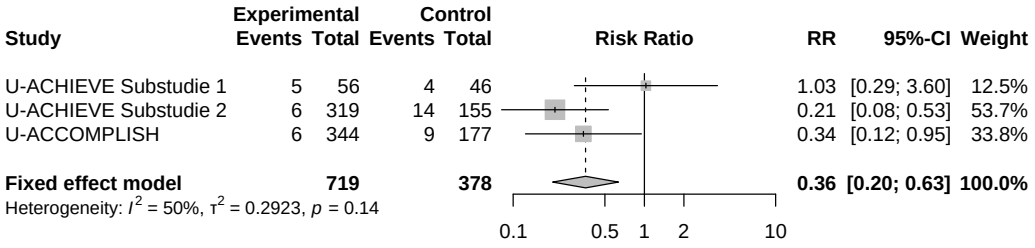

Abbildung 29: Forest-Plot für UE, die zum Therapieabbruch führten zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

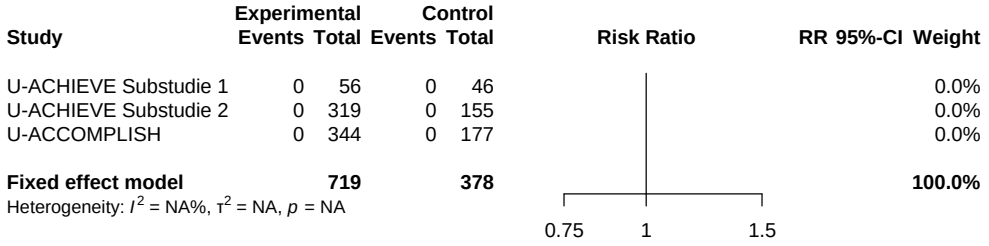

Abbildung 30: Forest-Plot für Todesfälle zu Woche 8 aus RCT (U-ACHIEVE Substudie 1, U-ACHIEVE Substudie 2 und U-ACCOMPLISH); Upadacitinib vs. Placebo

Tabelle 4-126: Ergebnisse für UE Gesamtraten – Einzelstudie Ustekinumab – aus RCT für indirekte Vergleiche

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
Alle UE					
Induktionsphase					
UNIFI [6mg/kg]	162/320 (50,6)	153/319 (48,0)	1,06 [0,90; 1,24] 0,5011	1,11 [0,82; 1,52] 0,5009	0,03 [-0,05; 0,10] 0,5007
Erhaltungsphase					
UNIFI [90mg/12w]	119/172 (69,2)	138/175 (78,9)	0,88 [0,77; 1,00] 0,0415	0,60 [0,37; 0,98] 0,0408	-0,10 [-0,19; 0,00] 0,0389

Studie	Ustekinumab	Placebo	Ustekinumab vs. Placebo		
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	RR [95%-KI] p-Wert	OR [95%-KI] p-Wert	RD [95%-KI] p-Wert
UNIFI [90mg/8w]	136/176 (77,3)	138/175 (78,9)	0,98 [0,88; 1,09] 0,7199	0,91 [0,55; 1,51] 0,7199	-0,02 [-0,10; 0,07] 0,7198
Schwerwiegende UE					
Induktionsphase					
UNIFI [6mg/kg]	11/320 (3,4)	22/319 (6,9)	0,50 [0,25; 1,01] 0,0536	0,48 [0,23; 1,01] 0,0526	-0,03 [-0,07; 0,00] 0,0476
Erhaltungsphase					
UNIFI [90mg/12w]	13/172 (7,6)	17/175 (9,7)	0,78 [0,39; 1,55] 0,4764	0,76 [0,36; 1,62] 0,4759	-0,02 [-0,08; 0,04] 0,4741
UNIFI [90mg/8w]	15/176 (8,5)	17/175 (9,7)	0,88 [0,45; 1,70] 0,6984	0,87 [0,42; 1,79] 0,6984	-0,01 [-0,07; 0,05] 0,6982
UE, die zum Therapieabbruch führten					
Erhaltungsphase					
UNIFI [90mg/12w]	9/172 (5,2)	20/175 (11,4)	0,46 [0,21; 0,98] 0,0434	0,43 [0,19; 0,97] 0,0417	-0,06 [-0,12; 0,00] 0,0353
UNIFI [90mg/8w]	5/176 (2,8)	20/175 (11,4)	0,25 [0,10; 0,65] 0,0044	0,23 [0,08; 0,62] 0,0037	-0,09 [-0,14; -0,03] 0,0015
Tod					
Induktionsphase					
UNIFI [6mg/kg]	1/320 (0,3)	0/319 (0,0)	2,99 [0,12; 73,14] 0,5018	3,00 [0,12; 73,92] 0,5016	0,00 [-0,01; 0,01] 0,4776
Erhaltungsphase					
UNIFI [90mg/12w]	0/172 (0,0)	0/175 (0,0)	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	0,00 [-0,01; 0,01] 1,0000
UNIFI [90mg/8w]	0/176 (0,0)	0/175 (0,0)	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	0,00 [-0,01; 0,01] 1,0000

Analysepopulation: Safety Analysis Set

RR, OR, RD, sowie der zugehörige Heterogenitätstest, basieren auf einer nicht adjustierten Schätzung nach der Mantel-Haenszel-Methode mit einer Nullzellenkorrektur von 0,5.

Ein RR und OR kleiner Eins bzw. eine RD kleiner Null stellen einen Vorteil von Ustekinumab dar.

k. A.: Keine Angabe; KI: Konfidenzintervall; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; OR: Odds-Ratio; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis

Hinsichtlich der Gesamtraten zeigen sich bei UE jeglichen Schweregrades und bei den UE, die zum Therapieabbruch führten, für die Erhaltungsphase statistisch signifikante Vorteile unter einer Behandlung mit Ustekinumab gegenüber einer Behandlung mit Placebo. In beiden Behandlungsgruppen traten keine Todesfälle auf.

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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Tabelle 4-127: Ergebnisse für UE Gesamtraten - Upadacitinib gegenüber Ustekinumab - aus RCT für indirekte Vergleiche Induktionsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 45 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Alle UE	1,06 [0,95; 1,19] 0,3099	1,06 [0,90; 1,24] 0,5011	1,01 [0,83; 1,22] 0,9517
Schwerwiegende UE	0,55 [0,31; 0,98] 0,0417	0,50 [0,25; 1,01] 0,0536	1,10 [0,44; 2,74] 0,8417
UE, die zum Therapieabbruch führten	0,36 [0,20; 0,63] 0,0005	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.
Tod	k. A. [k. A.; k. A.] k. A.	2,99 [0,12; 73,14] 0,5018	k. A. [k. A.; k. A.] k. A.

	Upadacitinib	Ustekinumab	Indirekter Vergleich
	Upadacitinib 45 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 6 mg/kg vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo. Ein RR im indirekten Vergleich kleiner Eins stellt einen Vorteil von Upadacitinib dar. k. A.: Keine Angabe; KI: Konfidenzintervall; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko; UE: Unerwünschtes Ereignis			

Bezüglich der Gesamtraten zeigen sich keine Unterschiede zwischen einer Behandlung mit Upadacitinib und einer Behandlung mit Ustekinumab.

Tabelle 4-128: Ergebnisse für UE Gesamtraten – Upadacitinib gegenüber Ustekinumab – aus RCT für indirekte Vergleiche Erhaltungsphase

	Upadacitinib	**Ustekinumab**	**Indirekter Vergleich**	**Upadacitinib**	**Ustekinumab**	**Indirekter Vergleich**
	Upadacitinib 15 mg vs. Placebo RR [95%-KI] p-Wert	**Ustekinumab 90mg Q12W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**	**Upadacitinib 30 mg vs. Placebo RR [95%-KI] p-Wert**	**Ustekinumab 90mg Q8W vs. Placebo RR [95%-KI] p-Wert**	**Upadacitinib vs. Ustekinumab RR [95%-KI] p-Wert**
Alle UE	1,02 [0,92; 1,12] 0,7365	0,88 [0,77; 1,00] 0,0415	1,16 [0,99; 1,36] 0,0692	1,04 [0,94; 1,14] 0,4303	0,98 [0,88; 1,09] 0,7199	1,06 [0,92; 1,23] 0,4319
Schwerwiegende UE	0,89 [0,51; 1,57] 0,6996	0,78 [0,39; 1,55] 0,4764	1,15 [0,47; 2,81] 0,7588	0,89 [0,51; 1,57] 0,6894	0,88 [0,45; 1,70] 0,6984	1,02 [0,43; 2,43] 0,9718
UE, die zum Therapieabbruch führten	0,39 [0,19; 0,80] 0,0099	0,46 [0,21; 0,98] 0,0434	0,86 [0,30; 2,42] 0,7698	0,70 [0,39; 1,25] 0,2331	0,25 [0,10; 0,65] 0,0044	2,83 [0,92; 8,66] 0,0688
Tod	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.	k. A. [k. A.; k. A.] k. A.

Indirekter Vergleich nach Bucher et al. zwischen Upadacitinib und Ustekinumab über den gemeinsamen Brückenkomparator Placebo.

Ein RR im indirekten Vergleich kleiner Eins stellt einen Vorteil von Upadacitinib dar.

k. A.: Keine Angabe; KI: Konfidenzintervall; Q12W: Alle 12 Wochen; Q8W: Alle 8 Wochen; RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial); RR: Relatives Risiko; UE: Unerwünschtes Ereignis

Bezüglich der Gesamtraten zeigen sich keine Unterschiede zwischen einer Behandlung mit Upadacitinib und einer Behandlung mit Ustekinumab.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

UESI

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

Um ein vollständiges Bild des Verträglichkeitsprofils von Upadacitinib abzubilden werden im Folgenden für die Erhaltungsphase die Gesamtraten für Schwere UE und UE, die zum Tod führten ergänzend dargestellt. Zusätzlich werden, ebenfalls für die Erhaltungsphase, die im Prüfplan präspezifizierten UESI ergänzend berichtet.

Tabelle 4-129: Weitere Ergebnisse für Verträglichkeit – Einzelstudie Upadacitinib – aus RCT für indirekte Vergleiche Erhaltungsphase

Endpunkt	Upadacitinib [15mg]	Upadacitinib [30mg]	Placebo
	n/N (%)	n/N (%)	n/N (%)
UE Gesamtraten ergänzend			
Schwere UE	14/250 (5,6)	20/251 (8,0)	23/245 (9,4)
UE, die zum Tod führten	0/250 (0,0)	0/251 (0,0)	0/245 (0,0)
UESI			
Schwerwiegende Infektionen	9/250 (3,6)	7/251 (2,8)	8/245 (3,3)
Opportunistische Infektion (ohne Herpes Zoster und Tuberkulose)	2/250 (0,8)	2/251 (0,8)	2/245 (0,8)
Aktive Tuberkulose	0/250 (0,0)	0/251 (0,0)	0/245 (0,0)
Herpes Zoster	12/250 (4,8)	14/251 (5,6)	0/245 (0,0)
Neutropenie	7/250 (2,8)	15/251 (6,0)	5/245 (2,0)
Erhöhte Kreatinphosphokinase (CPK)	15/250 (6,0)	19/251 (7,6)	5/245 (2,0)
Mögliche Malignität (alle Typen)	1/250 (0,4)	5/251 (2,0)	2/245 (0,8)
Malignität	1/250 (0,4)	5/251 (2,0)	1/245 (0,4)
Malignität außer NMSC	1/250 (0,4)	2/251 (0,8)	1/245 (0,4)

Endpunkt	Upadacitinib [15mg]	Upadacitinib [30mg]	Placebo
	n/N (%)	n/N (%)	n/N (%)
Nicht melanozytärer Hautkrebs (NMSC)	0/250 (0,0)	3/251 (1,2)	0/245 (0,0)
Lymphome	0/250 (0,0)	0/251 (0,0)	0/245 (0,0)
Nierendysfunktion	1/250 (0,4)	1/251 (0,4)	1/245 (0,4)
Hepatische Störung	18/250 (7,2)	13/251 (5,2)	6/245 (2,4)
Anämie	11/250 (4,4)	9/251 (3,6)	15/245 (6,1)
Lymphopenie	7/250 (2,8)	5/251 (2,0)	4/245 (1,6)
Gastrointestinale Perforation	0/250 (0,0)	0/251 (0,0)	1/245 (0,4)
Adjudizierte kardiovaskuläre Ereignisse	0/250 (0,0)	1/251 (0,4)	1/245 (0,4)
Adjudizierte Venöse thromboembolische Ereignisse (VTE)	2/250 (0,8)	2/251 (0,8)	0/245 (0,0)

Analysepopulation: Safety Analysis Set

CPK: Kreatinphosphokinase; n: Anzahl Patienten mit Response; N: Anzahl Patienten in der Analyse; NMSC: Nicht melanozytärer Hautkrebs (Non Melanoma Skin Cancer); RCT: Randomisierte kontrollierte Studie (Randomized Controlled Trial) UE: Unerwünschtes Ereignis; VTE: Venöse thromboembolische Ereignisse

Die Gesamtraten der schweren UE waren unter einer Behandlung mit Upadacitinib für beide Dosierungen niedriger als unter einer Behandlung mit Placebo.

Der Anteil der Patienten mit UE war im Allgemeinen ähnlich in den Upadacitinib- und Placebo-Gruppen. Insgesamt werden die beschriebenen Unterschiede als geringfügig eingestuft. Das Sicherheitsprofil von Upadacitinib entspricht dem bekannten und gut charakterisierten Sicherheitsprofil von Upadacitinib ohne neue unerwartete UE.

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.2.1.2.1 dargestellt.

4.3.2.1.4 Subgruppenanalysen – indirekte Vergleiche aus RCT

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

Aus den im Abschnitt 4.2.5.5 formulierten Gründen werden keine Subgruppenergebnisse dargestellt.

Allerdings werden, soweit verfügbar, deskriptive Subgruppenanalysen für die Subgruppe Bio-IR-Status präsentiert, um eine Einschätzung der Ergebnisse hinsichtlich bio-naiven und bio-erfahrenen Patienten zu ermöglichen. Die Darstellung der Subgruppenergebnisse erfolgt nur für die Erhaltungsphase.

4.3.2.1.4.1 Klinische Remission – Subgruppen

Im Folgenden werden die deskriptiven Subgruppenergebnisse für die klinische Remission berichtet.

Tabelle 4-130: Subgruppenergebnisse für Anteil der Patienten mit klinischer Remission (Adaptierter Mayo Score) aus RCT (U-ACHIEVE Substudie 3) mit dem zu bewertenden Arzneimittel

Endpunkt	Subgruppe	Upadacitinib	Placebo
		Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
Erhaltungsphase			
U-ACHIEVE Substudie 3 [15 mg]	Bio-IR	29/71 (40,8)	6/81 (7,4)
	Non-Bio-IR	34/77 (44,1)	12/68 (17,6)
U-ACHIEVE Substudie 3 [30 mg]	Bio-IR	36/73 (49,3)	6/81 (7,4)
	Non-Bio-IR	44/81 (54,3)	12/68 (17,6)

Analysepopulation: ITT

Imputationsmethode: NRI-C

Bio-IR: Intoleranz oder unzureichendes Ansprechen auf eine vorherige biologische Therapie; ITT: Intention-to-Treat; MI: Multiple Imputation; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Response; NRI: Non-Responder-Imputation; NRI-C: NRI incorporating MI to handle missing data due to COVID-19; RCT: Randomisierte kontrollierte Studie

Bezüglich der Subgruppenanalyse zeigten sich für die klinische Remission mittels Adaptiertem Mayo Score in der Erhaltungsphase vergleichbare Ansprechraten zwischen Bio-IR- und

Non-Bio-IR-Patienten (Upadacitinib 15 mg: 40,8 % im Vergleich zu 44,1 %; Upadacitinib 30 mg: 49,3 % im Vergleich zu 54,3 %). Aufgrund der vergleichbaren Ansprechraten erscheint eine gemeinsame Betrachtung im Rahmen der Zielpopulation als adäquat.

4.3.2.1.4.2 Mukosaheilung - Subgruppen

Im Folgenden werden die deskriptiven Subgruppenergebnisse für die Mukosaheilung berichtet.

Tabelle 4-131: Subgruppenergebnisse für Anteil der Patienten mit Mukosaheilung aus RCT (U-ACHIEVE Substudie 3) mit dem zu bewertenden Arzneimittel

Endpunkt	Subgruppe	Upadacitinib	Placebo
		Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
Erhaltungsphase			
U-ACHIEVE Substudie 3 [15 mg]	Bio-IR	23/71 (32,4)	4/81 (4,9)
	Non-Bio-IR	28/77 (36,4)	14/68 (20,6)
U-ACHIEVE Substudie 3 [30 mg]	Bio-IR	35/73 (47,9)	4/81 (4,9)
	Non-Bio-IR	42/81 (51,9)	14/68 (20,6)
Analysepopulation: ITT Imputationsmethode: NRI-C Bio-IR: Intoleranz oder unzureichendes Ansprechen auf eine vorherige biologische Therapie; ITT: Intention-to-Treat; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Response; RCT: Randomisierte kontrollierte Studie			

Bezüglich der Subgruppenanalyse zeigten sich für die Mukosaheilung in der Erhaltungsphase vergleichbare Ansprechraten zwischen Bio-IR- und Non-Bio-IR-Patienten (Upadacitinib 15 mg: 32,4 % im Vergleich zu 36,4 %; Upadacitinib 30 mg: 47,9 % im Vergleich zu 51,9 %). Aufgrund der vergleichbaren Ansprechraten erscheint eine gemeinsame Betrachtung im Rahmen der Zielpopulation als adäquat.

4.3.2.1.5 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

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).

Studie	Quelle
Upadacitinib	
U-ACCOMPLISH (M14-675)	Studienbericht, inklusive Studienprotokoll und statistischen Analyseplans (82) Studienregistereinträge (71-74)
U-ACHIEVE (M14-234)	Studienbericht, inklusive Studienprotokoll und statistischen Analyseplans (83) Studienregistereinträge (75-77) Publikationen (66, 84, 85)
Ustekinumab	
UNIFI	Publikationen (inkl. Studienprotokoll und statistischem Analyseplan) (35) Studienregistereinträge (78-80)

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/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

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

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

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

Tabelle 4-132: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		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 zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus nicht randomisierten vergleichenden Studien beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-133: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

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

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-134: Operationalisierung von <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 4-135: 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 zutreffend.

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 zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende 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).

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*

- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

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

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus weiteren Untersuchungen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-136: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

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

Nicht zutreffend.

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 4-137: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. 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.

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

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

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen

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).

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

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

Zum Nachweis des Nutzens und Zusatznutzens von Upadacitinib im oben genannten AWG liegen keine direkt zur ZVT vergleichenden klinischen Studien vor. Für die Bewertung des medizinischen Nutzens und medizinischen Zusatznutzens von Upadacitinib im AWG wird daher ein adjustierter, indirekter Vergleich nach Bucher durchgeführt und vorgelegt.

Zur Beurteilung des Zusatznutzens von Upadacitinib wurden vom G-BA zwei Fragestellungen (Patienten mit oder ohne unzureichendes Ansprechen auf eine biologische Therapie: Bio-IR, Non-Bio-IR) formuliert. Im Rahmen des indirekten Vergleichs ist eine getrennte Beantwortung dieser Fragestellungen nicht möglich, da die Daten nicht für alle Endpunktdimensionen bzw. patientenrelevante Endpunkte, insbesondere nicht für die Verträglichkeit, getrennt nach Vorbehandlung vorliegen. Analysen zur Zielpopulation sind aber im Rahmen des indirekten Vergleichs adäquat, da ein gemeinsames Therapieziel zu Grunde liegt und für beide Fragestellungen Ustekinumab als ZVT durch den G-BA definiert wurde (2). Darüber hinaus zeigen sich bei einer Behandlung unter Upadacitinib für die Erhaltungsphase, vergleichbare Ansprechraten für Bio-IR- und Non-Bio-IR-Patienten. Daher werden beide Fragestellungen zu einer Fragestellung zusammengefasst und Ergebnisse für die Zielpopulation dargestellt.

Evidenzstufe

Upadacitinib

Für das zu bewertende Arzneimittel Upadacitinib liegt für die zur Ableitung des Zusatznutzens in der Indikation CU relevante Erhaltungsphase die Studie U-ACHIEVE (M14-234) vor. Hierbei handelt es sich um eine multizentrische, randomisierte, doppelt-verblindete, placebokontrollierte Phase-III-Studie der Evidenzstufe Ib (nach §5 5. Kapitel, 1. Abschnitt, Absatz 6 der VerfO des G-BA). Die Studie U-ACHIEVE umfasst drei Substudien: eine Phase-IIb-Dosierungsstudie für die Induktionsphase (Substudie 1), eine Phase-III-Induktionsstudie (Substudie 2) und eine Phase-III-Erhaltungsstudie (Substudie 3). Die Studie U-ACCOMPLISH ist eine weitere Phase-III-Induktionsstudie.

Die Induktionsstudien werden als ergänzende Evidenz zur Rekapitulation des deutschen Versorgungskontext und damit einhergehend zur Ableitung des Zusatznutzen dargestellt.

Es liegen keine direkt zur ZVT vergleichenden Studien vor. Um Aussagen über den Nutzen bzw. Zusatznutzen von Upadacitinib treffen zu können, wurde ein indirekter Vergleich nach Bucher mit der ZVT Ustekinumab über den Brückenkomparator Placebo durchgeführt.

Ustekinumab

Bei der Studie UNIFI handelt es sich um eine randomisierte, doppelblinde, multizentrische, placebokontrollierte Phase-III-Studie, die die Sicherheit und Wirksamkeit von Ustekinumab in der Induktions- und Erhaltungsphase bei Patienten mit CU untersucht. Die Studie entspricht der Evidenzstufe Ib (nach §5 5. Kapitel, 1. Abschnitt, Absatz 6 der VerfO des G-BA) und wird für den adjustierten indirekten Vergleich über den Brückenkomparator Placebo herangezogen.

Studienqualität

Upadacitinib

Die Studie U-ACHIEVE wurden bei erwachsenen Patienten mit mittelschwerer bis schwerer aktiver CU, die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben durchgeführt.

Insgesamt wurden 847 Patienten in der Erhaltungsphase der Studie U-ACHIEVE mittels einer computergenerierten Randomisierungssequenz in einem Verhältnis von 1:1:1 auf die drei Behandlungsarme randomisiert (Upadacitinib 15 mg täglich, Upadacitinib 30 mg täglich oder Placebo). Die Zuteilung erfolgte jeweils durch ein IRT, wodurch die Verdeckung der Gruppenzuteilung gegeben war. Die Randomisierung erfolgte stratifiziert nach dem Bio-IR-Status (Bio-IR oder nicht Bio-IR) zu Baseline der Induktionsstudien, Vorhandensein einer klinischen Remission zu Beginn der Erhaltungsphase (ja, nein) und Einnahme von Kortikosteroiden zu Beginn der Erhaltungsphase (ja, nein).

Sowohl die Patienten als auch die behandelnden Personen waren über die gesamte bewertungsrelevante Erhaltungsphase verblindet. Endpunkte wurden, soweit im Rahmen des indirekten Vergleichs möglich, gemäß dem SAP berichtet. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet und entsprechend in der Operationalisierung beschrieben und somit nicht ergebnisgesteuert. Daher kann insgesamt von einer ergebnisunabhängigen Berichterstattung ausgegangen werden. Es lagen auch keine sonstigen Aspekte vor, die auf ein Verzerrungspotenzial auf Studienebene hindeuten. Daher wird das Verzerrungspotenzial für die Erhaltungsphase der Studie U-ACHIEVE als niedrig bewertet.

Zum Nachweis des Zusatznutzens von Upadacitinib im vorliegenden AWG wurde ein adjustierter indirekter Vergleich der Studie U-ACHIEVE mit der Studie UNIFI durchgeführt. Dabei wurde das Verfahren nach Bucher verwendet (68). Als Brückenkomparator wurde der Placebo-Arm gewählt, da sowohl die Studie U-ACHIEVE als auch die Studie UNIFI placebokontrolliert sind.

Die herangezogene klinische Studie wurden anhand des CONSORT-Statements (Anhang 4-E) und der Cochrane-Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse (Anhang 4-F) auf die Planungs-, Durchführungs- und Auswertungsqualität überprüft.

Ustekinumab

Die Studie UNIFI wurde bei erwachsenen Patienten mit mittelschwerer bis schwerer aktiver CU, welche auf eine vorangegangene Therapie unzureichend angesprochen haben, nicht mehr darauf ansprechen oder eine Unverträglichkeit oder Kontraindikation aufweisen, durchgeführt.

Insgesamt wurden 523 Patienten in der Studie UNIFI in der Erhaltungsphase mittels einer computergenerierten Randomisierungssequenz in einem Verhältnis von 1:1:1 auf die drei Behandlungsarme randomisiert (Ustekinumab 90 mg alle 12 Wochen, Ustekinumab 90 mg alle 8 Wochen oder Placebo). Die Zuteilung erfolgte jeweils durch ein IRT, wodurch die Verdeckung der Gruppenzuteilung gegeben war. Die Randomisierung erfolgte stratifiziert nach der Vorbehandlung mit einer biologischen Therapie (ja, nein), und der geografischen Herkunft (Osteuropa, Asien, oder der Rest der Welt).

Sowohl die Patienten als auch die behandelnden Personen waren über die gesamte bewertungsrelevante Erhaltungsphase verblindet. Endpunkte wurden, soweit im Rahmen des indirekten Vergleichs möglich, gemäß dem SAP berichtet. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet und entsprechend in der Operationalisierung beschrieben und somit nicht ergebnisgesteuert. Daher kann insgesamt von einer ergebnisunabhängigen Berichterstattung ausgegangen werden. Es lagen auch keine sonstigen Aspekte vor, die auf ein Verzerrungspotenzial auf Studienebene hindeuten. Daher wird das Verzerrungspotenzial für die Erhaltungsphase der Studie UNIFI als niedrig bewertet.

Die herangezogene klinische Studie wurden anhand des CONSORT-Statements (Anhang 4-E) und der Cochrane-Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse (Anhang 4-F) auf die Planungs-, Durchführungs- und Auswertungsqualität überprüft.

Fazit zum adjustierten indirekten Vergleich

Die Studien U-ACHIEVE und UNIFI sind hinsichtlich demografischer und krankheitsspezifischer Charakteristika weitgehend homogen (Abschnitt 4.3.2.1.2). Bei den Studien handelt es sich um randomisierte, kontrollierte, doppelt-verblindete Phase-III-Studien mit parallelen Behandlungsarmen. In beide Studien wurden erwachsene Patienten mit mittelschwerer bis schwerer aktiver CU eingeschlossen, welche auf eine vorangegangene Therapie unzureichend angesprochen haben, nicht mehr darauf ansprechen oder eine Unverträglichkeit oder Kontraindikation aufweisen. Die Studien untersuchen die Effekte in der Erhaltungsphase bei Patienten, die während der Induktionsstudie ein klinisches Ansprechen auf die entsprechende Behandlung gezeigt haben.

Aufgrund des konsequent umgesetzten Studiendesigns (Randomisierung, verdeckte Gruppenzuteilung, doppelte Verblindung, placebokontrolliert), sowie aufgrund der umfassenden und transparenten Auswertung und Berichterstattung, ist das Verzerrungspotenzial der Studien auf Studienebene zusammenfassend als niedrig zu bewerten. Die Aussagekraft der Nachweise für einen Zusatznutzen der Studien ist zusammenfassend als hoch anzusehen.

Validität der herangezogenen Endpunkte

Für die Bewertung des Zusatznutzens wurden Daten aus den Kategorien Mortalität, Morbidität, Lebensqualität und Verträglichkeit berücksichtigt. Für die Operationalisierung der entsprechenden Kategorien wurden Daten zu patientenrelevanten Endpunkten verwendet.

Im vorliegenden Dossier wird jeweils, sofern nicht anders angegeben, der letzte verfügbare Analysezeitpunkt (primär Woche 52 in der Studie U-ACHIEVE bzw. Woche 44 in der Studie UNIFI) in der Erhaltungsphase indirekt verglichen und berichtet. Die untersuchten Endpunkte wurden mittels validierter und krankheitsspezifischer Instrumente erhoben und gemäß der im Studienprotokoll festgelegten Vorgehensweise ausgewertet (ITT-Prinzip). Das Verzerrungspotenzial auf Endpunktebene für die Endpunkte Mortalität, Morbidität, Lebensqualität und Verträglichkeit wird als niedrig verzerrt bewertet.

Ableitung der Ergebnissicherheit

Die für den adjustierten indirekten Vergleich nach Bucher herangezogenen Studien U-ACHIEVE und UNIFI sind RCTs und entsprechen somit der Evidenzstufe Ib; es liegen keine bekannten verzerrenden Aspekte vor. Das Studiendesign ist grundsätzlich vergleichbar, entspricht dem gängigen Vorgehen im AWG und den Anforderungen der Zulassungsbehörden. Bei den Studienpopulationen besteht weitgehend Strukturgleichheit, v. a. bezüglich klinischer Parameter. Die im adjustierten indirekten Vergleich beobachteten Effekte sind daher auch im Behandlungsalltag in Deutschland zu erwarten. Da ein adjustierter indirekter Vergleich jedoch eine niedrige Ergebnissicherheit gewährleistet, wird diese auf **Anhaltspunkt** festgesetzt.

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.

Hintergrund

Upadacitinib wird angewendet zur Behandlung der mittelschweren bis schweren aktiven Colitis ulcerosa bei erwachsenen Patienten, die auf eine konventionelle Therapie oder ein Biologikum unzureichend angesprochen haben, nicht mehr darauf ansprechen oder diese nicht vertragen haben.

Der Leidensdruck der Patienten in dieser Zielpopulation ist sehr groß. Insbesondere aufgrund des jungen Diagnosealters, des blutigen, imperativen Stuhlgangs und des kontinuierlichen Wiederaufflammens der Hauptsymptomatik nach Phasen der Remission, geht die CU mit dramatischen physischen und psychischen Belastungen einher (5, 16, 86). Dies hat eine extreme Beeinträchtigung der Lebensqualität der Patienten zur Folge, die zu Schwierigkeiten im sozialen Umfeld führen (16, 86).

Als primäres Ziel der Behandlung von Patienten mit CU formuliert die aktuelle deutsche S3-Leitlinie das rasche Erreichen einer klinischen Remission und den Erhalt einer langfristigen

steroidfreien klinischen und endoskopischen Remission und somit das vollständige Abklingen der mukosalen Entzündung (13, 14). Das Erreichen einer klinischen Remission hat, durch die Kontrolle der Kranheitssymptome, eine enorme Verbesserung der gesundheitsbezogenen Lebensqualität und sozialen Teilhabe dieser Patienten zur Folge (12, 23, 27, 87). Bisher ist keine Behandlungsoption verfügbar, mit der die Therapieziele einer Remission bei allen Patienten dauerhaft erreicht werden könnte. Es besteht daher weiterhin ein hoher therapeutischer Bedarf an alternativen Behandlungsmöglichkeiten, die eine Langzeittherapie bei stabiler Krankheitskontrolle inklusive Mukosaheilung und einem gut charakterisierten Sicherheitsprofil ermöglichen.

Upadacitinib ist ein neuer selektiver und reversibler JAK-Inhibitor, dessen Wirksamkeits- und Sicherheitsprofil in einem umfangreichen klinischen Studienprogramm untersucht wurde. Dabei zeigt Upadacitinib über alle Endpunkte hinweg konsistent gute Wirksamkeit und ein gut charakterisiertes Sicherheitsprofil. Für die Nutzenbewertung steht mit einem adjustierten indirekten Vergleich nach Bucher auf Basis von RCT über den Brückenkomparator Placebo zur ZVT Ustekinumab die best verfügbare Evidenz zur Verfügung.

Ableitung des Zusatznutzens von Upadacitinib gegenüber der ZVT Ustekinumab

Zur Beurteilung des Zusatznutzens von Upadacitinib wurden vom G-BA zwei Fragestellungen formuliert. Im Rahmen des indirekten Vergleichs ist eine getrennte Beantwortung dieser Fragestellungen nicht möglich, da die Daten nicht für alle Endpunktdimensionen bzw. patientenrelevante Endpunkte, insbesondere nicht für die Verträglichkeit, getrennt nach Vorbehandlung vorliegen. Analysen zur Zielpoulation sind aber im Rahmen des indirekten Vergleichs adäquat, da beiden Fragestellungen mit der klinischen Remission ein gemeinsames Therapieziel zugrunde liegt, und Ustekinumab durch den G-BA für beide Fragestellungen als ZVT definiert wurde (2). Darüber hinaus zeigen sich bei einer Behandlung unter Upadacitinib für die Erhaltungsphase vergleichbare Ansprechraten für Bio-IR- und Non-Bio-IR-Patienten. Daher werden beide Fragestellungen zu einer Fragestellung zusammengefasst und Ergebnisse für die Zielpopulation dargestellt.

Studie U-ACHIEVE (M14-234)

Die Studie U-ACHIEVE (M14-234) ist eine direkt vergleichende, multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie, welche die Induktions- und Erhaltungsphase der Upadacitinib-Behandlung untersucht. Die Studienpopulation der Studie U-ACHIEVE entspricht der Zielpopulation. Es liegen Daten zu den Analysezeitpunkten 8 Wochen nach Therapiebeginn (Induktionsphase) und 52 Wochen nach Therapiebeginn (Erhaltungsphase) vor.

Studie UNIFI

Die Studie UNIFI ist eine direkt vergleichende, multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie. Die Studienpopulation der Studie UNIFI entspricht der Fragestellung. Es liegen Daten zum Analysezeitpunkt 8 Wochen nach Therapiebeginn (Induktionsphase) und 44 Wochen nach der Induktionsphase (Erhaltungsphase) vor.

In der vorliegenden Nutzenbewertung wird der Zusatznutzen von Upadacitinib im Vergleich zur ZVT Ustekinumab mittels eines adjustierten, indirekten Vergleichs nach Bucher über den Brückenkomparator Placebo auf der Basis der Studien U-ACHIEVE und der Studie UNIFI in der Indikation CU abgeleitet.

Zur Beurteilung des medizinischen Nutzens und Zusatznutzens von Upadacitinib wurden die Kategorien Mortalität, Morbidität und Verträglichkeit berücksichtigt. Das Aussmaß und die Wahrscheinlichkeit des Zusatznutzens einer Upadacitinib-Behandlung in einzelnen Endpunkten wird in der nachfolgenden Tabelle 4-138 berichtet.

Tabelle 4-138: Zusammenfassung der Ergebnisse des adjustierten indirekten Vergleichs der Studien U-ACHIEVE, U-ACCOMPLISH und UNIFI und daraus resultierender Zusatznutzen

Endpunkt	**Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert**		**Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert**	**Wahrscheinlichkeit und Ausmaß des Zusatznutzens**
	Upadacitinib	**Ustekinumab**		
Klinische Remission (Mayo Gesamtscore)				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit klinischer Remission	RR: 10,74 [5,47; 21,09] p < 0,0001	RR: 2,91 [1,72; 4,94] p < 0,0001	RR: 3,69 [1,57; 8,68] p = 0,0028	
Erhaltungsphase				
Anteil der Patienten mit klinischer Remission (niedrigere Dosierung)	RR: 3,78 [2,28; 6,24] p < 0,0001	RR: 1,60 [1,16; 2,21] p = 0,0046	RR: 2,36 [1,30; 4,29] p = 0,0049	
Anteil der Patienten mit klinischer Remission (höhere Dosierung)	RR: 4,41 [2,70; 7,22] p < 0,0001	RR: 1,82 [1,33; 2,49] p = 0,0002	RR: 2,42 [1,35; 4,34] p = 0,0029	
Klinische Remission (Adaptierter Mayo Score)				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit klinischer Remission	RR: 7,39 [4,42; 12,36] p < 0,0001	RR: 3,02 [1,87; 4,89] p < 0,0001	RR: 2,45 [1,21; 4,95] p = 0,0127	
Erhaltungsphase				
Anteil der Patienten mit klinischer Remission (niedrigere Dosierung)	RR: 3,52 [2,20; 5,65] p < 0,0001	RR: 1,61 [1,17; 2,21] p = 0,0034	RR: 2,19 [1,24; 3,87] p = 0,0070	
Anteil der Patienten mit klinischer Remission (höhere Dosierung)	RR: 4,30 [2,72; 6,81] p < 0,0001	RR: 1,73 [1,27; 2,37] p = 0,0005	RR: 2,48 [1,42; 4,32] p = 0,0013	

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Klinisches Ansprechen (Mayo Gesamtscore)				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit klinischem Ansprechen	RR: 3,04 [2,50; 3,70] p < 0,0001	RR: 1,97 [1,64; 2,37] p < 0,0001	RR: 1,55 [1,19; 2,02] p = 0,0011	
Erhaltungsphase				
Anteil der Patienten mit klinischem Ansprechen (niedrigere Dosierung)	RR: 2,81 [2,03; 3,89] p < 0,0001	RR: 1,53 [1,26; 1,85] p < 0,0001	RR: 1,84 [1,26; 2,69] p = 0,0017	
Anteil der Patienten mit klinischem Ansprechen (höhere Dosierung)	RR: 3,40 [2,48; 4,66] p < 0,0001	RR: 1,59 [1,32; 1,93] p < 0,0001	RR: 2,13 [1,48; 3,08] p < 0,0001	
Mukosaheilung				
Induktionsphase (ergänzend)				Anhaltspunkt für einen beträchtlichen Zusatznutzen
Anteil der Patienten mit Mukosaheilung	RR: 5,50 [3,55; 8,52] p < 0,0001	RR: 2,08 [1,36; 3,17] p = 0,0007	RR: 2,65 [1,44; 4,86] p = 0,0017	
Erhaltungsphase				
Anteil der Patienten mit Mukosaheilung (niedrigere Dosierung)	RR: 2,91 [1,79; 4,73] p < 0,0001	RR: 1,61 [1,16; 2,23] p = 0,0043	RR: 1,81 [1,01; 3,24] p = 0,0475	
Anteil der Patienten mit Mukosaheilung (höhere Dosierung)	RR: 4,14 [2,61; 6,56] p < 0,0001	RR: 1,90 [1,39; 2,60] p < 0,0001	RR: 2,17 [1,25; 3,79] p = 0,0063	
Stuhlfrequenz				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit Stuhlfrequenz ≤ 1	RR: 2,76 [2,25; 3,38] p < 0,0001	RR: 1,49 [1,25; 1,76] p < ,0001	RR: 1,86 [1,42; 2,42] p < 0,0001	
Erhaltungsphase				
Anteil der Patienten mit Stuhlfrequenz ≤ 1 (niedrigere Dosierung)	RR: 2,62 [1,88; 3,65] p < 0,0001	RR: 1,22 [1,04; 1,44] p = 0,0160	RR: 2,15 [1,48; 3,10] p < 0,0001	
Anteil der Patienten mit Stuhlfrequenz ≤ 1 (höhere Dosierung)	RR: 3,11 [2,26; 4,28] p < 0,0001	RR: 1,44 [1,24; 1,67] p < 0,0001	RR: 2,16 [1,52; 3,08] p < 0,0001	

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Rektale Blutungen				
Induktionsphase (ergänzend)				Anhaltspunkt für einen erheblichen Zusatznutzen
Anteil der Patienten mit Rektale Blutung = 0	RR: 2,33 [1,98; 2,74] p < 0,0001	RR: 1,70 [1,44; 2,00] p < 0,0001	RR: 1,37 [1,09; 1,73] p = 0,0073	
Erhaltungsphase				
Anteil der Patienten mit Rektale Blutung = 0 (niedrigere Dosierung)	RR: 2,64 [1,91; 3,64] p < 0,0001	RR: 1,38 [1,19; 1,60] p < 0,0001	RR: 1,91 [1,34; 2,72] p = 0,0004	
Anteil der Patienten mit Rektale Blutung = 0 (höhere Dosierung)	RR: 3,24 [2,38; 4,42] p < 0,0001	RR: 1,37 [1,18; 1,59] p < 0,0001	RR: 2,37 [1,68; 3,34] p < 0,0001	
EQ-5D VAS				
Induktionsphase (ergänzend)				Zusatznutzen nicht belegt
Mittlere Veränderung der EQ-5D VAS	Hedges‘ g: 0,78 [0,64; 0,91] p < 0,0001	Hedges‘ g: 0,41 [0,25; 0,57] p < 0,0001	Hedges‘ g: 0,37 [0,16; 0,57] p = 0,0006	
Erhaltungsphase				
Mittlere Veränderung der EQ-5D VAS (niedrigere Dosierung)	Hedges‘ g : 0,48 [0,24; 0,71] p < 0,0001	Hedges‘ g : 0,28 [0,07; 0,50] p = 0,0085	Hedges‘ g : 0,19 [-0,12; 0,50] p = 0,2310	
Mittlere Veränderung der EQ-5D VAS (höhere Dosierung)	Hedges‘ g: 0,63 [0,40; 0,86] p < 0,0001	Hedges‘ g: 0,56 [0,35; 0,77] p < 0,0001	Hedges‘ g: 0,07 [-0,24; 0,38] p = 0,6594	
Gesundheitsbezogene Lebensqualität: IBDQ				
Induktionsphase (ergänzend)				Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen
Mittlere Veränderung des IBDQ	Hedges‘ g: 1,01 [0,87; 1,15] p < 0,0001	Hedges‘ g: 0,60 [0,44; 0,76] p < 0,0001	Hedges‘ g: 0,41 [0,20; 0,62] p = 0,0002	
Erhaltungsphase				
Mittlere Veränderung des IBDQ (niedrigere Dosierung)	Hedges‘ g: 0,73 [0,50; 0,97] p < 0,0001	Hedges‘ g: 0,35 [0,14; 0,57] p = 0,0011	Hedges‘ g: 0,38 [0,06; 0,69] p = 0,0195	
Mittlere Veränderung des IBDQ (höhere Dosierung)	Hedges‘ g: 0,94 [0,71; 1,18] p < 0,0001	Hedges‘ g: 0,57 [0,35; 0,78] p < 0,0001	Hedges‘ g: 0,38 [0,06; 0,70] p = 0,0205	

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Gesundheitsbezogene Lebensqualität: SF-36 PCS				
Induktionsphase (ergänzend)				Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen
Mittlere Veränderung des SF-36 PCS	Hedges‘ g: 0,82 [0,68; 0,96] p < 0,0001	Hedges‘ g: 0,49 [0,34; 0,65] p < 0,0001	Hedges‘ g: 0,33 [0,12; 0,54] p = 0,0023	
Erhaltungsphase				
Mittlere Veränderung des SF-36 PCS(niedrigere Dosierung)	Hedges‘ g: 0,69 [0,45; 0,92] p < 0,0001	Hedges‘ g: 0,19 [-0,02; 0,40] p = 0,0763	Hedges‘ g: 0,49 [0,18; 0,81] p = 0,0021	
Mittlere Veränderung des SF-36 PCS (höhere Dosierung)	Hedges‘ g: 0,82 [0,58; 1,05] p < 0,0001	Hedges‘ g: 0,49 [0,28; 0,71] p < 0,0001	Hedges‘ g: 0,33 [0,01; 0,64] p = 0,0432	
Gesundheitsbezogene Lebensqualität: SF-36 MCS				
Induktionsphase (ergänzend)				Zusatznutzen nicht belegt
Mittlere Veränderung des SF-36 MCS	Hedges‘ g : 0,61 [0,47; 0,74] p < 0,0001	Hedges‘ g : 0,29 [0,14; 0,45] p = 0,0003	Hedges‘ g : 0,32 [0,11; 0,52] p = 0,0028	
Erhaltungsphase				
Mittlere Veränderung des SF-36 MCS (niedrigere Dosierung)	Hedges‘ g : 0,38 [0,15; 0,61] p = 0,0013	Hedges‘ g : 0,29 [0,08; 0,50] p = 0,0066	Hedges‘ g : 0,08 [-0,23; 0,39] p = 0,6037	
Mittlere Veränderung des SF-36 MCS (höhere Dosierung)	Hedges‘ g : 0,49 [0,26; 0,72] p < 0,0001	Hedges‘ g : 0,28 [0,07; 0,49] p = 0,0096	Hedges‘ g : 0,21 [-0,10; 0,52] p = 0,1797	
Gesamtraten UE				
Induktionsphase (ergänzend)				Zusatznutzen/ Schaden nicht belegt
Alle UE	RR: 1,06 [0,95; 1,19] p = 0,3099	RR: 1,06 [0,90; 1,24] p = 0,5011	RR: 1,01 [0,83; 1,22] p = 0,9517	
SUE	RR: 0,55 [0,31; 0,98] p = 0,0417	RR: 0,50 [0,25; 1,01] p = 0,0536	RR: 1,10 [0,44; 2,74] p = 0,8417	
UE, die zum Therapie-abbruch führten	RR: 0,36 [0,20; 0,63] p = 0,0005	RR: k. A. [k. A.; k. A.] p = k. A.	RR: k. A. [k. A.; k. A.] p = k. A.	

Endpunkt	Vergleich gegenüber Placebo Effektschätzer [95 %-KI] p-Wert		Indirekter Vergleich Effektschätzer [95 %-KI] p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
	Upadacitinib	Ustekinumab		
Erhaltungsphase				Zusatznutzen/ Schaden nicht belegt
Alle UE (niedrigere Dosierung)	RR: 1,02 [0,92; 1,12] p = 0,7365	RR: 0,88 [0,77; 1,00] p = 0,0415	RR: 1,16 [0,99; 1,36] p = 0,0692	
Alle UE (höhere Dosierung)	RR: 1,04 [0,94; 1,14] p = 0,4303	RR: 0,98 [0,88; 1,09] p = 0,7199	RR: 1,06 [0,92; 1,23] p = 0,4319	
SUE (niedrigere Dosierung)	RR: 0,89 [0,51; 1,57] p = 0,6996	RR: 0,78 [0,39; 1,55] p = 0,4764	RR: 1,15 [0,47; 2,81] p = 0,7588	
SUE (höhere Dosierung)	RR: 0,89 [0,51; 1,57] p = 0,6894	RR: 0,88 [0,45; 1,70] p = 0,6984	RR: 1,02 [0,43; 2,43] p = 0,9718	
UE, die zum Therapie-abbruch führten (niedrigere Dosierung)	RR: 0,39 [0,19; 0,80] p = 0,0099	RR: 0,46 [0,21; 0,98] p = 0,0434	RR: 0,86 [0,30; 2,42] p = 0,7698	
UE, die zum Therapie-abbruch führten (höhere Dosierung)	RR: 0,70 [0,39; 1,25] p = 0,2331	RR: 0,25 [0,10; 0,65] p = 0,0044	RR: 2,83 [0,92; 8,66] p = 0,0688	

Für binäre Endpunkte werden das RR mit zugehörigem 95 %-KI und p-Wert, sowohl für die Einzelstudien als auch für den indirekten Vergleich angegeben. Für kontinuierliche Endpunkte werden das Hedges‘ g mit zugehörigem 95 %-KI und p-Wert, sowohl für die Einzelstudien als auch für den indirekten Vergleich angegeben.

Bei den Ergebnissen des indirekten Vergleichs stellt ein RR größer Eins bzw. ein Hedges‘ g größer Null einen Vorteil von Upadacitinib dar. Bei den Ergebnissen zur Verträglichkeit stellt ein RR kleiner Eins einen Vorteil von Upadacitinib dar.

Bezogen auf die Einzelstudien, stellen für Endpunkte zur Morbidität und gesundheitsbezogenen Lebensqualität ein RR größer Eins bzw. ein Hedges‘ g größer Null einen Vorteil von Upadacitinib beziehungsweise Ustekinumab gegenüber Placebo dar. Für Endpunkte zur Verträglichkeit, stellt für die Einzelstudien ein RR kleiner Eins einen Vorteil von Upadacitinib beziehungsweise Ustekinumab gegenüber Placebo dar.

Niedrigere Dosierung: Adjustierter indirekter Vergleich der Behandlungsarme Upadacitinib 15 mg einmal täglich und Ustekinumab 90 mg alle 12 Wochen.

Höhere Dosierung: Adjustierter indirekter Vergleich der Behandlungsarme Upadacitinib 30 mg einmal täglich und Ustekinumab 90 mg alle 8 Wochen.

EQ-5D: EuroQol-5 dimensions, IBDQ: Inflammatory Bowel Disease Questionnaire; k. A.: Keine Angabe; KI: Konfidenzintervall; MCS: Mental Health Component Summary Score; PCS: Physical Health Component Summary Score; RR: Relatives Risiko; SF-36: Short Form – 36; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala

Der adjustierte indirekte Vergleich zwischen Upadacitinib und Ustekinumab, auf Basis der RCT U-ACHIEVE für Upadacitinib und UNIFI für Ustekinumab, zeigt ein

endpunktübergreifend sehr vergleichbares Bild mit deutlichen Behandlungsvorteilen für Upadacitinib sowohl in der Induktions- als auch in der Erhaltungsphase. Das Studiendesign mit Induktions- und Erhaltungsphase rekapituliert den deutschen Versorgungskontext, die beschriebenen Ergebnisse sind somit auch im Behandlungsalltag von CU-Patienten in Deutschland zu erwarten. Eine Upadacitinib-Behandlung resultiert somit häufig in einem schnellen und umfassenden Ansprechen sowie einer langanhaltenden klinischen Remission.

Morbidität

Klinische Remission (Mayo Gesamtscore und Adaptierter Mayo Score)

Bezüglich des Endpunkts klinische Remission (nach Mayo Gesamtscore und Adaptiertem Mayo Score) zeigte sich im vorliegenden indirekten Vergleich ein statistisch signifikanter Vorteil zugunsten von Upadacitinib gegenüber Ustekinumab. Der Mayo Gesamtscore wird, insbesondere aufgrund der Darstellung von vier Krankheitsdomänen, Stuhlfrequenz, rektale Blutungen, endoskopisches Erscheinungsbild und ärztliche Beurteilung, als stringentes und geeignetes Messinstrument angesehen, um das Therapieziel klinische Remission darzustellen (13, 14). Zum Analysezeitpunkt der Induktionsphase, bereits 8 Wochen nach Therapiebeginn, zeigte der indirekte Vergleich deutliche Vorteile von Upadacitinib gegenüber Ustekinumab in beiden betrachteten Dosierungen hinsichtlich des Endpunkts klinische Remission sowohl nach Mayo Gesamtscore (RR: 3,69 [1,57; 8,68]; p = 0,0028) als auch nach Adaptiertem Mayo Score (RR: 2,45 [1,21; 4,95]; p = 0,0127). Diese nachhaltigen Effekte konnten auch noch zum Analysezeitpunkt, 52 Wochen nach Therapiebeginn, gezeigt werden. Es ergaben sich für den Endpunkt klinische Remission sowohl nach Mayo Gesamtscore (Upadacitinib 15 mg: RR: 2,36 [1,30; 4,29]; p = 0,0049; Upadacitinib 30 mg: RR: 2,42 [1,35; 4,34]; p = 0,0029) als auch nach Adaptiertem Mayo Score (Upadacitinib 15 mg: RR: 2,19 [1,24; 3,87]; p = 0,0070; Upadacitinib 30 mg: RR: 2,48 [1,42; 4,32]; p = 0,0013) deutliche Vorteile zugunsten von Upadacitinib. Somit liegt in der Induktions- und in der Erhaltungsphase ein statistisch signifikanter und klinisch relevanter Vorteil bezüglich der klinischen Remission zugunsten von Upadacitinib gegenüber Ustekinumab vor.

Die Ergebnisse zur klinischen Remission gemäß Mayo Gesamtscore zu Woche 52 stellen im Sinne von § 5 Satz 7 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) eine Verringerung von schwerwiegenden Symptomen und somit eine für die Patienten bisher nicht erreichte große Verbesserung des therapierelevanten Nutzens im Sinne von § 2 Absatz 3 einer Upadacitinib-Behandlung gegenüber einer Ustekinumab-Behandlung dar. Dies rechtfertigt die Ableitung eines **Anhaltspunkts** für einen **erheblichen Zusatznutzen** bezüglich des Endpunkts klinische Remission.

Klinisches Ansprechen (Mayo Gesamtscore)

Bezüglich des klinischen Ansprechens per Mayo Gesamtscore zeigte sich bei der Behandlung der CU im vorliegenden indirekten Vergleich sowohl in der Induktionsphase also auch in der Erhaltungsphase ein nachhaltiger statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab. In der Induktionsphase (8 Wochen nach Therapiebeginn) ergaben sich folgende Effekte: RR: 1,55 [1,19; 2,02]; p = 0,0011. In der Erhaltungsphase (52 Wochen nach

Therapiebeginn) konnten die folgenden Effekte beobachtet werden: Upadacitinib 15 mg: RR: 1,84 [1,26; 2,69]; p = 0,0017; Upadacitinib 30 mg: RR: 2,13 [1,48; 3,08]; p < 0,0001. Somit liegt sowohl in der Induktions- als auch in der Erhaltungsphase für beide Dosierungen ein statistisch signifikanter und klinisch relevanter, nachhaltiger Vorteil bezüglich des klinischen Ansprechens zugunsten von Upadacitinib gegenüber Ustekinumab vor.

Die Ergebnisse bezüglich des klinischen Ansprechens gemäß Mayo Gesamtscore zu Woche 52 stellen im Sinne von § 5 Satz 7 der AM-NutzenV eine Verringerung von schwerwiegenden Symptomen und somit eine für die Patienten bisher nicht erreichte große Verbesserung des therapierelevanten Nutzens im Sinne von § 2 Absatz 3 einer Upadacitinib-Behandlung gegenüber einer Ustekinumab-Behandlung dar. Dies rechtfertigt die Ableitung eines **Anhaltspunkts** für einen **erheblichen Zusatznutzen** bezüglich des Endpunkts klinische Remission.

Mukosaheilung (Endoskopischer Subscore und Geboes Score)

Bezüglich des Endpunkts Mukosaheilung, definiert als endoskopischer Subscore des Mayo Gesamtscores von null oder eins und einem Geboes Score ≤ 3,1, zeigte sich bei der Behandlung der CU im vorliegenden indirekten Vergleich ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab. Der endoskopische Mayo Subscore erfasst das Erscheinungsbild der mukosalen Entzündung. Der Geboes Score umfasst insgesamt sechs Kategorien, um das histologische Erscheinungsbild der mukosalen Entzündung zu evaluieren (31). Das Erfassen der Mukosaheilung per histo-endoskopischer Mukosaverbesserung wird als stringentes und geeignetes Messinstrument angesehen, um das Therapieziel der Mukosaheilung darzustellen. Bereits zum Analysezeitpunkt der Induktionsphase, 8 Wochen nach Therapiebeginn, zeigte der indirekte Vergleich statistisch signifikante Vorteile von Upadacitinib gegenüber Ustekinumab (RR: 2,65 [1,44; 4,86]; p = 0,0017). Ferner zeigte der indirekte Vergleich von Upadacitinib gegenüber Ustekinumab in beiden Dosierungen zum Analysezeitpunkt, 52 Wochen nach Therapiebeginn, Vorteile für Upadacitinib hinsichtlich des Endpunkts Mukosaheilung (Upadacitinib 15 mg: RR: 1,81 [1,01; 3,24]; p = 0,0475; Upadacitinib 30 mg: RR: 2,17 [1,25; 3,79]; p = 0,0063). Somit liegt in der Induktions- und in der Erhaltungsphase ein statistisch signifikanter und klinisch relevanter Vorteil bezüglich der Mukosaheilung zugunsten von Upadacitinib gegenüber Ustekinumab vor.

Die Ergebnisse zur Mukosaheilung gemäß histo-endoskopischer Mukosaverbesserung zu Woche 52 stellen im Sinne von § 5 Satz 7 der AM-NutzenV eine spürbare Linderung der Erkrankung und somit eine für die Patienten bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens im Sinne von § 2 Absatz 3 einer Upadacitinib-Behandlung gegenüber einer Ustekinumab-Behandlung dar. Dies rechtfertigt die Ableitung eines **Anhaltspunkts** für einen **beträchtlichen Zusatznutzen** bezüglich des Endpunkts Mukosaheilung.

Stuhlfrequenz (Mayo Subscore)

Bezüglich des Endpunkts Stuhlfrequenz per Mayo Subscore zeigte sich bei der Behandlung der CU im vorliegenden indirekten Vergleich ein statistisch signifikanter Vorteil von Upadacitinib

gegenüber Ustekinumab. Bereits 8 Wochen nach Therapiebeginn, also zum Analysezeitpunkt der Induktionsphase, zeigte der indirekte Vergleich statistisch signifikante Vorteile von Upadacitinib gegenüber Ustekinumab hinsichtlich des Endpunkts Stuhlfrequenz (RR: 1,86 [1,42; 2,42]; p < 0,0001). Ferner zeigte der indirekte Vergleich von Upadacitinib gegenüber Ustekinumab in beiden Dosierungen 44 Wochen nach Therapiebeginn, zum Analysezeitpunkt der Erhaltungsphase, statistisch signifikante Vorteile für Upadacitinib hinsichtlich des Endpunkts Stuhlfrequenz (Upadacitinib 15 mg: RR: 2,15 [1,48; 3,10]; p < 0,0001; Upadacitinib 30 mg: RR: 2,16 [1,52; 3,08]; p < 0,0001). Somit liegt in der Induktions- und in der Erhaltungsphase ein statistisch signifikanter und klinisch relevanter Vorteil bezüglich des Endpunkts Stuhlfrequenz zugunsten von Upadacitinib gegenüber Ustekinumab vor.

Die Ergebnisse bezüglich des Endpunkts Stuhlfrequenz gemäß Mayo Subscore zu Woche 44 stellen im Sinne von § 5 Satz 7 der AM-NutzenV eine Verringerung von schwerwiegenden Symptomen und somit eine für die Patienten bisher nicht erreichte große Verbesserung des therapierelevanten Nutzens im Sinne von § 2 Absatz 3 einer Upadacitinib-Behandlung gegenüber einer Ustekinumab-Behandlung dar. Dies rechtfertigt die Ableitung eines **Anhaltspunkts** für einen **erheblichen Zusatznutzen** bezüglich des Endpunkts Stuhlfrequenz gemäß Mayo Subscore.

Rektale Blutungen (Mayo Subscore)

Bezüglich des Endpunkts rektale Blutung per Mayo Subscore zeigte sich bei der Behandlung der CU im vorliegenden indirekten Vergleich ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab. Der Mayo Subscore für rektale Blutungen erfasst die Anzahl blutiger Stuhlgänge pro Tag. Das Erfassen rektaler Blutungen per Mayo Subscore wird als stringentes und geeignetes Messinstrument angesehen, um eine Reduktion der Symptomatik der CU darzustellen. Bereits zum Analysezeitpunkt der Induktionsphase, 8 Wochen nach Therapiebeginn, zeigte der indirekte Vergleich statistisch signifikante Vorteile von Upadacitinib gegenüber Ustekinumab hinsichtlich des Endpunkts rektale Blutungen (RR: 1,37 [1,09; 1,73]; p = 0,0073). Ferner zeigte der indirekte Vergleich von Upadacitinib gegenüber Ustekinumab in beiden Dosierungen zum Analysezeitpunkt, 44 Wochen nach Therapiebeginn, statistisch signifikante Vorteile für Upadacitinib hinsichtlich des Endpunkts rektale Blutungen (Upadacitinib 15 mg: RR: 1,91 [1,34; 2,72]; p = 0,0004; Upadacitinib 30 mg: RR: 2,37 [1,68; 3,34]; p < 0,0001). Somit liegt in der Induktions- und in der Erhaltungsphase ein statistisch signifikanter und klinisch relevanter Vorteil bezüglich des Endpunkts rektale Blutungen zugunsten von Upadacitinib gegenüber Ustekinumab vor.

Die Ergebnisse bezüglich rektale Blutungen gemäß Mayo Subscore zu Woche 44 stellen im Sinne von § 5 Satz 7 der AM-NutzenV eine Verringerung von schwerwiegenden Symptomen und somit eine für die Patienten bisher nicht erreichte große Verbesserung des therapierelevanten Nutzens im Sinne von § 2 Absatz 3 einer Upadacitinib-Behandlung gegenüber einer Ustekinumab-Behandlung dar. Dies rechtfertigt die Ableitung eines **Anhaltspunkts** für einen **erheblichen Zusatznutzen** bezüglich des Endpunkts klinische Remission.

EQ-5D VAS

Bezüglich des Endpunkts allgemeiner Gesundheitszustand gemessen anhand der EQ-5D VAS zeigte sich bei der Behandlung der CU im vorliegenden indirekten Vergleich kein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab zum Analysezeitpunkt der Erhaltungsphase, 52 Wochen nach Therapiebeginn, in beiden Dosierungen (Erhaltungsphase: Upadacitinib 15 mg: Hedges‘ g: 0,19 [-0,12; 0,50]; p = 0,2310; Upadacitinib 30 mg: Hedges‘ g: 0,07 [-0,24; 0,38]; p = 0,6594). In der Erhaltungsphase liegt somit kein statistisch signifikanter und klinisch relevanter Vorteil bezüglich des allgemeinen Gesundheitszustands per EQ-5D VAS zugunsten von Upadacitinib gegenüber Ustekinumab vor.

Im Sinne von § 5 Absatz 7 AM-NutzenV kann aufgrund dieser Ergebnisse **kein Zusatznutzen** im Endpunkt allgemeiner Gesundheitszustand (EQ-5D VAS) festgestellt werden.

Gesundheitsbezogene Lebensqualität

IBDQ

Bezüglich des Endpunkts gesundheitsbezogene Lebensqualität gemessen anhand des IBDQ zeigte sich bei der Behandlung der CU im vorliegenden indirekten Vergleich ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab. Bereits in der Induktionsphase (Hedges‘ g: 0,41 [0,20; 0,62]; p = 0,0002), 8 Wochen nach Therapiebeginn, als auch zum Analysezeitpunkt, 52 Wochen nach Therapiebeginn, zeigte sich in der Betrachtung der mittleren Veränderung des IBDQ gegenüber Baseline eine größere Verbesserung der gesundheitsbezogenen Lebensqualität unter beiden Upadacitinib-Dosierungen verglichen mit den entsprechenden Ustekinumab-Dosierungen (Upadacitinib 15 mg: Hedges‘ g: 0,38 [0,06; 0,69]; p = 0,0195; Upadacitinib 30 mg: Hedges‘ g: 0,38 [0,06; 0,70]; p = 0,0205).

Die Ergebnisse zum IBDQ zeigen eine statistisch signifikante Verbesserung unter einer Upadacitinib-Behandlung gegenüber einer Ustekinumab-Behandlung zu Woche 52. Aus dem statistisch signifikanten Ergebnis lässt sich ein relevanter Effekt zugunsten von Upadacitinib, im Sinne von § 5 Satz 7 der AM-NutzenV, ableiten. Daher wird für den Endpunkt ein **Anhaltspunkt** für einen **nicht quantifizierbaren Zusatznutzen** abgeleitet.

Verbesserung des SF-36 PCS

Bezüglich des Endpunkts gesundheitsbezogene Lebensqualität gemessen anhand des SF-36 PCS zeigte sich bei der Behandlung der CU im vorliegenden indirekten Vergleich ein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab. Bereits in der Induktionsphase (Hedges‘ g: 0,33 [0,12; 0,54]; p = 0,0023), 8 Wochen nach Therapiebeginn, als auch zum Analysezeitpunkt, 52 Wochen nach Therapiebeginn, zeigte sich in der Betrachtung der mittleren Veränderung des SF-36 PCS gegenüber Baseline eine größere Verbesserung der gesundheitsbezogenen Lebensqualität unter beiden Upadacitinib-Dosierungen verglichen mit den entsprechenden Ustekinumab-Dosierungen (Upadacitinib 15 mg: Hedges‘ g: 0,49 [0,18; 0,81]; p = 0,0021; Upadacitinib 30 mg: Hedges‘ g: 0,33 [0,01; 0,64]; p = 0,0432).

Die Ergebnisse zur Verbesserung des SF-36 PCS zeigen eine statistisch signifikante Verbesserung unter einer Upadacitinib-Behandlung gegenüber einer Ustekinumab-Behandlung zu Woche 52. Aus dem statistisch signifikanten Ergebnis lässt sich ein relevanter Effekt zugunsten von Upadacitinib, im Sinne von § 5 Satz 7 der AM-NutzenV, ableiten. Daher wird für den Endpunkt ein **Anhaltspunkt** für einen **nicht quantifizierbaren Zusatznutzen** abgeleitet.

Verbesserung des SF-36 MCS

Bezüglich des Endpunkts gesundheitsbezogene Lebensqualität gemessen anhand des SF-36 MCS zeigte sich bei der Behandlung der CU im vorliegenden indirekten Vergleich kein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab zum Analysezeitpunkt der Erhaltungsphase, 52 Wochen nach Therapiebeginn. In der Betrachtung der mittleren Veränderung des SF-36 MCS zu Baseline zeigten sich keine signifikanten Vorteile für Upadacitinib gegenüber Ustekinumab in beiden Dosierungen (Upadacitinib 15 mg: Hedges‘ g: 0,08 [-0,23; 0,39]; p = 0,6037; Upadacitinib 30 mg: Hedges‘ g: 0,21 [-0,10; 0,52]; p = 0,1797). In der Erhaltungsphase liegt somit kein statistisch signifikanter und klinisch relevanter Vorteil bezüglich der gesundheitsbezogenen Lebensqualität zugunsten von Upadacitinib gegenüber Ustekinumab vor.

Im Sinne von § 5 Absatz 7 AM-NutzenV kann aufgrund dieser Ergebnisse **kein Zusatznutzen** im Endpunkt gesundheitsbezogene Lebensqualität gemessen anhand des SF-36 MCS festgestellt werden.

Verträglichkeit

Bezüglich der Gesamtraten der UE, bei den SUE, und bei UE die zum Therapieabbruch führten zeigten sich bei der Behandlung der CU im vorliegenden indirekten Vergleich jeweils kein statistisch signifikanter Vorteil von Upadacitinib gegenüber Ustekinumab zum Analysezeitpunkt der Erhaltungsphase, 52 Wochen nach Therapiebeginn.

Im Sinne von § 5 Absatz 7 AM-NutzenV kann aufgrund dieser Ergebnisse **kein Zusatznutzen** im Endpunkt Verträglichkeit festgestellt werden.

Subgruppenanalysen

Aus den in Abschnitt 4.2.5.5 genannten Gründen werden keine Subgruppen dargestellt. In den deskriptiven Auswertungen der Subgruppe Bio-IR-Status zeigten sich bei einer Behandlung unter Upadacitinib für die Erhaltungsphase vergleichbare Ansprechraten für Bio-IR- und Non-Bio-IR-Patienten. Diese Ergebnisse rechtfertigen die Ableitung eines Zusatznutzens für die Zielpopulation von Upadacitinib in der Indikation CU.

Zusammenfassung

Wie in Tabelle 4-139 beschrieben, ergeben sich zahlreiche, deutliche und statistisch signifikante Vorteile zugunsten einer Upadacitinib-Therapie. In den Nutzendimensionen Morbidität und gesundheitsbezogene Lebensqualität weist Upadacitinib eine bisher nicht erreichte Wirksamkeit auf und damit verglichen mit Ustekinumab eine deutliche Überlegenheit

bezüglich dem Erreichen von patientenrelevanten Therapiezielen wie z. B. einer langfristigen klinischen Remission und der nachhaltigen Verbesserung der Lebensqualität. Besonders wichtig für die Patienten ist auch die Mukosaheilung, da die chronische Entzündung mukosaler Zellen die pathophysiologische Ursache der Erkrankung ist, ergibt sich aus der Mukosaheilung insgesamt eine günstige Prognose für die Patienten. Auch die besonders belastende Symptomatik der Stuhlfrequenz und der rektalen Blutung wurde unter einer Upadacitinib Therapie nachhaltig und deutlich verbessert. Für einen Großteil der Patienten bedeutet dies eine signifikante Verbesserung der physischen aber auch psychischen Einschränkungen durch die Erkrankung. Als direkte Konsequenz folgt eine signifikante Verbesserung der patientenbezogenen Lebensqualität.

Unter Behandlung mit Upadacitinib ist somit die Wahrscheinlichkeit signifikant höher, das patientenrelevante Haupttherapieziel einer raschen klinischen Remission zu erreichen und die Remission langfristig zu erhalten, und somit das vollständige Abklingen der mukosalen Entzündung. Zudem erfahren unter einer Behandlung mit Upadacitinib mehr Patienten eine Verbesserung der gesundheitsbezogenen Lebensqualität als unter Ustekinumab.

Mit einer Upadacitinib-Therapie kann somit eine umfassende Kontrolle der Krankheitsaktivität (klinische Remission und Mukosaheilung) bei Patienten mit CU erreicht werden. Dies zeigt sich auch anhand der signifikant höheren klinischen Ansprechraten als unter der ZVT Ustekinumab. Damit kann mit Upadacitinib eine langfristige Kontrolle der Krankheitsaktivität erreicht werden. Aufgrund seiner überlegenen Wirksamkeit gegenüber Ustekinumab und eines günstigen Nutzen-Risiko-Profils bietet Upadacitinib daher einen erheblichen Mehrwert für die Behandlung von Patienten mit CU.

Das Verträglichkeitsprofil von Upadacitinib ist vergleichbar mit dem von Ustekinumab und entspricht dem bekannten und konsistenten Sicherheitsprofil von Upadacitinib in den bisher zugelassenen AWG. Upadacitinib ist daher wegen seines guten Wirksamkeits- und Verträglichkeitsprofils eine Bereicherung der Therapielandschaft in Deutschland.

In der Gesamtschau zeigt sich für eine Upadacitinib-Therapie bei vorbehandelten Patienten mit CU im Vergleich zu Ustekinumab zum Analysezeitpunkt 52 Wochen (Erhaltungsphase) ein **Anhaltspunkt** für einen **erheblichen Zusatznutzen**.

Tabelle 4-139: Zusammenfassung der positiven und negativen Effekte von Upadacitinib gegenüber Ustekinumab aus dem adjustierten indirekten Vergleich

Positive Effekte (statistisch signifikante Effekte)	**Negative Effekte (statistisch signifikante Effekte)**
• Klinische Remission (Mayo Gesamtscore) **Anhaltspunkt** für einen **erheblichen** Zusatznutzen • Klinische Remission (Adaptierter Mayo Score) **Anhaltspunkt** für einen **erheblichen** Zusatznutzen	-

Positive Effekte (statistisch signifikante Effekte)	**Negative Effekte (statistisch signifikante Effekte)**
• Klinisches Ansprechen (Mayo Gesamtscore) **Anhaltspunkt** für einen **erheblichen** Zusatznutzen • Mukosaheilung **Anhaltspunkt** für einen **beträchtlichen** Zusatznutzen • Stuhlfrequenz **Anhaltspunkt** für einen **erheblichen** Zusatznutzen • Rektale Blutungen **Anhaltspunkt** für einen **erheblichen** Zusatznutzen • Inflammatory Bowel Disease Questionnaire (IBDQ) **Anhaltspunkt** für einen **nicht quantifizierbaren** Zusatznutzen • SF-36 PCS **Anhaltspunkt** für einen **nicht quantifizierbaren** Zusatznutzen	
IBDQ: Inflammatory Bowel Disease Questionnaire; PCS: Physical Health Component Summary Score; SF-36: Short Form – 36	

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 4-140: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	**Ausmaß des Zusatznutzens**
Erwachsene Patienten mit mittelschwerer bis schwerer aktiver Colitis ulcerosa, die auf eine konventionelle oder biologische Therapie unzureichend angesprochen haben, nicht mehr ansprechen oder diese nicht vertragen haben.	Erheblich (Anhaltspunkt)

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.

Für die Patientenpopulation mit mittelschwerer bis schwerer aktiver CU, die auf eine konventionelle oder biologische Therapie unzureichend angesprochen haben, nicht mehr ansprechen oder diese nicht vertragen haben, lagen für den direkten Vergleich von Upadacitinib gegenüber der vom G-BA festgelegten ZVT keine direkt vergleichenden Studien vor. Es wurde ein adjustierter indirekter Vergleich nach Bucher durchgeführt, welcher die beste verfügbare Evidenz für einen Vergleich von Upadacitinib gegenüber der ZVT Ustekinumab darstellt. Dazu wurden RCTs zu Upadacitinib bzw. Ustekinumab in systematischen Literatur- und Registerrecherchen identifiziert, die jeweils den gleichen Komparator untersuchen, welcher anschließend im indirekten Vergleich als Brückenkomparator herangezogen werden kann. Mit den dabei identifizierten Studien U-ACHIEVE und U-ACCOMPLISH, sowie der Studie UNIFI konnte Upadacitinib über den Brückenkomparator Placebo mit der ZVT Ustekinumab verglichen werden. Die eingeschlossenen Studien weisen ein geringes Verzerrungspotenzial auf und sind in Studiendesign, Population, Studiendauer und berichteten Endpunkten vergleichbar. Das Studiendesign sah sowohl bei Upadacitinib, als auch bei Ustekinumab eine Induktionsphase und eine anschließende Re-Randomisierung der Therapieansprecher in der Erhaltungsphase vor. Auch im Versorgungsalltag wird die Erhaltungsphase bei Patienten nach der Induktion nur fortgeführt, wenn diese ausreichend angesprochen haben. Die erneute Re-Randomisierung in der Erhaltungsphase ist zwischen den Studien ausreichend ähnlich, sodass die Durchführung eines indirekten Vergleichs adäquat ist. Die Studienpopulation schließt bei beiden Studien Patienten mit aktiver Colitis ein, die entsprechend bei Upadacitinib gemäß dem Adaptierten Mayo Score, beziehungsweise bei Ustekinumab gemäß dem Mayo Gesamtscore eine mittelschwere bis schwere Colitis ulcerosa aufweisen mussten. Bei beiden Studien musste der Wert der endoskopischen Subdomäne ≥ 2 sein, sodass von einer hinreichend ähnlichen Patientenpopulation gesprochen werden kann. Die Studiendauer ist mit einem Erhebungszeitpunkt von 52 Wochen bei Upadacitinib und einem Erhebungszeitpunkt von 44 Wochen bei Ustekinumab adäquat. Die EMA-Guideline empfiehlt eine Studiendauer in der vorliegenden Indikation von 52 Wochen, welche sowohl die Induktions- als auch die Erhaltungsphase einschließen sollten. Vor diesem Hintergrund erscheint der Analysezeitpunkt von Woche 44 adäquat. Der Verlauf der Krankheitsaktivität zeigt konstante Wirksamkeitsdaten zwischen dem Zeitraum Woche 44 und Woche 52 bei Upadacitinib, sodass ein Vergleich der Erhaltungsphase zu den Zeitpunkten Woche 44 und Woche 52 adäquat ist. Die berichteten Endpunkte sind zwischen den Studien hinreichend ähnlich erhoben und operationalisiert worden.

In der Gesamtschau sind die Studien hinsichtlich der Strukturmerkmale für einen indirekten Vergleich ausreichend ähnlich und für den Nachweis des Zusatznutzens mittels eines indirekten Vergleichs als beste verfügbare Evidenz geeignet.

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.

Nicht zutreffend.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005[17], Molenberghs 2010[18]). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrundeliegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

[17] Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.
[18] 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.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006[19]) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%- Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006[20]) 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.

Nicht zutreffend.

4.6 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.

[19] Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.
[20] Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

1. AbbVie Deutschland GmbH & Co. KG. Fachinformation für RINVOQ® 15 mg Retardtabletten, RINVOQ® 30 mg Retardtabletten, RINVOQ® 45 mg Retardtabletten Stand: Juli 2022.
2. Gemeinsamer Bundesausschuss (G-BA). Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie: Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V) Filgotinib (neues Anwendungsgebiet: Colitis ulcerosa). Vom 19. Mai 2022. 2022. Verfügbar unter: https://www.g-ba.de/downloads/39-261-5432/2022-05-19_AM-RL-XII_Filgotinib_D-743.pdf. [Zugriff am: 08.06.2022]
3. Lammert F, Lynen Jansen P, Lerch MM. Weissbuch Gastroenterologie 2020/2021. 2021.
4. Argyriou K, Kapsoritakis A, Oikonomou K, Manolakis A, Tsakiridou E, Potamianos S. Disability in Patients with Inflammatory Bowel Disease: Correlations with Quality of Life and Patient's Characteristics. Can J Gastroenterol Hepatol. 2017;2017:6138105.
5. Armuzzi A, Liguori G. Quality of life in patients with moderate to severe ulcerative colitis and the impact of treatment: A narrative review. Dig Liver Dis. 2021;53(7):803-8.
6. Bernklev T, Jahnsen J, Henriksen M, Lygren I, Aadland E, Sauar J, et al. Relationship between sick leave, unemployment, disability, and health-related quality of life in patients with inflammatory bowel disease. Inflamm Bowel Dis. 2006;12(5):402-12.
7. Knowles SR, Keefer L, Wilding H, Hewitt C, Graff LA, Mikocka-Walus A. Quality of Life in Inflammatory Bowel Disease: A Systematic Review and Meta-analyses-Part II. Inflamm Bowel Dis. 2018;24(5):966-76.
8. Van Assche G, Peyrin-Biroulet L, Sturm A, Gisbert JP, Gaya DR, Bokemeyer B, et al. Burden of disease and patient-reported outcomes in patients with moderate to severe ulcerative colitis in the last 12 months - Multicenter European cohort study. Dig Liver Dis. 2016;48(6):592-600.
9. Gemeinsamer Bundesausschuss (G-BA). Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Vedolizumab, Stand: 8. Januar 2015 2015. Verfügbar unter: https://www.g-ba.de/downloads/40-268-3074/2015-01-08_AM-RL-XII_Vedolizumab_2014-07-15-D-122_TrG.pdf. [Zugriff am: 08.06.2022]
10. European Medicines Agency (EMA). Guideline on the development of new medicinal products for the treatment of Ulcerative Colitis. Stand: 28. Juni 2018. 2018. Verfügbar unter: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-development-new-medicinal-products-treatment-ulcerative-colitis-revision-1_en.pdf. [Zugriff am: 08.06.2022]
11. Stange EF, Travis SP, Vermeire S, Reinisch W, Geboes K, Barakauskiene A, et al. European evidence-based Consensus on the diagnosis and management of ulcerative colitis: Definitions and diagnosis. J Crohns Colitis. 2008;2(1):1-23.
12. Kucharzik T, Koletzko S, Kannengiesser K, Dignass A. Ulcerative Colitis-Diagnostic and Therapeutic Algorithms. Dtsch Arztebl Int. 2020;117(33-34):564-74.
13. Kucharzik T, Dignass AU, Atreya R, Bokemeyer B, Esters P, Herrlinger K, et al. Aktualisierte S3-Leitlinie Colitis ulcerosa - Living Guideline. Z Gastroenterol. 2020;58(12):e241-e326.

14. Sturm A, Maaser C, Calabrese E, Annese V, Fiorino G, Kucharzik T, et al. ECCO-ESGAR Guideline for Diagnostic Assessment in IBD Part 2: IBD scores and general principles and technical aspects. J Crohns Colitis. 2019;13(3):273-84.
15. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. Lancet. 2017;389(10080):1756-70.
16. Raine T, Bonovas S, Burisch J, Kucharzik T, Adamina M, Annese V, et al. ECCO Guidelines on Therapeutics in Ulcerative Colitis: Medical Treatment. J Crohns Colitis. 2022;16(1):2-17.
17. Lewis JD, Chuai S, Nessel L, Lichtenstein GR, Aberra FN, Ellenberg JH. Use of the noninvasive components of the Mayo score to assess clinical response in ulcerative colitis. Inflamm Bowel Dis. 2008;14(12):1660-6.
18. Naegeli AN, Hunter T, Dong Y, Hoskin B, Middleton-Dalby C, Hetherington J, et al. Full, Partial, and Modified Permutations of the Mayo Score: Characterizing Clinical and Patient-Reported Outcomes in Ulcerative Colitis Patients. 2021.
19. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG Clinical Guideline: Ulcerative Colitis in Adults. Am J Gastroenterol. 2019;114(3):384-413.
20. D'Haens G, Sandborn WJ, Feagan BG, Geboes K, Hanauer SB, Irvine EJ, et al. A review of activity indices and efficacy end points for clinical trials of medical therapy in adults with ulcerative colitis. Gastroenterology. 2007;132(2):763-86.
21. U.S. Department of Health and Human Services. Ulcerative Colitis: Clinical Trial Endpoints Guidance for Industry Stand: August 2016. 2016. Verfügbar unter: https://www.fda.gov/files/drugs/published/Ulcerative-Colitis--Clinical-Trial-Endpoints-Guidance-for-Industry.pdf. [Zugriff am: 08.06.2022]
22. Cazzato G, Colagrande A, Andriola V, Lettini T, Cicco S, Candance PMV, et al. Histological Hallmarks of Mucosal Healing in Inflammatory Bowel Diseases in the Era of Monoclonal Antibodies Therapy: New Insights and Perspectives. Diagnostics (Basel). 2021;11(9).
23. Boal Carvalho P, Cotter J. Mucosal Healing in Ulcerative Colitis: A Comprehensive Review. Drugs. 2017;77(2):159-73.
24. Colombel JF, Rutgeerts P, Reinisch W, Esser D, Wang Y, Lang Y, et al. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. Gastroenterology. 2011;141(4):1194-201.
25. Lasa JS, Olivera PA, Danese S, Peyrin-Biroulet L. Efficacy and safety of biologics and small molecule drugs for patients with moderate-to-severe ulcerative colitis: a systematic review and network meta-analysis. Lancet Gastroenterol Hepatol. 2022;7(2):161-70.
26. Burr NE, Gracie DJ, Black CJ, Ford AC. Efficacy of biological therapies and small molecules in moderate to severe ulcerative colitis: systematic review and network meta-analysis. Gut. 2021.
27. Shah SC, Colombel JF, Sands BE, Narula N. Mucosal Healing Is Associated With Improved Long-term Outcomes of Patients With Ulcerative Colitis: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol. 2016;14(9):1245-55 e8.
28. Monstad IL, Solberg IC, Cvancarova M, Hovde O, Henriksen M, Huppertz-Hauss G, et al. Outcome of Ulcerative Colitis 20 Years after Diagnosis in a Prospective Population-based Inception Cohort from South-Eastern Norway, the IBSEN Study. J Crohns Colitis. 2021;15(6):969-79.

29. Reinink AR, Lee TC, Higgins PD. Endoscopic Mucosal Healing Predicts Favorable Clinical Outcomes in Inflammatory Bowel Disease: A Meta-analysis. Inflamm Bowel Dis. 2016;22(8):1859-69.
30. Jauregui-Amezaga A, Geerits A, Das Y, Lemmens B, Sagaert X, Bessissow T, et al. A Simplified Geboes Score for Ulcerative Colitis. J Crohns Colitis. 2017;11(3):305-13.
31. Geboes K, Riddell R, Ost A, Jensfelt B, Persson T, Lofberg R. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. Gut. 2000;47(3):404-9.
32. Tursi A, Elisei W, Picchio M, Forti G, Penna A, Inchingolo CD, et al. Histological inflammation in ulcerative colitis in deep remission under treatment with infliximab. Clin Res Hepatol Gastroenterol. 2015;39(1):107-13.
33. Marchal Bressenot A, Riddell RH, Boulagnon-Rombi C, Reinisch W, Danese S, Schreiber S, et al. Review article: the histological assessment of disease activity in ulcerative colitis. Aliment Pharmacol Ther. 2015;42(8):957-67.
34. Sandborn WJ, Su C, Sands BE, D'Haens GR, Vermeire S, Schreiber S, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2017;376(18):1723-36.
35. Sands BE, Sandborn WJ, Panaccione R, O'Brien CD, Zhang H, Johanns J, et al. Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2019;381(13):1201-14.
36. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel JF, Sandborn WJ, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2013;369(8):699-710.
37. Ko JK, Auyeung KK. Inflammatory bowel disease: etiology, pathogenesis and current therapy. Curr Pharm Des. 2014;20(7):1082-96.
38. Arijs I, De Hertogh G, Lemmens B, Van Lommel L, de Bruyn M, Vanhove W, et al. Effect of vedolizumab (anti-alpha4beta7-integrin) therapy on histological healing and mucosal gene expression in patients with UC. Gut. 2018;67(1):43-52.
39. Fiorino G, Cesarini M, Indriolo A, Malesci A. Mucosal healing in ulcerative colitis: where do we stand? Curr Drug Targets. 2011;12(10):1417-23.
40. Schnitzler F, Fidder H, Ferrante M, Noman M, Arijs I, Van Assche G, et al. Mucosal healing predicts long-term outcome of maintenance therapy with infliximab in Crohn's disease. Inflamm Bowel Dis. 2009;15(9):1295-301.
41. Theede K, Kiszka-Kanowitz M, Nordgaard-Lassen I, Mertz Nielsen A. The Impact of Endoscopic Inflammation and Mucosal Healing on Health-related Quality of Life in Ulcerative Colitis Patients. J Crohns Colitis. 2015;9(8):625-32.
42. Mazzuoli S, Guglielmi FW, Antonelli E, Salemme M, Bassotti G, Villanacci V. Definition and evaluation of mucosal healing in clinical practice. Dig Liver Dis. 2013;45(12):969-77.
43. Cottone M, Scimeca D, Mocciaro F, Civitavecchia G, Perricone G, Orlando A. Clinical course of ulcerative colitis. Dig Liver Dis. 2008;40 Suppl 2:S247-52.
44. Ben Ghezala I, Charkaoui M, Michiels C, Bardou M, Luu M. Small Molecule Drugs in Inflammatory Bowel Diseases. Pharmaceuticals (Basel). 2021;14(7).
45. Gemeinsamer Bundesausschuss (G-BA). Niederschrift zum Beratungsgespräch gemäß § 8 Absatz 1 AM-NutzenV, Beratungsanforderung B-2017-083. 2017.
46. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.

47. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337-43.
48. Szende A, Janssen B. Cross-Country Analysis of EQ-5D Data. In: Szende A, Janssen B , Cabases J (Hrsg.). Self-Reported Population Health: An International Perspective based on EQ-5D. Dordrecht (NL)2014. S. 31-6.
49. Gemeinsamer Bundesausschuss (G-BA). Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V: Tofacitinib (neues Anwendungsgebiet: Colitis ulcerosa), Stand: 21. Februar 2019 2019. Verfügbar unter: https://www.g-ba.de/downloads/40-268-5587/2019-02-21_AM-RL-XII_Tofacitinib_D-374_TrG.pdf. [Zugriff am: 08.06.2022]
50. EuroQol Research Foundation. EQ-5D-5L User Guide, Stand: September 2019. 2019. Verfügbar unter: file:///C:/Users/niebauer/AppData/Local/Temp/EQ-5D-5LUserguide-08-0421.pdf. [Zugriff am: 08.06.2022]
51. Williet N, Sarter H, Gower-Rousseau C, Adrianjafy C, Olympie A, Buisson A, et al. Patient-reported Outcomes in a French Nationwide Survey of Inflammatory Bowel Disease Patients. J Crohns Colitis. 2017;11(2):165-74.
52. Stark RG, Reitmeir P, Leidl R, Konig HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. Inflamm Bowel Dis. 2010;16(1):42-51.
53. Sammut J, Scerri J, Xuereb RB. The lived experience of adults with ulcerative colitis. J Clin Nurs. 2015;24(17-18):2659-67.
54. Kobayashi T, Siegmund B, Le Berre C, Wei SC, Ferrante M, Shen B, et al. Ulcerative colitis. Nat Rev Dis Primers. 2020;6(1):74.
55. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden. Version 6.1 vom 24. Januar 2022. Verfügbar unter: https://www.iqwig.de/methoden/allgemeine-methoden-v6-1.pdf. [Zugriff am: 08.06.2022]
56. Guyatt G, Mitchell A, Irvine EJ, Singer J, Williams N, Goodacre R, et al. A new measure of health status for clinical trials in inflammatory bowel disease. Gastroenterology. 1989;96(3):804-10.
57. Mitchell A, Guyatt G, Singer J, Irvine EJ, Goodacre R, Tompkins C, et al. Quality of life in patients with inflammatory bowel disease. J Clin Gastroenterol. 1988;10(3):306-10.
58. Irvine EJ, Feagan B, Rochon J, Archambault A, Fedorak RN, Groll A, et al. Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. Canadian Crohn's Relapse Prevention Trial Study Group. Gastroenterology. 1994;106(2):287-96.
59. Higgins PD, Schwartz M, Mapili J, Krokos I, Leung J, Zimmermann EM. Patient defined dichotomous end points for remission and clinical improvement in ulcerative colitis. Gut. 2005;54(6):782-8.
60. Bernklev T, Jahnsen J, Lygren I, Henriksen M, Vatn M, Moum B. Health-related quality of life in patients with inflammatory bowel disease measured with the short form-36: psychometric assessments and a comparison with general population norms. Inflamm Bowel Dis. 2005;11(10):909-18.

61. Pallis AG, Mouzas IA, Vlachonikolis IG. The inflammatory bowel disease questionnaire: a review of its national validation studies. Inflamm Bowel Dis. 2004;10(3):261-9.
62. K-H. Janke, B. Klump, U. Steder-Neukamm, J. Hoffmann, W. Häuser. Validierung der Deutschen Version (Kompetenznetz „Chronisch entzündliche Darmerkrankungen") des Inflammatory Bowel Disease Questionnaire IBDQ-D. 2006. Verfügbar unter: https://www.thieme-connect.com/products/ejournals/abstract/10.1055/s-2006-932661. [Zugriff am: 08.06.2022]
63. Ellert U, Kurth BM. [Methodological views on the SF-36 summary scores based on the adult German population]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2004;47(11):1027-32. Methodische Betrachtungen zu den Summenscores des SF-36 anhand der erwachsenen bundesdeutschen Bevolkerung.
64. Tarlov AR, Ware JE, Jr., Greenfield S, Nelson EC, Perrin E, Zubkoff M. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. JAMA. 1989;262(7):925-30.
65. Coteur G, Feagan B, Keininger DL, Kosinski M. Evaluation of the meaningfulness of health-related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients with active Crohn's disease. Aliment Pharmacol Ther. 2009;29(9):1032-41.
66. Danese S, Vermeire S, Zhou W, Pangan AL, Siffledeen J, Greenbloom S, et al. Upadacitinib as induction and maintenance therapy for moderately to severely active ulcerative colitis: results from three phase 3, multicentre, double-blind, randomised trials. Lancet. 2022;399(10341):2113-28.
67. Panaccione R, Danese S, Sandborn WJ, O'Brien CD, Zhou Y, Zhang H, et al. Ustekinumab is effective and safe for ulcerative colitis through 2 years of maintenance therapy. Aliment Pharmacol Ther. 2020;52(11-12):1658-75.
68. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997;50(6):683-91.
69. 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):1-134, iii-iv.
70. Kiefer C, Sturtz S, Bender R. Indirect Comparisons and Network Meta-Analyses. Dtsch Arztebl Int. 2015;112(47):803-8.
71. Clinicaltrials.gov. A Study of the Efficacy and Safety of Upadacitinib (ABT-494) in Participants With Moderately to Severely Active Ulcerative Colitis (U-ACCOMPLISH). 2022. Verfügbar unter: https://clinicaltrials.gov/ct2/show/NCT03653026. [Zugriff am: 08.06.2022]
72. EU-CTR. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Induction Study to Evaluate the Efficacy and Safety of Upadacitinib (ABT-494) in Subjects with Moderately to Severely Active Ulcerative Colitis (U-ACCOMPLISH). 2022. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/trial/2016-000642-62/NL. [Zugriff am: 08.06.2022]
73. WHO-ICTRP. NCT03653026 - A Multicenter, Randomized, Double-Blind, Placebo-Controlled Induction Study to Evaluate the Efficacy and Safety of Upadacitinib (ABT-494) in Subjects With Moderately to Severely Active Ulcerative Colitis (U-ACCOMPLISH). 2022. Verfügbar unter: https://trialsearch.who.int/?TrialID=NCT03653026. [Zugriff am: 08.06.2022]

74. WHO-ICTRP. JPRN-JapicCTI-184131 - A Multicenter, Randomized, Double-Blind, Placebo-Controlled Induction Study to Evaluate the Efficacy and Safety of Upadacitinib (ABT-494) in Subjects With Moderately to Severely Active Ulcerative Colitis (U-ACCOMPLISH). 2022. Verfügbar unter: https://trialsearch.who.int/?TrialID=JPRN-JapicCTI-184131. [Zugriff am: 08.06.2022]
75. Clinicaltrials.gov. A Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) for Induction and Maintenance Therapy in Participants With Moderately to Severely Active Ulcerative Colitis (U-ACHIEVE). 2022. Verfügbar unter: https://clinicaltrials.gov/ct2/show/NCT02819635. [Zugriff am: 08.06.2022]
76. EU-CTR. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) for Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Ulcerative Colitis (U-ACHIEVE). 2022. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/trial/2016-000641-31/SK. [Zugriff am: 08.06.2022]
77. WHO-ICTRP. NCT02819635 - A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) for Induction and Maintenance Therapy in Subjects With Moderately to Severely Active Ulcerative Colitis (U-ACHIEVE). 2022. Verfügbar unter: https://trialsearch.who.int/?TrialID=NCT02819635. [Zugriff am: 08.06.2022]
78. Clinicaltrials.gov. A Study to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Participants With Moderately to Severely Active Ulcerative Colitis (UNIFI). 2022. Verfügbar unter: https://clinicaltrials.gov/ct2/show/NCT02407236. [Zugriff am: 08.06.2022]
79. EU-CTR. A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Protocol to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Ulcerative Colitis (UNIFI). 2022. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/trial/2014-005606-38/DE. [Zugriff am: 08.06.2022]
80. WHO-ICTRP. NCT02407236 - A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Protocol to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Subjects With Moderately to Severely Active Ulcerative Colitis (UNIFI). 2022. Verfügbar unter: https://trialsearch.who.int/?TrialID=NCT02407236. [Zugriff am: 08.06.2022]
81. Janssen Research & Development L. A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Study to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Ulcerative Colitis - UNIFI. 2018. Verfügbar unter: https://portal.dimdi.de/data/ctr/O-2670573-3-0-65E8C6-20200304151728.pdf. [Zugriff am: 09.06.2022]
82. AbbVie Inc. Clinical Study Report - U-ACCOMPLISH (M14-675) - A Multicenter, Randomized, Double-Blind, Placebo-Controlled Induction Study to Evaluate the Efficacy and Safety of Upadacitinib (ABT-494) in Subjects with Moderately to Severely Active Ulcerative Colitis - Date of Report: 29. June 2021. 2021.
83. AbbVie Inc. Clinical Study Report - U-ACHIEVE (M14-234) - A Multicenter, Randomized, Double-Blind, Placebo- Controlled Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) for Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Ulcerative Colitis - Substudy 3: Phase 3 Maintenance - Date of Report: 03. August 2021.

84. Sandborn WJ, Ghosh S, Panes J, Schreiber S, D'Haens G, Tanida S, et al. Efficacy of Upadacitinib in a Randomized Trial of Patients With Active Ulcerative Colitis. Gastroenterology. 2020;158(8):2139-49 e14.
85. Ghosh S, Sanchez Gonzalez Y, Zhou W, Clark R, Xie W, Louis E, et al. Upadacitinib Treatment Improves Symptoms of Bowel Urgency and Abdominal Pain, and Correlates With Quality of Life Improvements in Patients With Moderate to Severe Ulcerative Colitis. J Crohns Colitis. 2021;15(12):2022-30.
86. O'Hagan P, Limdi J, Akbar A, Tucknott S, Kahol DN. Ulcerative colitis: understanding the impact of ulcerative colitis on everyday life and exploring the unmet needs of patients. Current Medical Research and Opinion. 2021;37(11):1901-11.
87. Kim JH, Cheon JH, Park Y, Lee HJ, Park SJ, Kim TI, et al. Effect of mucosal healing (Mayo 0) on clinical relapse in patients with ulcerative colitis in clinical remission. Scand J Gastroenterol. 2016;51(9):1069-74.

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	07.11.2016	
Zeitsegment	1974 to 2016 November 04	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[21]] – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	Diabetes Mellitus/	552986
2	Non Insulin Dependent Diabetes Mellitus/	195234
3	(diabet* or niddm or t2dm).ab,ti.	714228
4	or/1-3	847068
5	linagliptin*.mp.	1562
6	(random* or double-blind*).tw.	1193849
7	placebo*.mp.	388057
8	or/6-7	1382838
9	and/4,5,8	633

[21] 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	Embase Classic+Embase 1947 bis 01. Juni 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	1947 bis 01. Juni 2022
Suchfilter	nach Wong et al. 2006

#	Suchbegriffe	Ergebnis
1	exp upadacitinib/	1.200
2	upadacitinib.mp.	1.233
3	(ABT494 or ABT-494 or ABT 494).mp.	124
4	rinvoq.mp.	33
5	1 or 2 or 3 or 4	1.245
6	exp ulcerative colitis/ or ulcer* coliti*.mp.	99.361
7	exp Colitis, Ulcerative/ or coliti* ulcer*.mp.	90.527
8	ulcer* colorectitis*.mp.	4
9	exp proctocolitis/ or ulcer* proctocolitis*.mp.	1.417
10	ulcer* procto-colitis*.mp.	3
11	idiopathic proctocolitis*.mp.	47
12	idiopathic procto-colitis*.mp.	1
13	colitis gravis*.mp.	8
14	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	206.904
15	6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14	212.201
16	random:.tw. or placebo:.mp. or double-blind:.tw.	2.088.336
17	5 and 15 and 16	130

Datenbankname	Ovid MEDLINE® and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions 1946 bis 01. Juni 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	1946 to June 01, 2022
Suchfilter	nach Wong et al. 2006

#	**Suchbegriffe**	**Ergebnis**
1	upadacitinib.mp.	337
2	(ABT494 or ABT-494 or ABT 494).mp.	20
3	rinvoq.mp.	10
4	1 or 2 or 3	342
5	exp Colitis, Ulcerative/ or ulcer* coliti*.mp.	54.820
6	coliti* ulcer*.mp.	3.8913
7	ulcer* colorectitis*.mp.	2
8	ulcer* proctocolitis*.mp.	57
9	ulcer* procto-colitis*.mp.	1
10	exp Proctocolitis/ or idiopathic proctocolitis.mp.	861
11	idiopathic procto-colitis*.mp.	2
12	colitis gravis*.mp.	7
13	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	110.784
14	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	117.737
15	randomized controlled trial.pt. or randomized.mp. or placebo.mp.	1.029.852
16	4 and 14 and 15	26

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials April 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	April 2022
Suchfilter	Kein Suchfilter verwendet

#	Suchbegriffe	Ergebnis
1	upadacitinib.mp.	477
2	(ABT494 or ABT-494 or ABT 494).mp.	99
3	rinvoq.mp.	3
4	1 or 2 or 3	506
5	ulcer* coliti*.mp. or exp Colitis, Ulcerative/	5.814
6	coliti* ulcer*.mp.	1.912
7	ulcer* colorectitis*.mp.	0
8	ulcer* proctocolitis*.mp.	2
9	ulcer* procto-colitis*.mp.	0
10	idiopathic proctocolitis*.mp.	2
11	idiopathic procto-colitis*.mp.	0
12	colitis gravis*.mp.	1
13	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	6.416
14	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	9.290
15	4 and 14	65

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Suche nach dem zu bewertenden Arzneimittel (Upadacitinib)

Datenbankname	Embase Classic+Embase 1947 bis 01. Juni 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	1947 bis 01. Juni 2022
Suchfilter	nach Wong et al. 2006

#	Suchbegriffe	Ergebnis
18	exp upadacitinib/	1.200
19	upadacitinib.mp.	1.233
20	(ABT494 or ABT-494 or ABT 494).mp.	124
21	rinvoq.mp.	33
22	1 or 2 or 3 or 4	1.245
23	exp ulcerative colitis/ or ulcer* coliti*.mp.	99.361
24	exp Colitis, Ulcerative/ or coliti* ulcer*.mp.	90.527
25	ulcer* colorectitis*.mp.	4
26	exp proctocolitis/ or ulcer* proctocolitis*.mp.	1.417
27	ulcer* procto-colitis*.mp.	3
28	idiopathic proctocolitis*.mp.	47
29	idiopathic procto-colitis*.mp.	1
30	colitis gravis*.mp.	8
31	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	206.904
32	6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14	212.201
33	random:.tw. or placebo:.mp. or double-blind:.tw.	2.088.336
34	5 and 15 and 16	130

Datenbankname	Ovid MEDLINE® and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions 1946 bis 01. Juni 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	1946 bis 01. Juni 2022
Suchfilter	nach Wong et al. 2006

#	Suchbegriffe	Ergebnis
17	upadacitinib.mp.	337
18	(ABT494 or ABT-494 or ABT 494).mp.	20
19	rinvoq.mp.	10
20	1 or 2 or 3	342
21	exp Colitis, Ulcerative/ or ulcer* coliti*.mp.	54.820
22	coliti* ulcer*.mp.	3.8913
23	ulcer* colorectitis*.mp.	2
24	ulcer* proctocolitis*.mp.	57
25	ulcer* procto-colitis*.mp.	1
26	exp Proctocolitis/ or idiopathic proctocolitis.mp.	861
27	idiopathic procto-colitis*.mp.	2
28	colitis gravis*.mp.	7
29	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	110.784
30	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	117.737
31	randomized controlled trial.pt. or randomized.mp. or placebo.mp.	1.029.852
32	4 and 14 and 15	26

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials April 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	April 2022
Suchfilter	Kein Suchfilter verwendet

#	Suchbegriffe	Ergebnis
16	upadacitinib.mp.	477
17	(ABT494 or ABT-494 or ABT 494).mp.	99
18	rinvoq.mp.	3
19	1 or 2 or 3	506
20	ulcer* coliti*.mp. or exp Colitis, Ulcerative/	5.814
21	coliti* ulcer*.mp.	1.912
22	ulcer* colorectitis*.mp.	0
23	ulcer* proctocolitis*.mp.	2
24	ulcer* procto-colitis*.mp.	0
25	idiopathic proctocolitis*.mp.	2
26	idiopathic procto-colitis*.mp.	0
27	colitis gravis*.mp.	1
28	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	6.416
29	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	9.290
30	4 and 14	65

Suche nach der ZVT (Ustekinumab)

Datenbankname	Embase Classic+Embase 1947 bis 01. Juni 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	1947 bis 01. Juni 2022
Suchfilter	RCT-Filter nach Wong et al. 2006

#	Suchbegriffe	Ergebnis
1	exp ustekinumab/	10.035
2	ustekinumab.mp.	10.295
3	(CNTO1275 or CNTO-1275 or CNTO 1275).mp.	146
4	stelara.mp.	560
5	1 or 2 or 3 or 4	10.356
6	exp ulcerative colitis/ or ulcer* coliti*.mp.	99.361
7	exp Colitis, Ulcerative/ or coliti* ulcer*.mp.	90.527
8	ulcer* colorectitis*.mp.	4
9	exp proctocolitis/ or ulcer* proctocolitis*.mp.	1.417
10	ulcer* procto-colitis*.mp.	3
11	idiopathic proctocolitis*.mp.	47
12	idiopathic procto-colitis*.mp.	1
13	colitis gravis*.mp.	8
14	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	206.904
15	6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14	212.201
16	random:.tw. or placebo:.mp. or double-blind:.tw.	2.088.336
17	5 and 15 and 16	814

Datenbankname	Ovid MEDLINE® and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions 1946 bis 01. Juni 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	1946 bis 01. Juni 2022
Suchfilter	RCT-Filter nach Wong et al. 2006

#	Suchbegriffe	Ergebnis
1	ustekinumab.mp.	2.710
2	(CNTO1275 or CNTO-1275 or CNTO 1275).mp.	17
3	stelara.mp.	47
4	1 or 2 or 3	2.713
5	exp Colitis, Ulcerative/ or ulcer* coliti*.mp.	54.820
6	coliti* ulcer*.mp.	38.913
7	ulcer* colorectitis*.mp.	2
8	ulcer* proctocolitis*.mp.	57
9	ulcer* procto-colitis*.mp.	1
10	exp Proctocolitis/ or idiopathic proctocolitis.mp.	861
11	idiopathic procto-colitis*.mp.	2
12	colitis gravis*.mp.	7
13	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	110.784
14	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	117.737
15	randomized controlled trial.pt. or randomized.mp. or placebo.mp.	1.029.852
16	4 and 14 and 15	108

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials April 2022
Suchoberfläche	Ovid
Datum der Suche	02.06.2022
Zeitsegment	April 2022
Suchfilter	Kein Suchfilter verwendet

#	Suchbegriffe	Ergebnis
1	ustekinumab.mp.	1.108
2	(CNTO1275 or CNTO-1275 or CNTO 1275).mp.	57
3	stelara.mp.	95
4	1 or 2 or 3	1.115
5	ulcer* coliti*.mp. or exp Colitis, Ulcerative/	5.814
6	coliti* ulcer*.mp.	2.057
7	ulcer* colorectitis*.mp.	1
8	ulcer* proctocolitis*.mp.	7
9	ulcer* procto-colitis*.mp.	0
10	idiopathic proctocolitis*.mp.	3
11	idiopathic procto-colitis*.mp.	0
12	colitis gravis*.mp.	3
13	inflammatory bowel disease.mp. or exp inflammatory bowel disease/	6.416
14	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	9.437
15	4 and 14	214

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken 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 jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. clinicaltrials.gov), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank 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/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07.11.2016
Eingabeoberfläche	Standard Search
Suchstrategie	linagliptin OR BI 1356
Treffer	169

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Studienregister	ClinicalTrials.gov (CT.gov)
Internetadresse	https://clinicaltrials.gov/ct2/search/advanced
Datum der Suche	02.06.2022
Suchstrategie	Ulcerative Colitis OR UC OR colitis ulcerosa [disease], Upadacitinib OR ABT494 OR ABT 494 OR ABT-494 OR Rinvoq [Intervention]
Treffer	**3**

Studienregister	EU Clinical Trials Register (EU-CTR)
Internetadresse	http://www.clinicaltrialsregister.eu
Datum der Suche	02.06.2022
Suchstrategie	(ulcerative colitis) AND (Upadacitinib OR ABT494 OR "ABT 494" OR ABT-494 OR Rinvoq)
Treffer	**3**

Studienregister	WHO International Clinical Trials Registry Platform Search Portal (ICTRP) (advanced search)
Internetadresse	https://trialsearch.who.int/AdvSearch.aspx
Datum der Suche	02.06.2022
Eingabeoberfläche	Advanced Search
Suchstrategie	Ulcerative colitis OR UC OR colitis ulcerosa [Condition] Upadacitinib OR ABT494 OR ABT 494 OR ABT-494 OR Rinvoq [Intervention]
Treffer	**5**

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Suche nach dem zu bewertenden Arzneimittel (Upadacitinib)

Studienregister	ClinicalTrials.gov (CT.gov)
Internetadresse	https://clinicaltrials.gov/ct2/search/advanced
Datum der Suche	02.06.2022
Suchstrategie	Ulcerative Colitis OR UC OR colitis ulcerosa [disease], Upadacitinib OR ABT494 OR ABT 494 OR ABT-494 OR Rinvoq [Intervention]
Treffer	**3**

Studienregister	EU Clinical Trials Register (EU-CTR)
Internetadresse	http://www.clinicaltrialsregister.eu
Datum der Suche	02.06.2022
Suchstrategie	(ulcerative colitis) AND (Upadacitinib OR ABT494 OR "ABT 494" OR ABT-494 OR Rinvoq)
Treffer	**3**

Studienregister	WHO International Clinical Trials Registry Platform Search Portal (ICTRP) (advanced search)
Internetadresse	https://trialsearch.who.int/AdvSearch.aspx
Datum der Suche	02.06.2022
Eingabeoberfläche	Advanced Search
Suchstrategie	Ulcerative colitis OR UC OR colitis ulcerosa [Condition] Upadacitinib OR ABT494 OR ABT 494 OR ABT-494 OR Rinvoq [Intervention]
Treffer	**5**

Suche nach der ZVT (Ustekinumab)

Studienregister	ClinicalTrials.gov (CT.gov)
Internetadresse	https://clinicaltrials.gov/ct2/search/advanced
Datum der Suche	02.06.2022
Suchstrategie	Ulcerative Colitis OR UC OR colitis ulcerosa [disease] Ustekinumab OR CNTO1275 OR CNTO-1275 OR CNTO 1275 OR Stelara [Intervention]
Treffer	**14**

Studienregister	EU Clinical Trials Register (EU-CTR)
Internetadresse	http://www.clinicaltrialsregister.eu
Datum der Suche	02.06.2022
Suchstrategie	(ulcerative colitis) AND (Ustekinumab OR CNTO1275 OR CNTO-1275 OR "CNTO 1275" OR Stelara)
Treffer	**22**

Studienregister	WHO International Clinical Trials Registry Platform Search Portal (ICTRP)
Internetadresse	https://trialsearch.who.int/AdvSearch.aspx
Datum der Suche	02.06.2022
Eingabeoberfläche	Advanced Search
Suchstrategie	Ulcerative colitis OR UC OR colitis ulcerosa [Condition] Ustekinumab OR CNTO1275 OR CNTO-1275 OR CNTO 1275 OR Stelara [Intervention]
Treffer	**18**

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

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.	Quelle	Ausschlussgrund
1.	Danese S, Vermeire S, Zhou W, Pangan AL, Siffledeen J, Greenbloom S, Hebuterne X, D'Haens G, Nakase H, Panes J, Higgins PDR, Juillerat P, Lindsay JO, Loftus EV Jr, Sandborn WJ, Reinisch W, Chen MH, Sanchez Gonzalez Y, Huang B, Xie W, Liu J, Weinreich MA, Panaccione R, (et al.), 2022, Upadacitinib as induction and maintenance therapy for moderately to severely active ulcerative colitis: results from three phase 3, multicentre, double-blind, randomised trials., Lancet. 2022 May 26.,	A3 andere Vergleichstherapie

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Suche nach dem zu bewertenden Arzneimittel (Upadacitinib)

Es wurden keine Volltexte ausgeschlossen.

Suche nach der ZVT (Ustekinumab)

Nr.	Quelle	Ausschlussgrund
1.	Abreu M.T., Rowbotham D.S., Danese S., Sandborn W.J., Miao Y., Zhang H., Tikhonov I., Panaccione R., Hisamatsu T., Scherl E.J., Leong R.W., Arasaradnam R.P., Afif W., Peyrin-Biroulet L., Sands B.E., Marano C., (et al.), 2022, Efficacy and Safety of Maintenance Ustekinumab for Ulcerative Colitis Through 3 Years: UNIFI Long-term Extension., Journal of Crohn's & colitis., (no pagination)	A3 andere Vergleichstherapie
2.	Adedokun O.J., Xu Z., Marano C., O'Brien C., Szapary P., Zhang H., Johanns J., Leong R.W., Hisamatsu T., Van Assche G., Danese S., Abreu M.T., Sands B.E., Sandborn W.J., (et al.), 2020, Ustekinumab Pharmacokinetics and Exposure Response in a Phase 3 Randomized Trial of Patients With Ulcerative Colitis., Clinical Gastroenterology and Hepatology. , 18(10) (pp 2244-2255.e9)	A4 andere Endpunkte
3.	Hisamatsu T, Kim HJ, Motoya S, Suzuki Y, Ohnishi Y, Fujii N, Matsushima N, Zheng R, Marano CW, (et al.), 2021, Efficacy and safety of ustekinumab in East Asian patients with moderately to severely active ulcerative colitis: a subpopulation analysis of global phase 3 induction and maintenance studies (UNIFI)., Intestinal Research. , 19(4):386-397	A4 andere Endpunkte

Nr.	Quelle	Ausschlussgrund
4.	Hu C, Zhou H, Sharma A, (et al.), 2020, Facilitating Longitudinal Exposure-Response Modeling of a Composite Endpoint Using the Joint Modeling of Sparsely and Frequently Collected Subcomponents., AAPS Journal. , 22(4):79	A4 andere Endpunkte
5.	Li K., Marano C., Zhang H., Yang F., Sandborn W.J., Sands B.E., Feagan B.G., Rubin D.T., Peyrin-Biroulet L., Friedman J.R., De Hertogh G., (et al.), 2020, Relationship Between Combined Histologic and Endoscopic Endpoints and Efficacy of Ustekinumab Treatment in Patients With Ulcerative Colitis., Gastroenterology. , 159(6) (pp 2052-2064)	A4 andere Endpunkte
6.	Panaccione R., Danese S., Sandborn W.J., O'Brien C.D., Zhou Y., Zhang H., Adedokun O.J., Tikhonov I., Targan S., Abreu M.T., Hisamatsu T., Scherl E.J., Leong R.W., Rowbotham D.S., Arasaradnam R.P., Sands B.E., Marano C., (et al.), 2020, Ustekinumab is effective and safe for ulcerative colitis through 2 years of maintenance therapy., Alimentary Pharmacology and Therapeutics. , 52(11-12) (pp 1658-1675)	A5 anderer Studientyp
7.	Sands BE, Sandborn WJ, Panaccione R, O'Brien CD, Zhang H, Johanns J, Adedokun OJ, Li K, Peyrin-Biroulet L, Van Assche G, Danese S, Targan S, Abreu MT, Hisamatsu T, Szapary P, Marano C, UNIFI Study Group, (et al.), 2019, Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis., New England Journal of Medicine. , 381(13):1201-1214	A7 anderer Publikationstyp
8.	Xu Y., Hu C., Chen Y., Miao X., Adedokun O.J., Xu Z., Sharma A., Zhou H., (et al.), 2020, Population Pharmacokinetics and Exposure-Response Modeling Analyses of Ustekinumab in Adults With Moderately to Severely Active Ulcerative Colitis., Journal of Clinical Pharmacology. , 60(7) (pp 889-902)	A4 andere Endpunkte

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken)

Listen Sie nachfolgend die durch die Studienregistersuche(n)/ Studienergebnisdatenbanksuche(n) identifizierten, aber ausgeschlossenen Registereinträgen 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.	Studien-ID	Titel	Ausschlussgrund
Clinicaltrials.gov			
(1)	NCT02819635	A Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) for Induction and Maintenance Therapy in Participants With Moderately to Severely Active Ulcerative Colitis (UC). ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://clinicaltrials.gov/ct2/show/NCT02819635	A3 Andere Vergeleichstherapie
(2)	NCT03006068	A Study to Evaluate the Long-Term Safety and Efficacy of Upadacitinib (ABT-494) in Participants With Ulcerative Colitis (UC). ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03006068	A5 Keine RCT
(3)	NCT03653026	A Study of the Efficacy and Safety of Upadacitinib (ABT-494) in Participants With Moderately to Severely Active Ulcerative Colitis. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03653026	A3 Andere Vergeleichstherapie
EU-CTR			
(4)	2016-000641-31	A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) for Induction and Maintenance Therapy in Subjects with Moderately to [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000641-31	A3 Andere Vergeleichstherapie
(5)	2016-000642-62	M14-675, A Multicenter, Randomized, Double-Blind, Placebo-Controlled Induction Study to Evaluate the Efficacy and Safety of Upadacitinib (ABT-494) in Subjects with Moderately to Severely Active Ulc [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000642-62	A3 Andere Vergeleichstherapie

Nr.	Studien-ID	Titel	Ausschlussgrund
(6)	2016-000674-38	A Phase 3 Multicenter, Long-Term Extension Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) in Subjects with Ulcerative Colitis (UC). EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000674-38	A5 Keine RCT
ICTRP			
(7)	EUCTR2016-000674-38-SK	Long Term Safety and Efficacy of ABT-494 in Subjects with Ulcerative Colitis. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2016-000674-38-SK	A5 Keine RCT
(8)	JPRN-JapicCTI-184131	A Induction Study to Evaluate the Efficacy and Safety of Upadacitinib (ABT-494) in Subjects With Ulcerative Colitis. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-JapicCTI-184131	A3 Andere Vergeleichstherapie
(9)	NCT02819635	A Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) for Induction and Maintenance Therapy in Participants With Moderately to Severely Active Ulcerative Colitis (UC). ICTRP. 2021. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02819635	A3 Andere Vergeleichstherapie
(10)	NCT03006068	A Study to Evaluate the Long-Term Safety and Efficacy of Upadacitinib (ABT-494) in Participants With Ulcerative Colitis (UC). ICTRP. 2021. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03006068	A5 Keine RCT
(11)	NCT03653026	A Study of the Efficacy and Safety of Upadacitinib (ABT-494) in Participants With Moderately to Severely Active Ulcerative Colitis. ICTRP. 2021. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03653026	A3 Andere Vergeleichstherapie

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Suche nach dem zu bewertenden Arzneimittel (Upadacitinib)

Nr.	Studien-ID	Titel	Ausschlussgrund
Clinicaltrials.gov			
(1)	NCT03006068	A Study to Evaluate the Long-Term Safety and Efficacy of Upadacitinib (ABT-494) in Participants With Ulcerative Colitis (UC). ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03006068	A5 Keine RCT
EU-CTR			
(2)	2016-000674-38	A Phase 3 Multicenter, Long-Term Extension Study to Evaluate the Safety and Efficacy of Upadacitinib (ABT-494) in Subjects with Ulcerative Colitis (UC). EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000674-38	A5 Keine RCT
ICTRP			
(3)	NCT03006068	A Study to Evaluate the Long-Term Safety and Efficacy of Upadacitinib (ABT-494) in Participants With Ulcerative Colitis (UC). ICTRP. 2021. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03006068	A5 Keine RCT

Suche nach der ZVT (Ustekinumab)

Nr.	Studien-ID	Titel	Ausschlussgrund
Clinicaltrials.gov			
(1)	NCT03006198	Tracking Biologics Along the Silk Road. ClinicalTrials.gov. 2019. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03006198	A5 Keine RCT
(2)	NCT03358706	A Study to Evaluate the Effect of Ustekinumab on Cytochrome P450 Enzyme Activities Following Induction and Maintenance Dosing in Participants With Active Crohn's Disease or Ulcerative Colitis. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03358706	A5 Keine RCT
(3)	NCT03606499	Real-world Effectiveness of Ustekinumab in Participants Suffering From Inflammatory Bowel Disease (Crohn's Disease or Ulcerative Colitis) With Extra-intestinal Manifestations or Immune-mediated Inflammatory Diseases. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03606499	A5 Keine RCT

Nr.	Studien-ID	Titel	Ausschlussgrund
(4)	NCT03885713	Identification of Predictive Biomarkers for Response to Biologic Therapies in Inflammatory Bowel Disease. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03885713	A5 Keine RCT
(5)	NCT04259138	Determination of the Optimal Treatment Target in Ulcerative Colitis. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT04259138	A2 Andere Intervention
(6)	NCT04372108	A Study to Assess the Long-Term Safety of Ustekinumab Versus Other Biologics in Patients With Crohn's Disease and Ulcerative Colitis. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT04372108	A5 Keine RCT
(7)	NCT04505410	Second Line Induction Therapy and Healthy Diet for Patients With Ulcerative Colitis. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT04505410	A2 Andere Intervention
(8)	NCT04630028	A Study of Ustekinumab in Pediatric Participants With Moderately to Severely Active Ulcerative Colitis (UC). ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT04630028	A1 Andere Patientenpopulation
(9)	NCT04963725	A Study of Symptom Improvement of Ulcerative Colitis After an Induction Dose of Ustekinumab in Japanese Clinical Practice. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT04963725	A5 Keine RCT
(10)	NCT05092269	A Long-term Extension Study of Ustekinumab in Pediatric Participants. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT05092269	A1 Andere Patientenpopulation
(11)	NCT05186623	Prediction Model for Response to Biologics and Small Molecular Agent for UC. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT05186623	A5 Keine RCT
(12)	NCT05327790	LFMT vs Placebo in New Biologic Start for Ulcerative Colitis. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT05327790	A2 Andere Intervention
(13)	NCT05386290	A Multicentered Prospective Cohort Study of Chinese IBD Patients. ClinicalTrials.gov. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT05386290	A5 Keine RCT

Nr.	Studien-ID	Titel	Ausschlussgrund
EU-CTR			
(14)	2010-020003-73	A PHASE 3, MULTI-SITE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-020003-73	A1 Andere Patientenpopulation
(15)	2017-000402-38	A Randomized, Double-blind, Placebo-controlled Phase 2 Study to Evaluate the Testicular Safety of Filgotinib in Adult Males with Moderately to Severely Active Inflammatory Bowel Disease. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-000402-38	A5 Keine RCT
(16)	2017-003703-22	A phase 2, multicenter, randomized, double-blind, placebo controlled, dose-finding study to evaluate the efficacy and safety of IMU 838 for induction and maintenance therapy in moderate-to-severe u [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-003703-22	A5 Keine RCT
(17)	2018-002925-47	Identification of predictive biomarkers for response to biologic therapies in inflammatory bowel disease by proteomic and mass cytometry approaches. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-002925-47	A2 Andere Intervention
(18)	2018-003985-15	A Phase 3, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Assess the Efficacy and Safety of Etrasimod in Subjects with Moderately to Severely Active Ulcerative Colitis. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-003985-15	A2 Andere Intervention
(19)	2018-003986-33	A Phase 3, Randomized, Double Blind, Placebo Controlled, 12 Week Study to Assess the Efficacy and Safety of Etrasimod in Subjects with Moderately to Severely Active Ulcerative Colitis. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-003986-33	A2 Andere Intervention

Nr.	Studien-ID	Titel	Ausschlussgrund
(20)	2018-004694-27	A Phase 2 Randomized, Double-Blind, Placebo-Controlled Study of the Safety and Efficacy of BMS-986165 in Subjects with Moderate to Severe Ulcerative Colitis. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-004694-27	A5 Keine RCT
(21)	2018-005086-39	A Randomized, Placebo-Controlled, Double-Blind, Multicenter Study to Evaluate Efficacy and Safety of Oral BT-11 in Mild to Moderate Ulcerative Colitis. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-005086-39	A2 Andere Intervention
(22)	2019-002698-74	A PHASE 2B, MULTICENTER, RANDOMIZED, DOUBLEBLIND, PLACEBO CONTROLLED DOSE-RANGING STUDY TO EVALUATE THE EFFICACY, SAFETY, AND PHARMACOKINETICS OF PF-06480605 IN ADULT PARTICIPANTS WITH MODERATE TO [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2019-002698-74	A5 Keine RCT
(23)	2019-003999-39	A PHASE 2B, RANDOMIZED, DOUBLE-BLIND, PARALLEL-GROUP, PLACEBO-CONTROLLED STUDY WITH AN OPEN LABEL EXTENSION TO EVALUATE THE SAFETY AND EFFICACY OF PF-06826647 IN PARTICIPANTS WITH MODERATE TO SEVER [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2019-003999-39	A5 Keine RCT
(24)	2019-004224-38	A Phase 3 Study of the Efficacy, Safety and Pharmacokinetics of Ustekinumab as Open-label Intravenous Induction Treatment Followed by Randomized Double-blind Subcutaneous Ustekinumab Maintenance in [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2019-004224-38	A1 Andere Patientenpopulation
(25)	2019-004652-11	RANDOMIZED, OPEN-LABEL PHASE 3B STUDY OF ETROLIZUMAB BASED INDUCTION THERAPY COMBINATIONS FOLLOWED BY ETROLIZUMAB MAINTENANCE THERAPY IN PATIENTS WITH MODERATE-TO-SEVERE ULCERATIVE COLITIS. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2019-004652-11	A5 Keine RCT

Nr.	Studien-ID	Titel	Ausschlussgrund
(26)	2020-000047-31	A Randomized, Placebo-controlled, Double-blind, Parallel-group, Multicenter, Phase 2a Study of the Efficacy and Safety of Oral AMT-101 in Subjects with Moderate to Severe Ulcerative Colitis. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2020-000047-31	A5 Keine RCT
(27)	2020-001398-59	Randomized, double-blind, Phase 2 study to evaluate the efficacy and the safety of OSE-127 versus placebo in subjects with moderate to severe active ulcerative colitis who have failed or are intole [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2020-001398-59	A5 Keine RCT
(28)	2020-002833-13	A Randomized, Placebo-controlled, Double-blind, Parallel-group, Exploratory, Phase 2 Study of the Efficacy and Safety of Oral AMT-101 in Combination with Adalimumab in Subjects with Moderate to Sev [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2020-002833-13	A5 Keine RCT
(29)	2020-004457-76	A Phase 3, Multicenter, Open-label, Basket, Long-term Extension Study of Ustekinumab in Pediatric Clinical Study Participants (2 to <18 Years of Age). EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2020-004457-76	A1 Andere Patientenpopulation
(30)	2021-001940-86	A Phase 3 Study to Evaluate the Efficacy and Long-term Safety of SHR0302 for Induction, Maintenance Therapy in Subjects with Moderately to Severely Active Ulcerative Colitis. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2021-001940-86	A2 Andere Intervention
(31)	2021-002537-41	A Phase 2, Dose-finding, Randomized, Double-blind, Placebo-controlled, Multicenter Study to Evaluate the Safety and Efficacy of Efavaleukin Alfa Induction Therapy in Subjects with Moderately to Sev [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2021-002537-41	A5 Keine RCT

Nr.	Studien-ID	Titel	Ausschlussgrund
(32)	2021-003050-23	A Phase 2, Multicenter, Randomized, Double-Blind,Placebo-Controlled, Parallel-Group Study to Evaluate the Clinical Efficacy and Safety of VTX002 in Subjects with Moderately to Severely Active Ulcer [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2021-003050-23	A5 Keine RCT
(33)	2021-003702-42	AN INTERVENTIONAL PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO CONTROLLED, PARALLEL GROUP STUDY OF THE EFFICACY AND SAFETY OF RITLECITINIB (PF-06651600) IN ADULT PARTICIPANTS WITH MODERATELY TO SEVER [...]. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2021-003702-42	A2 Andere Intervention
(34)	2022-000894-16	Predicting Response to Iron Supplementation in Patients with Inflammatory Bowel Disease during Induction Therapy. EU Clinical Trials Register. 0000. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2022-000894-16	A2 Andere Intervention
ICTRP			
(35)	DRKS00019082	Analysis of clinical, genetical, immunological and microbial factors in IBD patients to identify new prognostic and predictive factors to further optimize and personalize IBD management. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00019082	A5 Keine RCT
(36)	DRKS00022043	Effectiveness of Ustekinumab in Induction and Maintenance Therapy of ulcerative colitis (RUN-UC). ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00022043	A5 Keine RCT
(37)	EUCTR2018-002925-47-ES	prediction of response to therapy in inflammatory bowel disease. ICTRP. 2019. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-002925-47-ES	A2 Andere Intervention
(38)	EUCTR2019-004224-38-DE	A Clinical Study of the Efficacy, Safety and Pharmacokinetics of Ustekinumab as Intravenous Induction Treatment Followed by Subcutaneous Ustekinumab Maintenance in Pediatric Participants with Moderately to Severely Active Ulcerative Colitis. ICTRP. 2021. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2019-004224-38-DE	A1 Andere Patientenpopulation

Nr.	Studien-ID	Titel	Ausschlussgrund
(39)	EUCTR2019-004652-11-SI	A Study of Etrolizumab-Based Induction Therapy Combinations Followed by Etrolizumab Maintenance Therapy in Patients with Moderate-To-Severe Ulcerative Colitis. ICTRP. 2020. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2019-004652-11-SI	A2 Andere Intervention
(40)	EUCTR2020-004457-76-FR	A Long-term Extension Study of Ustekinumab in Pediatric Participants with different medical conditions. ICTRP. 2021. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2020-004457-76-FR	A1 Andere Patientenpopulation
(41)	EUCTR2020-004457-76-PL	A Long-term Extension Study of Ustekinumab in Pediatric Participants with different medical conditions. ICTRP. 2021. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2020-004457-76-PL	A1 Andere Patientenpopulation
(42)	JPRN-jRCT1031200329	UC Bio Study. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-jRCT1031200329	A2 Andere Intervention
(43)	JPRN-jRCT2041200098	A Study of Ustekinumab in Pediatric Participants with Moderately to Severely Active Ulcerative Colitis (UC). ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-jRCT2041200098	A1 Andere Patientenpopulation
(44)	NCT03358706	A Study to Evaluate the Effect of Ustekinumab on Cytochrome P450 Enzyme Activities Following Induction and Maintenance Dosing in Participants With Active Crohn's Disease or Ulcerative Colitis. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03358706	A5 Keine RCT
(45)	NCT03606499	Real-world Effectiveness of Ustekinumab in Participants Suffering From Inflammatory Bowel Disease (Crohn's Disease or Ulcerative Colitis) With Extra-intestinal Manifestations or Immune-mediated Inflammatory Diseases. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03606499	A1 Andere Patientenpopulation
(46)	NCT03885713	Identification of Predictive Biomarkers for Response to Biologic Therapies in Inflammatory Bowel Disease. ICTRP. 2021. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03885713	A2 Andere Intervention

Nr.	Studien-ID	Titel	Ausschlussgrund
(47)	NCT04372108	A Study to Assess the Long-Term Safety of Ustekinumab Versus Other Biologics in Patients With Crohn's Disease and Ulcerative Colitis. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04372108	A5 Keine RCT
(48)	NCT04630028	A Study of Ustekinumab in Pediatric Participants With Moderately to Severely Active Ulcerative Colitis (UC). ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04630028	A1 Andere Patientenpopulation
(49)	NCT04963725	A Study of Symptom Improvement of Ulcerative Colitis After an Induction Dose of Ustekinumab in Japanese Clinical Practice. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04963725	A6 Andere Studiendauer
(50)	NCT05092269	A Long-term Extension Study of Ustekinumab in Pediatric Participants. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT05092269	A1 Andere Patientenpopulation
(51)	NCT05186623	Prediction Model for Response to Biologics and Small Molecular Agent for UC. ICTRP. 2022. [Zugriffsdatum: 02.06.2022]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT05186623	A5 Keine RCT

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-141 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 4-141 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-141 (Anhang): Studiendesign und -methodik für Studie U-ACHIEVE Substudie 3

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Bewertung der Wirksamkeit und Sicherheit der Behandlung mit Upadacitinib 30 mg und Upadacitinib 15 mg (einmal täglich) im Vergleich mit Placebo hinsichtlich des Erreichens einer klinischen Remission, nach dem Adaptierten Mayo Score, bei Patienten mit mittelschwerer bis schwerer aktiver Colitis Ulcerosa, die nach der Induktionsphase der U-ACHIEVE (M14-234) Substudien 1 und 2 oder der Studie U-ACCOMPLISH (M14-675) ein klinisches Ansprechen, ebenfalls nach dem Adaptierten Mayo Score, erreicht hatten.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Eine multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Substudie mit zwei vorgelagerten Phase-II/b- und Phase-III-Substudien (U-ACHIEVE) und einer weiteren vorgelagerten Phase-III-Studie (U-ACCOMPLISH). Die Studie verfügt über vier Kohorten 1, 2, 3 und 4. Patienten in Kohorte 1 wurden in eine von drei Behandlungsgruppen mit Zuteilungsverhältnis 1:1:1 und jeweils ca. 150 Patienten eingeteilt: • Upadacitinib 15 mg (eimal täglich) • Upadacitinib 30 mg (einmal täglich) • Placebo (einmal täglich) Patienten in Kohorte 2 wurde nur ein Placebo (einmal täglich) verabreicht. In den Kohorten 3 und 4 wurden die Patienten hingegen mit Upadacitinib behandelt, wobei Patienten in Kohorte 3 in zwei Behandlungsgruppen eingeteilt wurden, Upadacitinib 15 mg (täglich) und Upadacitinib 30 mg (einmal täglich), während die Patienten in Kohorte 4 als eine Behandlungsgruppe lediglich Upadacitinib 7,5 mg (einmal täglich) erhielten.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<u>Wesentliche Änderungen aus Amendment 1 (29. September 2016):</u> • Korrektur der Definition des klinischen Ansprechens nach dem Adaptierten Mayo Score • Auschluss von Patienten, die traditionelle chinesische Arzneimittel innerhalb von 28 Tagen vor der Baseline und während der Studie erhalten haben • Enfernung des Anteils der Patienten mit einem Stool Frequency Subscore ≤ 1 und des Anteils der Patienten mit einem Rectal Bleeding Subscore von 0 als multiplizitätskontrollierte sekundäre Endpunkte und Einstufung als nicht multiplizitätskontrollierte sekundäre Endpunkte • Hinzufügen der Begründung für die Annahme einer durchschnittlichen Induktionsansprechrate von 60 % für Upadacitinib-Dosen bei der Bestimmung der Fallzahl <u>Wesentliche Änderungen aus Amendment 2 (10. Oktober 2017):</u> • Es wurden keine Änderungen bzgl. der Methodik in der Substudie 3 vorgenommen

Item[a]	Charakteristikum	Studieninformation
		Wesentliche Änderungen aus Amendment 3 (03. Juli 2018): • Hinzufügen von Kohorte 3 in Substudie 3 für Patienten, die für einen Zeitraum von 16 Wochen mit Upadacitinib 45 mg behandelt wurden und am Ende der erweiterten Behandlungsperiode zu Woche 16 ein klinisches Ansprechen erzielt hatten (Studie M14-234 Substudie 2 Teil 2 und Studie M14-675 Teil 2) • Hinzufügen der erweiterten Behandlungsperiode in Substudie 2, um Placebo-Patienten aus der 8-wöchigen Induktionsphase, die kein klinisches Ansprechen erzielt haben, eine Induktionsphase mit Upadacitinib anzubieten und das verzögerte klinische Ansprechen bei klinischen Non-Responder-Patienten aus der 8-wöchigen Induktionsphase zu evaluieren • Aktualisierung der Fallzahl auf Grundlage der Ergebnisse der Phase 2b • Verlängerung der Dauer der Maintenance-Periode von 44 auf 52 Wochen • Aktualisierung der primären und sekundären Endpunkte basierend auf der Richtlinien der Food and Drug Administration • Modifizierung der Stratifizierungsfaktoren Wesentliche Änderungen aus Amendment 4 (24. April 2019): • Änderung des Ausschlusskriteriums Nr. 14, um zu ermöglichen, dass nicht nachweisbare biologische Konzentrationen mit einem kommerziell erhältlichen Assay gemessen werden dürfen (als Alternative zur Washout-Periode) basierend auf dem Feedback der Behörden • Hinzufügen der länderspezifischen Anforderungen in den Einschluss-/Ausschlusskriterien (Einschluss von 16- und 17-jährigen Patienten, sofern lokal zulässig) • Hinzufügen einer Rechtfertigung für die Verwendung von Placebos Wesentliche Änderungen aus Amendment 5 (29. April 2020): • Aktualisierung der Kriterien für das klinische Ansprechen zur Aufnahme in die verlängerte Behandlungsperiode für Patienten mit fehlender Endoskopie zu Woche 8, wenn die Endoskopie aufgrund der COVID-19-Pandemie nicht durchgeführt werden kann • Neueinstufung von nicht multiplizitätskontrollierten sekundären Wirksamkeitsvariablen als zusätzliche Wirksamkeitsvariablen • Ausschluss von Patienten mit einer früheren Perforation des Magen-Darm-Trakts (eine Perforation aufgrund einer

Item[a]	Charakteristikum	Studieninformation
		mechanischen Verletzung sollte einen Patienten von der Teilnahme an der Studie nicht ausschließen) • Verbot einer Zytapherese-Behandlung innerhalb der letzten 60 Tage vor dem Screening • Auschluss von Patienten mit früheren thrombotischen Ereignissen, einschließlich tiefer Venenthrombose und Lungenembolie oder bekannten Erbkrankheiten, die zu Hyperkoagulabilität prädisponieren • Aktualisierung der Toxizitätsmanagement für Aspartat-Transaminase (AST) und Alanin-Transaminase (ALT), Thromboseereignisse und Herpes zoster <u>Wesentliche Änderungen aus Amendment 6 (31. Juli 2020):</u> • Es wurden keine Änderungen bzgl. der Methodik in der Substudie 3 vorgenommen <u>Wesentliche Änderungen aus Amendment 7 (10. Mai 2021):</u> • Entfernung von CU-bedingten Krankenhausaufenthalten und CU-bedingten Operationen aus den multiplizitätskontrollierten sekundären Endpunkten • Präzisierung der primären ITT-Populationen sowie der ITT-Populationen der Substudie 3 • Hinzufügen der Non-Responder-Imputation (NRI) unter Einbeziehung der multiplen Imputation (MI) zur Handhabung fehlender Daten aufgrund von COVID-19 (NRI-C) und MMRM-Methoden, um die Klarheit der statistischen Analyse zu gewährleisten • Hinzufügen einer neuen Subgruppenanalyse des primären Endpunkts: Aminosalicylat-Einnahme zur Baseline (ja, nein)
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien:</u> 1. Männliche oder weibliche Patienten im Alter von ≥ 16 und ≤ 75 Jahren zur Baseline. • Heranwachsende Patienten im Alter von 16 und 17 Jahren werden eingeschlossen, sofern das Land oder die Aufsichts-/Gesundheitsbehörde dies genehmigt. Wenn diese Genehmigungen nicht vorliegen, werden nur Patienten im Alter von ≥ 18 Jahren aufgenommen. • Heranwachsende Patienten im Alter von 16 und 17 Jahren müssen ≥ 40 kg wiegen und die Definition des Tanner-Stadiums 5 beim Screening erfüllen. 2. Diagnose von Colitis ulcerosa 90 Tage oder früher vor der Baseline, bestätigt durch Koloskopie während des Screening, mit Ausschluss einer aktuellen Infektion, Kolon-Dysplasie und/oder Malignität. Eine angemessene Dokumentation der Biopsieergebnisse, die nach Ermessen des Prüfarztes mit der Diagnose von Colitis ulcerosa übereinstimmen, müssen vorgelegt werden.

Item[a]	Charakteristikum	Studieninformation
		3. Aktive Colitis ulcerosa mit einem Adaptierten Mayo Score von 5 bis 9 Punkten und einem Endoskopie-Subscore von 2 bis 3. 4. Nachgewiesenes unzureichendes Ansprechen, Ausbleiben des Ansprechens auf oder Unverträglichkeit gegen mindestens eine der Behandlungen, einschließlich orale Aminosalicylate, Kortikosteroide, Immunsuppressiva und/oder biologische Therapien, nach Ermessen des Prüfarztes. 5. Weibliche Patienten im gebärfähigen Alter müssen einen negativen Serum-Schwangerschaftstest beim Screening und einen negativen Urintest bei der Erstuntersuchung vor Verabreichung der Studienmedikation vorlegen. 6. Weibliche Patienten müssen während der Behandlung mit der Studienmedikation und für mindestens 30 Tage nach Verabreichung der letzten Dosis der Studienmedikation eine Schwangerschaft vermeiden. Dazu müssen hochwirksame Verhütungsmethoden angewendet werden. 7. Patienten müssen sich nach Ermessen des Prüfarztes in einem ansonsten guten Gesundheitszustand basierend auf der Ergebnisse der Anamnese, des Laborprofils, der körperlichen Untersuchung sowie einem 12-Kanal-Elektrokardiogramm (EKG), das während des Screenings durchgeführt wird. 8. Patienten müssen sowohl in der Lage als auch bereit sein, eine schriftliche Einwilligungserklärung zu erteilen und die Anforderungen des Studienprotokolls zu erfüllen. In Japan müssen Patienten, die jünger als 20 Jahre alt sind, eine schriftliche Einwilligungserklärung der Eltern oder der Erziehungsberechtigten vorlegen. <u>Ausschlusskriterien:</u> 1. Patienten mit aktueller Diagnose von Morbus Crohn oder Colitis indeterminata. 2. Aktuelle Diagnose einer fulminanten Kolitis und/oder eines toxischen Megakolons. 3. Patienten mit einer auf das Rektum beschränkten Erkrankung (ulzerative Proktitis) bei der Screening-Endoskopie. 4. Eine frühere Kolektomie (vollständig oder partiell), Ileoanal-Pouches, Kock-Pouches, Ileostomie oder die Planung einer Darmoperation. 5. Behandlung mit rektalen Aminosalicylaten oder Kortikosteroiden, anderen Einläufen/Zäpfchen (sofern nicht für die Endoskopie erforderlich) innerhalb von 14 Tagen vor der Screening-Endoskopie und während der restlichen Screening-Periode. 6. Einnahme von Cyclosporin, Tacrolimus, Mycophenolatmofetil oder Thalidomid binnen 30 Tage vor der Baseline.

Item[a]	Charakteristikum	Studieninformation
		7. Patienten, die Azathioprin oder 6-Mercaptopurin binnen 10 Tage vor der Baseline erhielten. 8. Intravenöse Kortikosteroide binnen 14 Tage vor dem Screening oder während der Screening-Periode. 9. Patienten, die Methotrexat oder orale Aminosalicylate erhalten, die: • Die aktuelle Methotrexat-Behandlung nicht seit mindestens 42 Tagen vor der Baseline erhalten und diese nicht mindestens seit 28 Tagen vor der Baseline in stabiler Dosierung einnehmen; • Nicht seit mindestens 14 Tagen vor der Baseline stabile Dosen von oralen Aminosalicylaten erhalten; • Die Behandlung mit Aminosalicylaten binnen 14 Tage vor der Baseline abgesetzt haben. 10. Patienten, die mit Kortikosteroiden behandelt werden, die: • Eine orale Kortikosteroiddosis > 30 mg/Tag (Prednison oder Äquivalent) erhalten oder die Behandlung nicht seit mindestens 14 Tagen vor der Baseline und diese nicht in einer stabilen Dosis für mindestens 7 Tage vor der Baseline erhalten; • Eine orale Budesonid-Dosis > 9 mg/Tag erhalten oder die aktuelle Behandlung nicht mindestens 14 Tage vor der Baseline in einer stabilen Dosis für mindestens 7 Tage vor der Baseline erhalten; • Eine orale Beclomethason-Dosis > 5 mg/Tag erhalten oder die aktuelle Behandlung nicht seit mindestens 14 Tagen erhalten und diese nicht seit mindestens 7 Tagen vor der Baseline in stabiler Dosis einnehmen; • sowohl orales Budesonid (oder orales Beclomethason) als auch orales Prednison (oder Äquivalent) gleichzeitig eingenommen haben, mit Ausnahme von topischen oder Inhalatoren binnen 14 Tage vor dem Screening oder während der Screening-Periode. 11. Der Patient leidet unter Tuberkulose oder erfüllt die Tuberkulose-Ausschlussparameter. 12. Patienten, die innerhalb von 30 Tagen vor der Baseline eine fäkale mikrobielle Transplantation erhalten haben. 13. Patienten, die CU-bezogene Antibiotika einnehmen, die nicht mindestens 14 Tage vor der Baseline in stabiler Dosis eingenommen wurden oder diese Medikamente binnen 14 Tage nach der Baseline abgesetzt haben. 14. Patienten, die eine der folgenden biologischen Therapien erhalten haben: • Infliximab, Certolizumab, Adalimumab, Golimumab, Vedolizumab, Natalizumab, binnen 8 Wochen vor der Baseline; • Oder Ustekinumab binnen 12 Wochen vor der Baseline

Item[a]	Charakteristikum	Studieninformation
		15. Patienten mit früherer Exposition gegenüber einem Januskinase-Inhibitor, wie z. B. Tofacitinib, Baricitinib, Filgotinib, Upadacitinib. 16. Patienten, die nicht steroidale Antirheumatika erhalten haben (außer topische nicht steroidale Antirheumatika und die Einnahme von niedrig dosiertem Aspirin zum Schutz des Herz-Kreislauf-Systems) binnen 7 Tage vor der Baseline. 17. Der Patient hat binnen 30 Tage vor der Baseline traditionelle chinesische Medizin erhalten. 18. Der Patient erhielt Lebendimpfstoff(e) binnen 30 Tage (8 Wochen für Japan) vor der Baseline oder erwartet, dass er während der Studienteilnahme, einschließlich mindestens 30 Tage (8 Tage für Japan) oder länger (falls erforderlich) nach der letzten Dosis der Studienmedikation eine Lebendimpfung benötigt. 19. Systemische Anwendung von bekannten starken Cytochrom P450 (CYP)3A-Inhibitoren oder starken CYP3A-Induktoren während der Screening-Periode und bis zum Ende der Studie. 20. Patienten, die derzeit eine totale parenterale Ernährung (TPN) erhalten oder planen, eine TPN zu einem beliebigen Zeitpunkt während des Studienzeitraums. 21. Patienten, die binnen 30 Tage oder 5 Halbwertszeiten vor der Baseline ein Prüfpräparat erhalten haben oder derzeit an einer Interventionsstudie teilnehmen. 22. Patienten mit positivem C-difficile-Toxin-Stuhltest während des Screenings. 23. Infektion(en), die eine Behandlung mit intravenösen Antiinfektiva binnen 30 Tage vor dem Baseline-Besuch oder orale/intramuskuläre Antiinfektiva binnen 14 Tage vor dem Baseline-Besuch erfordern. 24. Eine chronisch wiederkehrende Infektion und/oder eine aktive Virusinfektion, die, nach der klinischen Einschätznug des Prüfarztes, den Patienten zu einem ungeeigneten Kandidaten für die Studie macht. 25. Der Patient leidet oder litt unter rezidivierender oder disseminierter Gürtelrose (Herpes zoster). 26. Der Patient leidet oder litt unter disseminiertem Herpes simplex. 27. Der Patient hat eine Infektion mit Hepatitis B, Hepatitis C oder dem humanen Immundefizienzvirus (HIV). 28. Frühere oder aktuelle gastrointestinale Dysplasien (mit Ausnahme von einer vollständig entfernten, niedriggradigen dysplastischen Läsion) in jeglicher Biopsie, die während oder vor der Screening-Endoskopie durchgeführt wurde. 29. Frühere bösartige Erkrankungen (mit Ausnahme von erfolgreich behandeltem Nicht-Melanom-Hautkrebsoder lokalisiertem in-situ-Carcinoma des Gebärmutterhalses). 30. Spontane gastrointestinale Perforation (mit Ausnahme von Appendizitis oder einer mechanischen Verletzung),

Item[a]	Charakteristikum	Studieninformation
		Divertikulitis oder signifikant erhöhtes Risiko einer Magen-Darm-Perforation nach Ermessen des Prüfarztes. 31. Screening-Untersuchungen und andere Analysen zeigen eines der folgenden abnormen Ergebnisse: • Serum-Aspartat-Transaminase oder Alanin-Transaminase > 2 × des oberen Normwertes (ULN); • Geschätzte glomeruläre Filtrationsrate nach simplifizierter 4-Variablen-MDRD-Formel < 30 ml/min/1,73 m^2; • Gesamtzahl der weißen Blutkörperchen < 2.500/µl; • Absolute Neutrophilenzahl < 1.200/µl; • Thrombozytenzahl < 100.000/µl; • Absolute Lymphozytenzahl < 750/µl; • Hämoglobin < 9 g/dl. 32. Patientinnen, die schwanger sind, stillen oder eine Schwangerschaft während der Studie und innerhalb von 30 Tagen nach der letzten Dosis des Studienmedikation in Betracht ziehen oder jene, die einen positiven Schwangerschaftstest bei dem Screening (Serum) oder zur Baseline (Urin) haben. 33. Allergische Reaktionen oder signifikante Überempfindlichkeit gegenüber Bestandteilen der Studienmedikation (und ihrer Hilfsstoffe) und/oder auf andere vergleichbare Produkte. 34. Klinisch signifikanter (nach Ermessen des Prüfarztes) Drogen- oder Alkoholmissbrauch in den letzten 6 Monaten. 35. Patienten, die zuvor eine Stammzellentransplantation erhalten haben. 36. Patienten, die eine Organtransplantation erhalten haben, die eine fortgesetzte Immunsuppression erfordert. 37. Nur für japanische Patienten: positives Ergebnis von Beta-D-Glucan oder zwei aufeinanderfolgende unbestimmte Ergebnisse von Beta-D-Glucan (Screening auf eine Pneumocystis-jiroveci-Infektion). 38. Zytapheresebehandlung (GCAP, LCAP usw.) binnen 60 Tage vor dem Screening. 39. Nur für japanische Patienten: ATM-Behandlung (antibiotische Kombinationstherapie mit Amoxicillin, Tetracyclin und Metronidazol) während der Screening-Periode. 40. Eine der folgenden kardiovaskulären oder thrombotischen Erkrankungen: • Kürzlicher (innerhalb der letzten 6 Monate) zerebrovaskulärer Unfall, Myokardinfarkt oder Koronarstenting • Kürzlich (innerhalb der letzten 6 Monate) aufgetretene mäßige bis schwere kongestive Herzinsuffizienz (New York Heart Association Klasse III oder IV);

Item[a]	Charakteristikum	Studieninformation
		• Frühere thrombotische Ereignisse, einschließlich eine tiefe Venenthrombose und eine Pulmonalembolie • Bekannte Erbkrankheiten, die zu einer Hyperkoagulabilität prädisponieren 41. Zustände, die die Absorption von Arzneimitteln beeinträchtigen könnten, einschließlich, aber nicht nur Kurzdarmsyndrom oder Magenbypass-Operation; Patienten mit früherem Magenband oder früherer Segmentierung sind nicht ausgeschlossen. 42. Vorheriger klinisch signifikanter medizinischer Zustand oder ein anderer Grund, der nach Ermessen des Prüfarztes die Teilnahme des Patienten an der Studie beeinträchtigen würde, den Patienten zu einem ungeeigneten Kandidaten für die Behandlung mit der Studienmedikation machen würde oder den Patienten durch die Teilnahme am Prüfverfahren einem Risiko aussetzen würde.
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie mit 302 Studienzentren in 43 Ländern (Argentinien, Australien, Belgien, Bosnien und Herzegowina, Brasilien, Kanada, Chile, China, Kolumbien, Kroatien, Tschechische Republik, Estland, Finnland, Frankreich, Deutschland, Griechenland, Ungarn, Irland, Israel, Italien, Japan, Südkorea, Lettland, Litauen, Malaysia, Mexiko, Niederlande, Norwegen, Österreich, Polen, Portugal, Russische Föderation, Serbien, Singapur, Slowakei, Südafrika, Spanien, Schweiz, Taiwan, Türkei, Vereinigtes Königreich, Ukraine, Vereinigte Staaten und ihr Territorium Puerto Rico)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Die Patienten in den vier Kohorten der Studie erhielten je nach Behandlungsgruppe pro Tag eine Tablette der Studienmedikation (Upadacitinib 7,5 mg oder Upadacitinib 15 mg oder Upadacitinib 30 mg oder ein entsprechendes Placebo). Die Interventionen variierten je nach Behandlungsgruppe, wie folgt: • Patienten, die Upadacitinib 30 mg oder 45 mg in der Substudie 1 erhielten und ein klinisches Ansprechen erzielten, werden erneut in eine der drei Behandlungsgruppen der Kohorte 1 randomisiert (Upadacitinib 15 mg oder Upadacitinib 30 mg oder Placebo) und erhalten eine Tablette der Studienmedikation pro Tag. • Patienten, die Upadacitinib 15 mg in der Substudie 1 erhielten und ein klinisches Ansprechen erzielten, werden erneut in eine von zwei Behandlungsgruppen der Kohorte 1 randomisiert (Upadacitinib 15 mg oder Placebo) und erhalten eine Tablette der Studienmedikation pro Tag. • Patienten, die Upadacitinib 7,5 mg oder Placebo erhielten und ein klinisches Ansprechen in der Substudie 1 erzielten, werden in der Studie mit der selben Behandlung fortfahren, erhalten jedoch nur eine Tablette der Studienmedikation pro Tag anstelle von zwei Tabletten. Dabei werden die Patienten, die Placebo erhielten, Kohorte 2 zugeordnet und die Patienten, die Upadacitinib 7,5 mg erhielten, Kohorte 4 zugeordnet. • Patienten der Substudie 2 Teil 1 und der Studie M14-675 Teil 1, die Upadacitinib 45 mg erhielten und ein klinisches

Item[a]	Charakteristikum	Studieninformation
		Ansprechen erzielten, werden randomisiert einer der drei Behandlungsgruppen der Kohorte 1 zugeteilt (Upadacitinib 15 mg oder Upadacitinib 30 mg oder Placebo) und erhalten eine Tablette der Studienmedikation pro Tag. • Patienten der Substudie 2 Teil 2 und der Studie M14-675 Teil 2, die in Teil 1 Placebo und in Teil 2 Upadacitinib erhielten und ein klinisches Ansprechen erzielten, werden randomisiert in eine der drei Behandlungsgruppen der Kohorte 1 eingeteilt und erhalten eine Tablette der Studienmedikation pro Tag. • Patienten der Substudie 2 Teil 1 und der Studie M14-675 Teil 1, die Placebo erhielten und ein klinisches Ansprechen erzielten, werden in Kohorte 2 der Studie weiterhin Placebo erhalten. • Patienten der Substudie 2 Teil 2 und der Studie M14-675 Teil 2, die in Teil 1 Upadacitinib und in Teil 2 Upadacitinib erhielten und ein klinisches Ansprechen erzielten, werden randomisiert in eine der zwei Behandlungsgruppen der Kohorte 3 eingeteilt (Upadacitinib 15 mg und 30 mg) und erhalten täglich eine Tablette der Studienmedikation.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<u>Primärer Endpunkt:</u> Anteil der Patienten, die zu Woche 52 eine klinische Remission nach dem Adaptierten Mayo Score erreichen <u>Sekundäre Endpunkte:</u> • Anteil der Patienten mit endoskopischer Verbesserung • Anteil der Patienten, die eine klinische Remission nach dem Adaptierten Mayo Score erreichen unter den Patienten, die eine klinische Remission nach dem Adaptierten Mayo Score in der Studie M14-234 (Substudie 1 oder 2) oder Studie M14-675 • Anteil der Patienten, die zu Woche 52 eine klinische Remission nach dem Adaptierten Mayo Score erreichten und für 90 Tage oder mehr kortikosteroidfrei waren, unter den Patienten in klinischer Remission am Ende der Induktionsphase der Studie M14-234 (Substudie 1 oder 2) oder der Studie M14-675 • Anteil der Patienten mit endoskopischer Verbesserung unter den Patienten mit endoskopischer Verbesserung in Studie M14-234 (Substudie 1 oder 2) oder Studie M14-675 • Anteil der Patienten mit endoskopischer Remission • Anteil der Patienten mit klinischem Ansprechen gemäß Adaptiertem Mayo Score • Anteil der Patienten, die eine histologisch-endoskopische Verbesserung der Schleimhaut erreichen • Veränderung des IBDQ-Gesamtscore im Vergleich zur Baseline • Anteil der Patienten mit Schleimhautheilung • Anteil der Patienten, die keinen Stuhldrang angeben

Item[a]	Charakteristikum	Studieninformation
		• Anteil der Patienten, die von keinen Unterleibsschmerzen berichten • Veränderung des FACIT-F-Scores im Vergleich zur Baseline <u>Sicherheitsendpunkte:</u> Die folgenden Endpunkte werden in die Sicherheitsanalysen einbezogen: • Therapiebedingte unerwünschte Ereignisse • Schwerwiegende unerwünschte Ereignisse • Unerwünschte Ereignisse von speziellem Interesse (UESI) • Unerwünschte Ereignisse, die zum Therapieabbruch führen • Vitalparameter und Laboruntersuchungen
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<u>Amendment 1 (29. September 2016):</u> • Korrektur der Definition des klinischen Ansprechens nach dem Adaptierten Mayo Score • Enfernung des Anteils der Patienten mit einem Stool-Frequency-Subscore ≤ 1 und des Anteils der Patienten mit einem Rectal-Bleeding-Subscore von 0 als multiplizitätskontrollierte sekundäre Endpunkte und Einstufung als nicht multiplizitätskontrollierte sekundäre Endpunkte <u>Amendment 3 (03. July 2018):</u> • Aktualisierung der primären und sekundären Endpunkte basierend auf der Richtlinien der Food-and-Drug-Administration • Verlängerung der Dauer der Maintenance-Periode von 44 auf 52 Wochen • Modifizierung der Stratifizierungsfaktoren <u>Amendment 5 (29. April 2020):</u> • Neueinstufung von nicht multiplizitätskontrollierten sekundären Wirksamkeitsvariablen als zusätzliche Wirksamkeitsvariablen <u>Amendment 7 (10. Mai 2021):</u> • Entfernung von UC-bedingten Krankenhausaufenthalten und UC-bedingten Operationen aus den multiplizitätskontrollierten sekundären Endpunkten • Präzisierung der primären ITT-Populationen sowie der ITT-Populationen der Substudie 3 • Hinzufügen einer neuen Subgruppenanalyse des primären Endpunkts: Aminosalicylat-Einnahme zur Baseline (ja, nein)
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl der Studie basiert auf dem erwarteten Anteil der Patienten, die nach dem Adaptierten Mayo Score nach Woche 52 eine klinische Remission erzielen. Unter der Annahme einer klinischen Remissionsrate von 12 % in der Placebogruppe und 40 % in einer der Behandlungsgruppen mit Upadacitinib (einmal

Item[a]	Charakteristikum	Studieninformation
		täglich) zu Woche 52, ergibt sich, dass eine Fallzahl von 150 Patienten in der Placebogruppe und 150 Patienten in jeder der Behandlungsgruppen mit Upadacitinib 15 mg und Upadacitinib 30 mg (jeweils einmal täglich) eine Power von > 95 % hat, um den Behandlungsunterschied von 28 % im primären Endpunkt zwischen einer Upadacitinib-Dosis und Placebo unter Verwendung des zweiseitigen exakten Tests von Fisher auf einem Signifikanzniveau von 0,025 mit Multiplizitätsanpassung aufzudecken. Unter der Annahme, dass die durchschnittliche Ansprechrate bei Upadacitinib-Dosen nach Abschluss der Substudien 1 und 2 50 % beträgt, werden circa 450 Patienten wie folgt erneut randomisiert: Die Patienten werden Behandlungsgruppen mit Upadacitinib 15 mg oder Upadacitinib 30 mg (einmal täglich) oder Placebo mit einem Verhältnis von 1:1:1 zugeteilt, wenn sie in der Substudie 1 oder der Substudie 2 oder der Studie M14-675 auf Upadacitinib 45 mg (einmal täglich) angesprochen haben. Die Annahme einer durchschnittlichen Ansprechrate von 50 % bei den Upadacitinib-Dosen nach der Induktion basiert auf den Ergebnissen der Phase 2b.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurden keine Zwischenanalysen für die Wirksamkeitsabschätzung geplant.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Alle Patienten werden mittels webbasierten IRT zentral randomisiert.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Substudie 3 umfasst insgesamt vier Kohorten. Die Behandlungszuweisung in der Studie hängt von der Behandlung in den Substudien 1 und 2 oder der Studie M14-675 ab. Bei Patienten, die für eine erneute Randomisierung in eine der drei Behandlungsgruppen der Kohorte 1 geeignet sind, wird die Randomisierung nach vorherigem Behandlungsversagen (Bio-IR oder Non-Bio-IR), klinischem Remissionsstatus zu Woche 0 der Substudie 3 (ja oder nein), erhaltener Dosis in der Induktionsphase (gilt nur für Patienten aus Substudie 1) und dem Einsatz von Kortikosteroid zu Woche 0 der Substudie 3 (ja oder nein) stratifiziert. Für Patienten, die für eine erneute Randomisierung in eine der zwei Behandlungsgruppen der Kohorte 3 geeignet sind, wird die Randomisierung nach vorherigem Behandlungsversagen (Bio-IR oder Non-Bio-IR), klinischem Remissionsstatus zu Woche 0 der Substudie 3 (ja oder nein) und dem Einsatz von Kortikosteroid zu Woche 0 der Substudie 3 (ja oder nein) stratifiziert. <u>Analyse Populationen:</u> 1. <u>Intent-to-Treat (ITT)-Populationen:</u>

Item[a]	Charakteristikum	Studieninformation
		• ITT-Population: Alle Patienten, die mindestens eine Dosis der Studienmedikation in der Maintenance-Studie erhalten haben. • ITT_A-Population: Die Subgruppe der ITT-Population, die die ersten 450 Patienten umfasst, die acht Wochen Upadacitinib 45 mg (einmal täglich) erhielten, auf die Induktionsphase ansprachen und gemäß dem Protokoll in der 52-wöchigen Maintenance-Therapie in Kohorte 1 eingeschrieben waren. Die ITT_A-Population ist die primäre Analysepopulation in Kohorte 1 für die Wirksamkeitsendpunkte. • ITT_B-Population: Die Subgruppe der ITT-Population in Kohorte 3, die alle Patienten umfasst, die 16 Wochen Upadacitinib 45 mg (einmal täglich) erhielten und auf die Induktionsbehandlung ansprachen. • ITT_C-Population: Die Subgruppe der ITT-Population, die alle Patienten umfasst, die gemäß dem ursprünglichen Protokoll (Amendment 1 oder 2) in der 44-wöchigen Maintenance-Therapie eingeschrieben waren. • ITT_D-Population: Die Subgruppe der ITT-Population, die alle Patienten umfasst, die auf die 8-wöchige Induktionsphase mit Upadacitinib 45 mg (einmal täglich) ansprachen und gemäß dem Protokoll in der 52-wöchigen Maintenance-Therapie in Kohorte 1 eingeschrieben waren. • ITT_E-Population: Die Subgruppe der ITT-Population, die alle Patienten umfasst, die auf die 8-wöchige Induktionsphase mit Placebo, Upadacitinib 7,5 mg (einmal täglich), 15 mg (einmal täglich) oder 30 mg (einmal täglich) ansprachen und gemäß dem Protokoll in der 52-wöchigen Maintenance-Therapie eingeschrieben waren. 2. <u>Safety-Populationen:</u> • SA-Population: Alle Patienten, die mindestens eine Dosis der Studienmedikation in der Maintenance-Studie erhalten haben. • SA_A-Population: Die Subgruppe der SA-Population, die die ersten 450 Patienten umfasst, die auf die 8-wöchige Induktionsphase mit Upadacitinib 45 mg (einmal täglich) ansprachen und gemäß dem Protokoll in der 52-wöchigen Maintenance-Therapie in Kohorte 1 eingeschrieben waren. • SA_B-Population: Die Subgruppe der SA-Population in Kohorte 3, die alle Patienten umfasst, die auf die 16-wöchige Induktionsphase mit Upadacitinib 45 mg (einmal täglich) ansprachen.

Item[a]	Charakteristikum	Studieninformation
		• SA_C-Population: Die Subgruppe der SA-Population, die alle Patienten umfasst, die auf die 8-wöchige Induktionsphase mit Upadacitinib 45 mg (einmal täglich) ansprachen und gemäß dem Protokoll für eine 44- oder 52-wöchige Maintenance-Therapie in Kohorte 1 eingeschrieben waren. • SA_D-Population: Die Subgruppe der SA-Population, die alle Patienten umfasst, die auf die 8-wöchige Induktionsphase mit Placebo, Upadacitinib 7,5 mg (einmal täglich), 15 mg (einmal täglich) oder 30 mg (einmal täglich) ansprachen und in der 44- oder 52-wöchigen Maintenance-Therapie eingeschrieben waren. • SA_UPA-Population: Alle Patienten, die mindestens eine Dosis der Studienmedikation (Upadacitinib) in der Maintenance-Studie erhielten.
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	Das IRT stellt die entsprechende(n) Nummer(n) des Medikamenten-Kits zur Verfügung, die an jeden/jede Patienten/Patientin ausgegeben werden sollen. Die Studiemedikation wird bei den Studienbesuchen verabreicht.
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?	Das IRT vergibt beim Screening eine eindeutige Randomisierungsnummer, die die Zuordnung des/der Patienten/Patientin zur Behandlungsgruppe gemäß dem von der Statistikabteilung von AbbVie erstellten Randomisierungsplan. Alle Patienten behalten während der gesamten Studie dieselbe eindeutige Randomisierungsnummer.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Ja b) Ja c) Ja Der/die Prüfarzt/Prüfärztin, das Personal des Prüfzentrums, die Patienten und alle Mitarbeiter von AbbVie, die die direkte Durchführung und Verwaltung der Studie beaufsichtigen, mit Ausnahme des AbbVie Drug Supply Management Teams, bleiben verblindet während des gesamten Verlaufs der Studie.

Item[a]	Charakteristikum	Studieninformation
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Identisch aussehende Upadacitinib- und Placebo-Tabletten werden eingesetzt
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Alle Wirksamkeitsanalysen werden basierend auf der ITT_A-Population durchgeführt. Sowhl der primäre als auch die sekundären Wirksamkeitsendpunkte werden mit grafischer Multiplizitätsanpassung getestet, um eine strenge Kontrolle Fehler I Art auf dem Signifikanzniveau $\alpha = 0,05$ (zweiseitig) zu gewährleisten. Für Analysen, die auf der ITT_A-Population basieren, werden kategoriale Variablen, sofern nicht anders angegeben, mit dem Cochran-Mantel-Haenszel (CMH)-Test analysiert, stratifiziert nach Bio-IR-Status (Bio-IR oder Non-Bio-IR) zur Baseline der Induktionsstudie, klinischem Remissionsstatus zu Woche 0 (ja oder nein) sowie dem Einsatz von Kortikosteroid zur Woche 0 (ja oder nein). Jeder Patient, der unter dem falschen Stratum randomisiert wurde, wird nach dem tatsächlichen Stratum, zu dem er gehört, analysiert. Stetige Variablen, die longitudinal erhoben wurden, werden mit dem Mixed-Effect-Model-Repeat-Measurement-Modell (MMRM) analysiert. Stetige Variablen, die nur bei einem Besuch nach der Baseline erhoben wurden, werden mit einem Kovarianzanalyse-Modell (ANCOVA) analysiert. Das MMRM-Modell umfasst die kategorialen festen Effekte der Behandlung, Besuch und Behandlung-Besuch-Interaktion, Randomisierungsstratifikationsfaktoren (Bio-IR-Status (Bio-IR oder Non-Bio-IR) zur Baseline der Induktionsstudie, den klinischen Remissionsstatus zu Woche 0 (ja oder nein) und den Einsatz von Kortikosteroid zu Woche 0 (ja oder nein)) sowie die stetigen festen Kovariaten der Baseline-Messungen. Es wird eine unstrukturierte Varianz-Kovarianz-Matrix verwendet. Falls das Modell nicht konvergieren kann, dann wird eine autoregressive Kovarianzstruktur-Matrix verwendet. Die Parameterschätzungen basieren auf der restriktiven Maximum-Likelihood-Methode (REML). Die festen Effekte werden verwendet, um modellbasierte Mittelwerte bei entsprechenden Besuchen zu vermelden. Das ANCOVA-Modell umfasst die kategorialen festen Effekte der Behandlung, Randomisierungsstratifizierungsfaktoren (Bio-IR-Status (Bio-IR oder Non-Bio-IR) zur Baseline der Induktionsstudie, klinischen Remissionsstatus zu Woche 0 (ja oder nein) und den Einsatz von Kortikosteroid zu Woche 0 (ja oder nein)) sowie die stetigen festen Kovariaten der Baseline-Messungen.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Der primäre Endpunkt wird für folgende Subgruppen der ITT_A-Population analysiert: • Geschlecht (männlich, weiblich) • Alter (≤ Median, > Median) • Abstammung (weiß, nicht-weiß) • Bio-IR-Status (Bio-IR, Non-Bio-IR) • Kortikosteroideinnahme zur Baseline (ja, nein) • Adaptierter Mayo Score zur Baseline (≤ 7, > 7)

Item[a]	Charakteristikum	Studieninformation
		• Full-Mayo-Score zur Baseline (≤ 9, > 9) • Vorherige Behandlung mit Anti-TNF bei Non-Bio-IR (ja, nein) • Vorherige Exposition gegenüber einer biologischen Therapie bei Non-Bio-IR (ja, nein) • Ausgangsgewicht (≤ Median, > Median) • Vorliegen einer Pankolitis zur Baseline (ja, nein) • Krankheitsdauer zur Baseline (≤ Median, > Median) • hs-CRP zur Baseline (≤ 5 mg/l und > 5 mg/l) • Region (USA, Nicht-USA) • Verwendung von Aminosalicylat zur Baseline (ja, nein)
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) Placebo: N = 386; Upadacitinib 7,5 mg: N = 20; Upadacitinib 15 mg: N = 324; Upadacitinib 30 mg: N = 316 b) Placebo: N = 385; Upadacitinib 7,5 mg: N = 20; Upadacitinib 15 mg: N = 323; Upadacitinib 30 mg: N = 316 c) ITT_A-Population: Placebo: N = 149; Upadacitinib 15 mg: N = 148 ; Upadacitinib 30 mg: N = 154 SA_C-Population: Placebo: N = 245; Upadacitinib 15 mg: N = 250; Upadacitinib 30 mg: N = 251
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Erster Patient erste Visite: 14. Dezember 2016 Letzter Patient letzte Visite: 13. Dezember 2021
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde planmäßig beendet.

a: nach CONSORT 2010.

ALT: Alanin-Transaminase; ANCOVA: Kovarianzanalyse; AST: Aspartat-Transaminase; Bio-IR: Intoleranz oder unzureichendes Ansprechen auf eine vorherige biologische Therapie; CMH: Cochran-Mantel-Haenszel; CONSORT: Consolidated Standards of Reporting Trials; CU: Colitis ulcerosa; CYP: Cytochrom P450; EKG: Elektrokardiogramm; FACIT-F: Functional Assessment of Chronic Illness Therapy – Fatigue;

Item[a]	Charakteristikum	Studieninformation
GCAP: Zytapheresebehandlung; HIV: Humanes Immundefizienz-Virus; hsCRP: Hoch sensitives C-reaktives Protein; IBDQ: Inflammatory Bowel Disease Questionnaire; IRT: Interaktives Response Technologie System; ITT: Intention-to-Treat; LCAP: Leukocytapheresis; MDRD: Modification of Diet in Renal Disease; MI: Multiple Imputation; MMRM: Mixed Model for Repeated Measures; N: Anzahl Patienten; NRI: Non-Responder-Imputation; NRI-C: NRI incorporating MI to handle missing data due to COVID-19; REML: Restriktive Maximum-Likelihood-Methode; SA: Safety Analyse Population; TNF: Tumornekrosefaktor; TPN: Totale parenterale Ernährung;UESI: Unerwünschtes Ereignis von speziellem Interesse; ULN: Obere Normgrenze		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

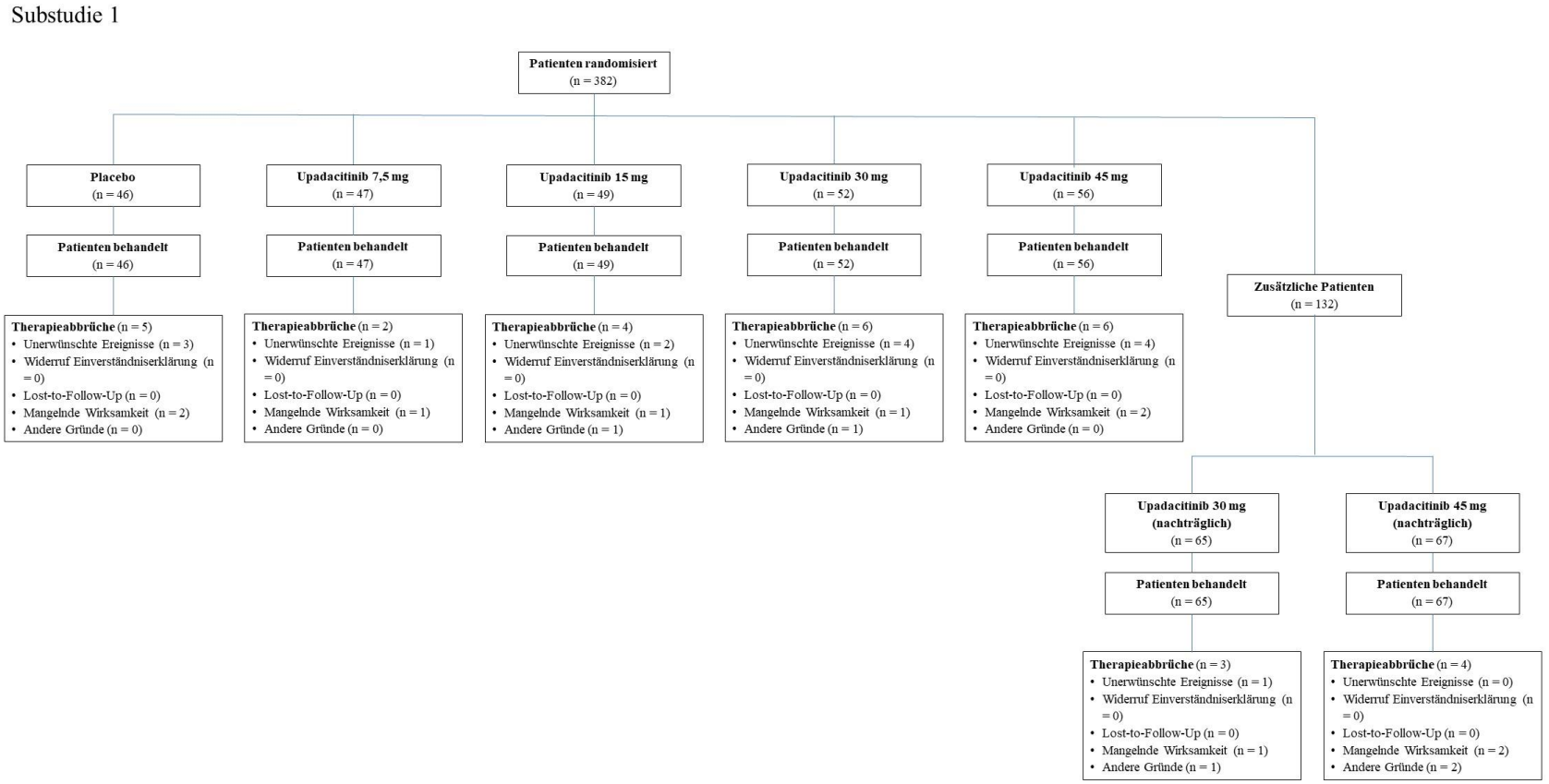

Abbildung 31: Flussdiagramm für Studie U-ACHIEVE Substudie 1

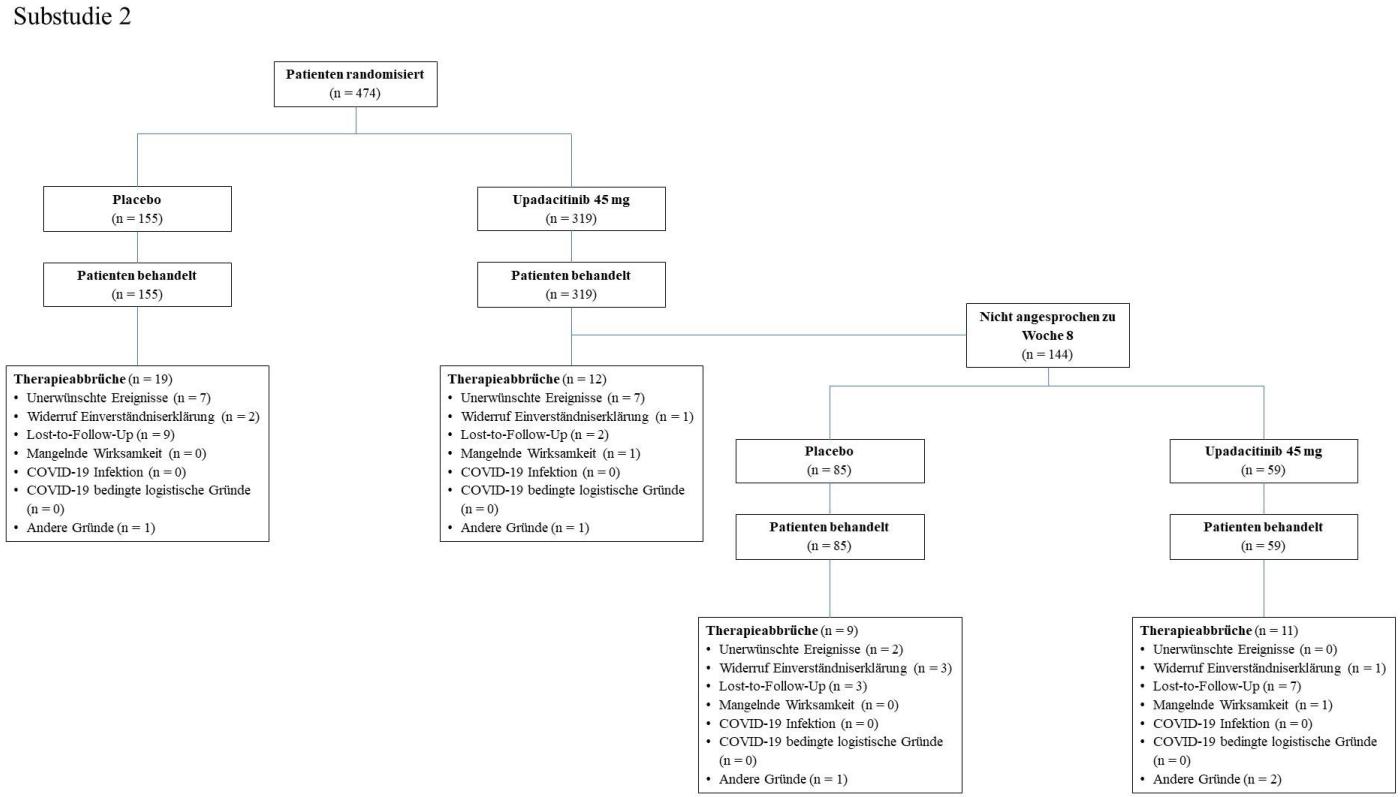

Abbildung 32: Flussdiagramm für Studie U-ACHIEVE Substudie 2

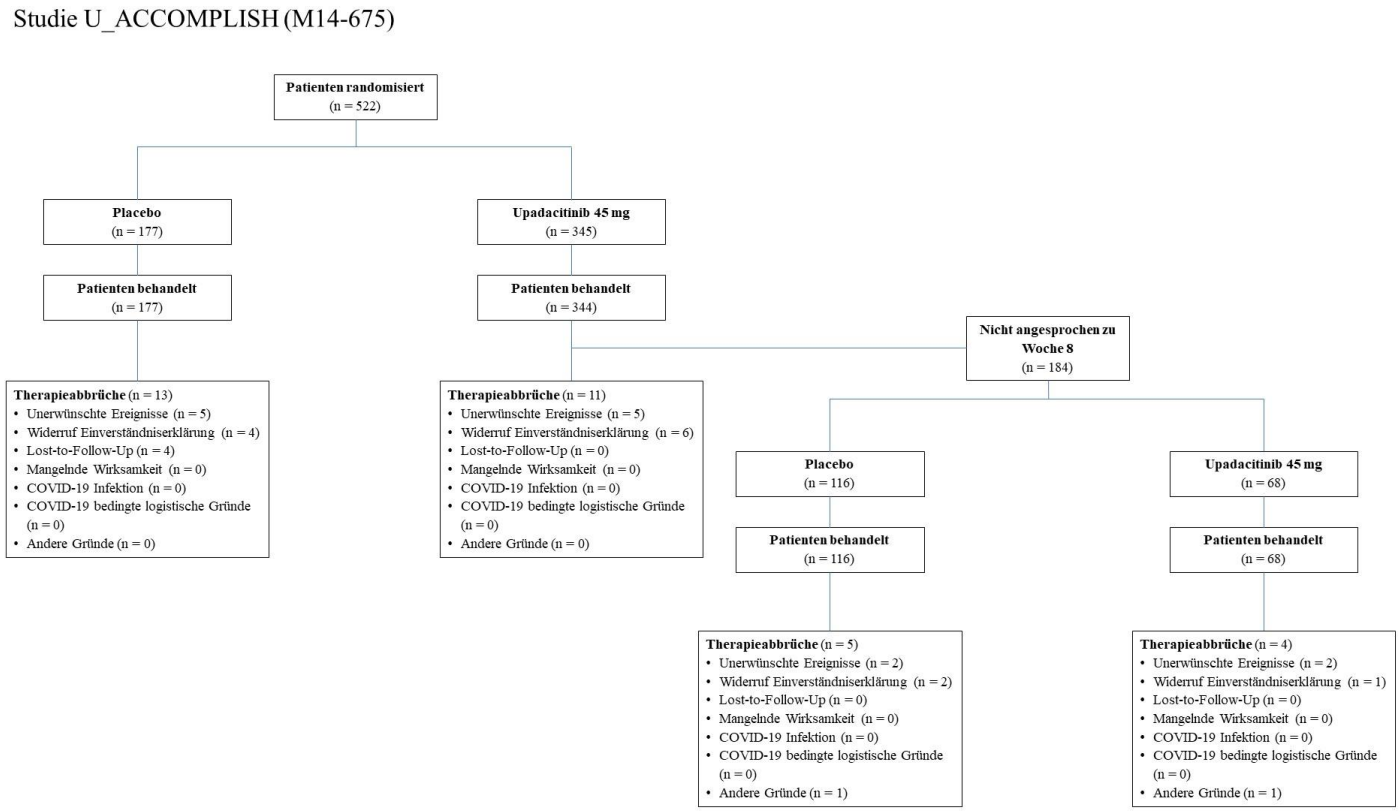

Abbildung 33: Flussdiagramm für Studie U-ACCOMPLISH

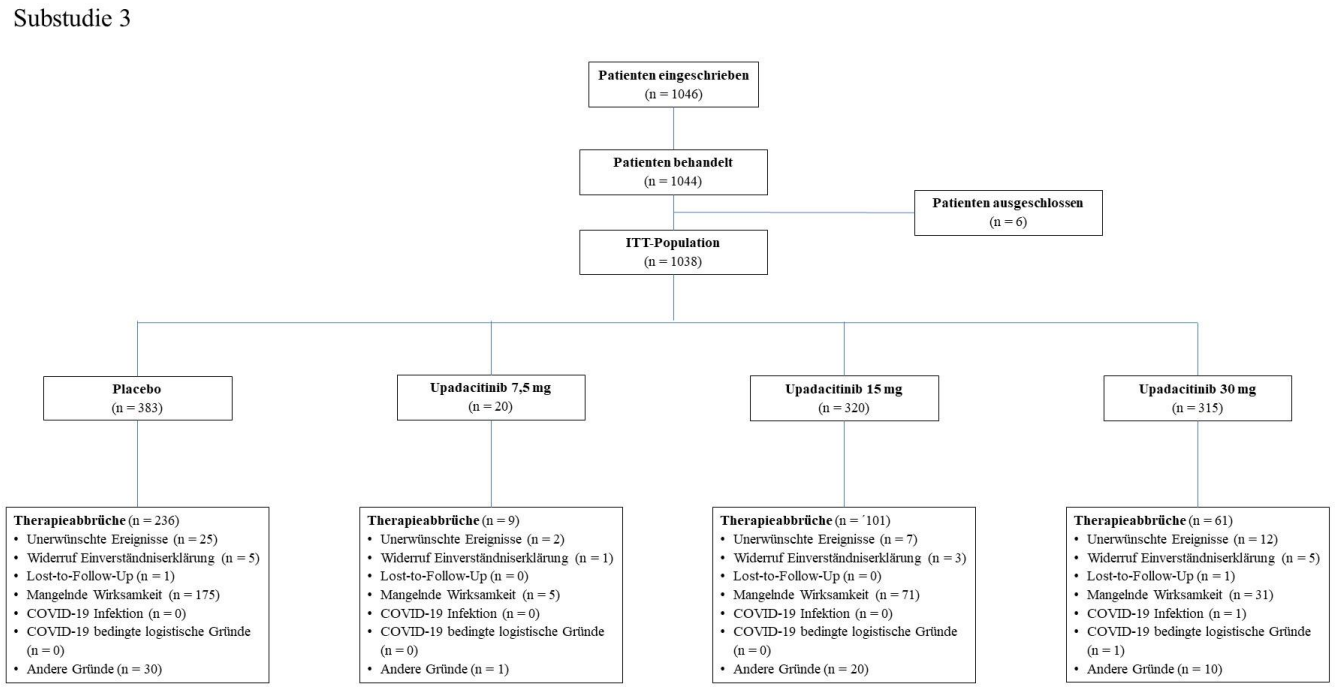

Abbildung 34: Flussdiagramm für Studie U-ACHIEVE Substudie 3

Tabelle 4-142 (Anhang): Studiendesign und -methodik für Studie UNIFI Erhaltungsstudie

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Bewertung der Wirksamkeit und Sicherheit der Behandlung mit Ustekinumab 130 mg oder einer auf dem Körpergewicht basierenden Ustekinumab-Dosis, die ca. 6 mg pro Kilogramm Körpergewicht beträgt, in der Induktionsphase; oder Ustekinumab 90 mg (alle 12 Wochen) oder Ustekinumab (alle 8 Wochen) in der Maintenance-Therapie im Vergleich mit Placebo hinsichtlich des Erreichens einer klinischen Remission, gemäß dem Mayo Score, bei Patienten mit mäßiger bis schwerer Colitis Ulcerosa.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Eine multizentrische, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie, die eine 8-wöchige Induktionsstudie sowie eine 44-wöchige Maintenance-Studie (insgesamt 52 Behandlungswochen) umfasst. Patienten wurden im Rahmen der Induktionsphase in eine von drei Behandlungsgruppen mit Zuteilungsverhältnis 1:1:1 eingeteilt: • Ustekinumab 130 mg • Ustekinumab (6 mg pro Kilgramm Körpergewicht) • Placebo Patienten, die zu Woche 8 der Induktionsphase auf Ustekinumab klinisch ansprachen, waren berechtigt, an der Maintenance-Studie teilzunehmen, ebenso wie diejenigen, die zunächst nicht auf Placebo ansprachen und die dann in Woche 8 eine Induktionsdosis von Ustekinumab (6 mg pro Kilogramm Körpergewicht) und zu Woche 16 ein Ansprechen erzielten. Für die Maintenance-Studie wurden berechtigte Patienten einer der drei Therapiegruppen mit einem Zuteilungsverhältnis von 1:1:1 randomisiert zugeordnet: • Ustekinumab 90 mg (alle 12 Wochen) • Ustekinumab 90 mg (alle 8 Wochen) • Placebo
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<u>Wesentliche Änderungen aus Amendment 1 (14. Juli 2015):</u> 1. <u>Einschlusskriterium 15b:</u> Die Common Terminology Criteria for Adverse Events (CTCAE) definieren ein unerwünschtes Ereignis vom Grad 1 für „verringerte weiße Blutkörperchen“, d. h. unterhalb der unteren Grenze des Normwerts, als < 3 × 109/l oder 3,0 × 103/µl; daher wurde die zulässige untere Grenze für die Anzahl der weißen Blutkörperchen eines/einer Patienen/Patientin auf mindestens 3 × 103/µl angehoben. 2. <u>Einschlusskriterium 15e:</u> Es wurde verdeutlicht, dass für die Teilnahme an der Studie berechtigte Patienten einen Serumkreatininwert von höchstens 1,5 mg/dl haben sollten. <u>Wesentliche Änderungen aus Amendment 1 (20. April 2016):</u>

Item[a]	Charakteristikum	Studieninformation
		1. Einschlusskriterium 6e: Beclometasondipropionat wurde bei dem Einschlusskriterium 6e ergänzt, da es in einigen Regionen, in denen diese Studie durchgeführt wird, für die Behandlung von Colitis Ulcerosa zugelassen ist. 2. Einschlusskriterien 6f und 7e: Der spezifische Hinweis auf Budesonid als Kortikosteroid wurde aus den Einschlusskriterien 6f und 7e gestrichen, da die Verwendung des Begriffs „Kortikosteroid‘‘ sowohl prednisonähnliche Kortikosteroide als auch Kortikosteroide, die für die gezielte Verabreichung an den Dickdarm entwickelt wurden, wie z. B. Budesonid und Beclometasondipropionat, umfasst und die in einigen Regionen, in denen die Studie durchgeführt wird, für die Behandlung von Colitis Ulcerosa zugelassen sind. 3. Einschlusskriterium 15: Um mehr Klarheit zu gewährleisten, wurden im Einschlusskriterium 15 Äquivalente des Internationalen Einheitensystems (SI) für die angegebenen Ergebnisse der Screening-Labortests hinzugefügt. 4. Ausschlusskriterium 12e: Eine weitere Apherese-Methode, die einen ähnlichen Wirkmechanismus hat, wurde im Ausschlusskriterium Kriterium 12e ergänzt. 5. Ausschlusskriterium 17: Das Ausschlusskriterium 17 wurde aktualisiert, um der Verfügbarkeit einer hochwirksamen antiviralen Therapie mit HCV-Clearance (gemessen an der HCV-RNA) zu reflektieren, die dazu führt, dass Patienten zwar seropositiv bleiben, aber als klinisch von ihrer Hepatitis C geheilt gelten. 6. Ausschlusskriterium 29: Das Ausschlusskriterium 29 wurde aktualisiert und bezieht sich nun auf die aktuellste Ausgabe des Diagnostic and Statistical Manual of Mental Disorders.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: 9. Männliche oder weibliche Patienten im Alter von ≥ 18 Jahren. 10. Diagnose von Colitis Ulcerosa mindestens drei Monate vor dem Screening. 11. Mittelschwere bis schwere Colitis Ulcerosa, definiert als ein Score von 6 bis 12 auf der Mayo-Skala (zur Baseline), einschließlich, unter Verwendung des Mayo-Subscore der Endoskopie, der bei der zentralen Auswertung der Videoendoskopie ermittelt wird. 12. Patienten haben eine Screening-Endoskopie mit ≥ 2 Subscores des Mayo Score, ermittelt durch eine zentrale Auswertung der Videoendoskopie. 13. Patienten hatten ein Behandlungsversagen nach Erhalt einer biologischen Therapie, d. h. Patienten haben nach Erhalt einer Behandlung mit einem oder mehreren TNF-Inhibitor

Item[a]	Charakteristikum	Studieninformation
		oder Vedolizumab, in einer für die Behandlung von Colitis Ulcerosa zugelassenen Dosis, einen dokumentierten Nachweis, dass sie auf eine solche Behandlung nicht angesprochen oder diese nicht vertragen haben; ODER sie sind naiv bezüglich einer biologischen Therapie, mit TNF-Inhibitor oder Vedolizumab, oder hatten kein Behandlungsversagen bezüglich eines Nicht-Ansprechens oder Unverträglichkeit nach Erhalt einer biologischen Therapie und verfügen über eine frühere oder aktuelle Medikamentenanamnese für Colitis Ulcerosa, die mindestens eines der folgenden Merkmale aufweist: a. Unzureichendes Ansprechen oder Unverträglichkeit der aktuellen Behandlung mit oralen Kortikosteroiden oder Immunmodulatoren (6-MP oder AZA); ODER b. Anamnese des Nicht-Ansprechens auf oder der Unverträglichkeit von mindestens einer der folgenden Therapien: orale oder intravenös-verabreichte Kortikosteroide oder Immunmodulatoren (6-MP oder AZA); ODER c. Eine frühere Kortikosteroidabhängigkeit, d. h. die Unfähigkeit, Kortikosteroide erfolgreich abzusetzen, ohne dass die Symptome von Colitis Ulcerosa zurückkehren. 14. Vor der ersten Verabreichung der Studienmedikation müssen die folgenden Bedingungen erfüllt sein: a. Im Fall der Einnahme von konventionellen Immunmodulatoren (d. h. AZA, 6-MP oder MTX), müssten diese bereits für einen Zeitraum von ≥ 12 Wochen eingenommen worden sein sowie für mindestens vier Wochen in einer stabilen Dosis. b. Falls AZA, 6-MP oder MTX vor kurzem abgesetzt wurden, dann müssen diese seit mindestens vier Wochen abgesetzt worden sein. c. Im Fall der Einnahme von oralen 5-ASA-Präparaten, muss die Dosis seit mindestens zwei Wochen stabil gewesen sein. d. Im Fall der Einnahme von oralen Kortikosteroiden, muss die Dosis ≤ 20 mg/Tag Prednison oder ein Äquivalent davon sein und seit mindestens zwei Wochen in stabiler Dosis eingenommen worden sein. e. Im Fall der Einnahme von Budesonid, muss die Dosis seit mindestens zwei Wochen in stabiler Dosis eingenommen worden sein. f. Falls orale 5-ASA-Verbindungen oder orale Kortikosteroide (einschließlich Budesonid) kürzlich abgesetzt wurden, dann müssen diese seit mindestens zwei Wochen abgesetzt worden sein.

Item[a]	Charakteristikum	Studieninformation
		15. Die folgenden Medikamente/Therapien müssen vor der ersten Verabreichung der Studienmedikation abgesetzt worden sein: a. Vedolizumab seit mindestens vier Monaten. b. TNF-Antagonisten-Therapie (z. B. Infliximab, Etanercept, Certolizumab, Adalimumab, Golimumab) seit mindestens acht Wochen. c. Cyclosporin, Tacrolimus oder Sirolimus seit mindestens vier Wochen. d. 6-Thioguanin (6-TG) seit mindestens vier Wochen. e. Rektale Kortikosteroide (d. h. Kortikosteroide, einschließlich Budesonid, die in den Rektum oder das Colon Sigmoideum über Schaum, Einlauf oder Zäpfchen verabreicht werden) seit mindestens zwei Wochen. f. Rektale 5-ASA-Präparate (d. h. 5-ASA, die in den Rektum oder das Colon Sigmoideum über Schaum, Einlauf oder Zäpfchen verabreicht werden) seit mindestens zwei Wochen. g. Parenterale Kortikosteroide seit mindestens zwei Wochen. h. Totale parenterale Ernährung (TPN) seit mindestens zwei Wochen. i. Antibiotika für die Behandlung von Colitis Ulcerosa (z. B. Ciprofloxacin, Metronidazol oder Rifaximin) seit mindestens zwei Wochen. 16. Patienten, die 45 Jahre alt oder älter sind, müssen sich entweder eine Koloskopie zur Feststellung adenomatöser Polypen innerhalb von fünf Jahren vor der ersten Verabreichung der Studienmedikation unterzogen haben, oder sich eine Koloskopie zur Feststellung adenomatöser Polypen bei dem Screening unterziehen. Die adenomatösen Polypen müssen vor der ersten Verabreichung der Studienmedikation entfernt worden sein. 17. Patienten, die seit acht Jahren oder länger an einer ausgedehnten Kolitis oder seit zehn Jahren oder länger an einer auf die linke Seite des Dickdarms begrenzten Erkrankung leiden, müssen sich entweder innerhalb von einem Jahr vor der ersten Verabreichung der Studienmedikation eine Koloskopie zur Feststellung von Dysplasien unterzogen haben oder sich eine Koloskopie zur Feststellung von Malignität bei dem Screening unterziehen. 18. Patienten werden nach den folgenden Tuberkulose-Screening-Kriterien als geeignet angesehen: a. Hatten keine frühere, latente oder aktive Tuberkulose vor dem Screening. Eine Ausnahme gilt für Patienten, die eine latente Tuberkulose in der Anamnese haben und derzeit eine Behandlung gegen latente Tuberkulose erhalten, die vor der ersten

Item[a]	Charakteristikum	Studieninformation
		Verabreichung der Studienmedikation eine Behandlung gegen latente Tuberkulose beginnen werden oder die nachweislich innerhalb von drei Jahren vor der ersten Verabreichung der Studienmedikation eine angemessene Behandlung gegen latente Tuberkulose abgeschlossen haben. Es liegt in der Verantwortung des/der Prüfarztes/Prüfärztin, die Angemessenheit der vorangegangenen antituberkulösen Behandlung zu überprüfen und entsprechende Unterlagen vorzulegen. b. Hatten keine Anzeichen oder Symptome, die auf aktive Tuberkulose hindeuten, weder in der Anamnese noch bei der körperlichen Untersuchung. c. Hatten kürzlich keinen engen Kontakt mit einer Person mit aktiver Tuberkulose oder werden, falls es einen solchen Kontakt gab, an einen auf Tuberkulose spezialisierten Arzt oder eine spezialisierte Ärztin überwiesen, um sich einer zusätzlichen Untersuchung zu unterziehen und gegebenenfalls vor der ersten Verabreichung der Studienmedikation eine angemessene Behandlung gegen latente Tuberkulose zu erhalten. d. Hatten innerhalb von zwei Monaten vor der ersten Verabreichung der Studienmedikation ein negatives QuantiFERON-TB-Gold-Testergebnis oder ein neu identifiziertes positives QuantiFERON-TB-Gold-Testergebnis, bei dem eine aktive Tuberculose ausgeschlossen wurde und für das vor der ersten Verabreichung der Studienmedikation eine angemessene Behandlung gegen latente Tuberkulose eingeleitet wurde; Hatten innerhalb von zwei Monaten vor der ersten Verabreichung der Studienmedikation ein negatives Tuberkulin-Hauttest-Ergebnis oder ein neu identifiziertes positives Tuberkulin-Hauttest-Ergebnis, bei dem eine aktive Tuberkulose ausgeschlossen wurde und eine angemessene Behandlung gegen latenten Tuberkulose vor der ersten Verabreichung der Studienmedikation eingeleitet wurde. Dies ist zusätzlich erforderlich, falls der QuantiFERON-TB-Gold-Test in dem entsprechenden Land nicht zugelassen oder registriert ist, oder der Tuberkulin-Hauttest von den örtlichen Gesundheitsbehörden vorgeschrieben ist. e. Verfügen über eine Röntgenaufnahme des Brustkorbs (Ansicht von hinten nach vorne), die innerhalb von drei Monaten vor der ersten Verabreichung der Studienmedikation durchgeführt und von einem/einer qualifizierten Radiologen/Radiologin gelesen wurde,

Item[a]	Charakteristikum	Studieninformation
		ohne Anzeichen von aktueller, aktiver Tuberkulose oder alter, inaktiver Tuberkulose. 19. Vor der Randomisierung ist eine für die Teilnahme an der Studie berechtigte Patientin entweder: a. Nicht im gebärfähigen Alter, d. h. entweder prämenarchal, postmenopausal, dauerhaft sterilisiert oder anderweitig schwangerschaftsunfähig; ODER b. Im gebärfähigen Alter, wendet jedoch, falls heterosexuell aktiv, eine hochwirksame Verhütungsmethode an, die den örtlichen Vorschriften über die Anwendung von Verhütungsmethoden für Patienten, die an klinischen Studien teilnehmen, entspricht, wie z. B. eine nachweisliche Anwendung von oralen, injizierten oder implantierten hormonellen Verhütungsmethoden, Einsetzen eines Intrauterinpessars oder Intrauterinsystems, Barrieremethoden, Sterilisation des männlichen Partners (der vasektomierte Partner sollte der einzige Partner der Patientin sein) oder echte Abstinenz (falls dies mit der bevorzugten und üblichen Lebensweise der Patientin übereinstimmt). 20. Eine Patientin im gebärfähigen Alter muss beim Screening ein negatives Serum-Schwangerschaftstestergebnis sowie ein negatives Urinschwangerschaftstestergebnis zu Woche 0 vorlegen. 21. Eine Patientin muss zustimmen, keine Eizellen oder Oozyten für die Zwecke der assistierten Reproduktion während der Studie und für insgesamt 20 Wochen nach der letzten Verabreichung der Studienmedikation zu spenden. 22. Ein Patient, der mit einer Frau im gebärfähigen Alter sexuell aktiv ist und sich einer Vasektomie nicht unterzogen hat, muss zustimmen, eine Barrieremethode der Geburtenkontrolle zu verwenden. Alle Patienten müssen sich außerdem verpflichten, kein Sperma während der Studie und bis insgesamt 20 Wochen nach der letzten Verabreichung der Studienmedikation zu spenden. 23. Die Ergebnisse der Screening-Labortests liegen innerhalb der folgenden Parameter: a. Hämoglobin ≥ 8,0 g/dl b. Anzahl der weißen Blutkörperchen ≥ 2,5 × 103/µl c. Neutrophile ≥ 1,5 × 103/µl d. Thrombozyten ≥ 100 × 103/µl e. Serum-Kreatinin < 1,5 mg/dl f. Konzentrationen von Alanin-Aminotransferase und Aspartat-Aminotransferase müssen innerhalb des 2-fachen, oberen Normwertes des Labors, das den Test durchführt

Item[a]	Charakteristikum	Studieninformation
		24. Patienten müssen bereit und in der Lage sein, die in dem Protokoll aufgeführten Prohibitionen und Restriktionen zu befolgen. 25. Jeder/jede Patient/Patientin muss eine Einverständniserklärung unterschreiben, in der er/sie bestätigt, dass er/sie Zweck und Verfahren der Studie verstanden hat und bereit ist, an der Studie teilzunehmen. In Regionen, in denen das gesetzliche Mindestalter für die Einwilligung älter als 18 Jahre ist, muss die Einwilligung nach Aufklärung sowohl von dem/der Patienten/Patientin als auch von seinem/ihrem gesetzlichen Vertreter erlangt und unterzeichnet werden. 26. Jeder/jede Patient/Patientin muss eine seperate Einwilligungserklärung unterzeichnen, wenn er/sie sich bereit erklärt, optionale DNA-Proben für die Forschung zur Verfügung zu stellen, sofern die örtlichen Vorschriften dies zulassen. In Regionen, in denen das gesetzliche Alter für die Einwilligung über 18 Jahren liegt, muss die Einwilligungserklärung sowohl von dem/der Patient/Patientin als auch von seinem/ihrem gesetzlichen Vertreter unterzeichnet werden. Die Verweigerung der Zustimmung für die optionalen DNA-Proben, schließt einen/eine Patienten/Patientin nicht von der Teilnahme an der Studie aus. <u>Ausschlusskriterien:</u> 1. Patienten haben eine schwere, ausgedehnte Kolitis, nachgewiesen durch: a. Aktuelle Hospitalisierung zur Behandlung von Colitis Ulcerosa; ODER b. Einschätzung eines/einer Prüfarztes/Prüfärztin, dass eine Kolektomie innerhalb von 12 Wochen nach der Baseline benötigt wird; ODER c. Symptomenkomplex bei dem Screening oder der Baseline-Untersuchung, der mindestens vier der folgenden Punkte umfasst: i. Diarrhoe mit ≥ 6 Stuhlgängen/Tag mit makroskopischem Blut im Stuhl ii. Schwere oder rückfällige fokale abdominelle Empfindlichkeit iii. Anhaltendes Fieber (≥ 37,5°C) iv. Tachykardie (> 90 Schläge/Minute) v. Anämie (Hämoglobin < 8,5 g/dl) 2. Colitis Ulcerosa ist nur auf das Rektum oder auf < 20 cm des Dickdarms beschränkt. 3. Vorliegen eines Stomas. 4. Aktueller oder früherer Einsatz von Fisteln.

Item[a]	Charakteristikum	Studieninformation
		5. Eine erforderliche oder eine innerhalb von zwei Monaten vor dem Screening geplante Operation wegen aktiver gastrointestinaler Magen-Darm-Blutung, Peritonitis, Darmverschluss oder intraabdominalem oder pankreatischem Abszess, der eine chirurgische Drainage erfordert; ODER andere Bedingungen, die möglicherweise die Bewertung des Nutzens der Behandlung mit der Studienmedikation beeinträchtigen könnten. 6. Vorliegen einer symptomatischen Kolon- oder Dünndarmobstruktion, bestätigt durch objektiven radiologischen oder endoskopischen Nachweis einer Striktur mit daraus resultierender Obstruktion. 7. Frühere ausgedehnte Kolonresektion, die eine adäquate Bewertung der Wirkung der Studienmedikation auf die klinische Krankheitsaktivität verhindern würde. 8. Frühere Dysplasie der Dickdarmschleimhaut. Patienten werden nicht aufgrund eines pathologischen Befundes einer „unbestimmten Dysplasie mit reaktiver Atypie‘‘ von der Studie ausgeschlossen. 9. Vorliegen adenomatöser Kolonpolypen bei der Screening-Endoskopie, falls diese nicht vor Studienbeginn entfernt wurden oder frühere adenomatöse Kolonpolypen, die nicht entfernt wurden. 10. Diagnose von unbestimmter Kolitis, mikroskopischer Kolitis, ischämischer Kolitis oder Morbus Crohn oder klinische Befunde, die auf eine Erkrankung durch Morbus Crohn hindeuten. 11. Positiver Nachweis einer Stuhlkultur oder einer anderen Untersuchung auf einen enterischen Erreger, einschließlich Clostridium-Difficile-Toxin, in den letzten vier Monaten, es sei denn, eine erneute Untersuchung ist negativ ausgefallen und es gibt keinerlei Anzeichen für eine aktuelle Infektion mit diesem Erreger. 12. Patienten haben die folgenden begleitenden oder früheren medizinischen Therapien erhalten: a. Eine biologische Therapie, die auf IL-12 und/oder IL-23 abzielt, wie z. B. Ustekinumab, Briakinumab, Guselkumab. b. Natalizumab innerhalb von zwölf Monaten vor der ersten Verabreichung der Studienmedikation. c. Mittel, die B- oder T-Zellen abbauen, wie z. B. Rituximab, Alemtuzumab, innerhalb von zwölf Monaten nach der ersten Verabreichung der Studienmedikation oder eine fortgestezte Depletion von B- oder T-Zellen länger als zwölf Monate nach Beendigung der Therapie mit Lymphozyten-depletierenden Mitteln.

Item[a]	Charakteristikum	Studieninformation
		d. Jedes Prüfpräparat innerhalb von vier Wochen vor der ersten Verabreichung der Studienmedikation oder innerhalb von fünf Halbwertszeiten des Prüfpräparats, je nachdem, welches länger ist. e. Apherese, d. h. Adacolumn-Apherese, innerhalb von zwei Wochen vor der ersten Verabreichung der Studienmedikation. 13. Vor dem Screening bestand eine latente oder aktive granulomatöse Infektion, einschließlich Histoplasmose oder Kokzidioidomykose. 14. Eine frühere oder aktuelle, chronische oder rezidivierende Infektionskrankheit, einschließlich, aber nicht nur chronische Niereninfektion, chronische Brustinfektion, Sinusitis, rezidivierende Harnwegsinfektion, eine offene, drainierende oder infizierte Hautwunde oder ein Geschwür. 15. Eine Röntgenaufnahme des Brustkorbs innerhalb von drei Monaten vor der ersten Verabreichung der Studienmedikation, die eine Anomalie zeigt, welche auf ein malignes Geschehen oder eine aktive Infektion, einschließlich Tuberkulose, hinweist. 16. Patienten, die positiv für Antikörper des Humanen Immundefizienz-Virus (HIV) waren oder die beim Screening positiv auf HIV getestet wurden. 17. Patienten sind sero-positiv für Antikörper gegen das Hepatitis-C-Virus (HCV). 18. Patienten müssen sich einem Screening auf das Hepatitis-B-Virus (HBV) unterziehen. Dies umfasst mindestens Tests auf HBV-Oberflächenantigen (HBsAg), HBV-Oberflächenantikörper (Anti-HBs) und HBV-Core Antikörper (Anti-HBc): a. Patienten, die bei allen HBV-Screening-Tests, d. h. HBsAg-, Anti-HBc- und Anti HBs-, negativ sind, sind für die Studie zugelassen. b. Patienten, die positiv auf das Oberflächenantigen (HBsAg+) getestet wurden, dürfen nicht an der Studie teilnehmen, unabhängig von den Ergebnissen anderer Hepatitis-B-Tests. c. Patienten mit negativem Test auf Oberflächenantigen (HBsAg-) und positivem Test auf Core-Antikörper (anti-HBc+) sowie Oberflächenantikörper (anti-HBs+), sind für die Studie zugelassen. d. Patienten, die nur auf Oberflächenantikörper (anti-HBs+) positiv getestet wurden, sind für die Studie zugelassen. e. Patienten, die nur auf Core-Antikörper (anti-HBc+) positiv getestet wurden, müssen sich einem weiteren Test auf Hepatitis-B-Desoxyribonukleinsäure

Item[a]	Charakteristikum	Studieninformation
		(HBV-DNA-Test) unterziehen. Fällt der HBV-DNA-Test positiv aus, dann ist der/die Patient/Patientin nicht für die Studie zugelassen. Fällt der HBV-DNA-Test hingegen negativ aus, so ist der/die Patient/Patientin für die Studie zugelassen. Falls der HBV-DNA-Test nicht durchgeführt werden kann, ist der/die Patient/Patientin nicht für die Studie zugelassen. 19. Patienten erhielten eine Bacille-Calmette-Guerin (BCG)-Impfung innerhalb von zwölf Monaten vor der Baseline oder eine andere bakterielle oder virale Lebendimpfung innerhalb von zwölf Wochen vor der Baseline. 20. Der/die Patient/Patientin hat oder hatte jemals eine nicht-tuberkulöse mykobakterielle Infektion oder eine schwerwiegende opportunistische Infektion, wie z. B. Cytomegalovirus-Kolitis, Pneumocystis Carinii, Aspergillose. 21. Der/die Patient/Patientin hatte eine schwerwiegende Infektion, wie z. B. Hepatitis, Lungenentzündung oder Pyelonephritis, war wegen einer Infektion im Krankenhaus oder erhielt eine parenterale Antibiotikabehandlung wegen einer Infektion innerhalb von zwei Monaten vor der ersten Verabreichung der Studienmedikation. Weniger schwerwiegende Infektionen, wie z. B. eine akute Infektion der oberen Atemwege und eine einfache Harnwegsinfektion, müssen nach Ermessen des/der Prüfarztes/Prüfärztin nicht als Ausschlusskriterium gelten. 22. Nachweis einer Herpes-Zoster-Infektion innerhalb von den letzten acht Wochen vor der Baseline. 23. Frühere oder aktuell bekannte bösartige Erkrankungen, mit Ausnahme von Basalzellkarzinomen, Plattenepithelkarzinomen-in-situ der Haut oder Gebärmutterhalskarzinomen-in-situ, die ohne Anzeichen eines Rezidivs behandelt wurden, oder Plattenepithelkarzinomen der Haut, die innerhalb von fünf Jahren vor dem Screening ohne Anzeichen eines Rezidivs behandelt wurden. 24. Eine frühere oder aktuelle lymphoproliferativen Erkrankung, einschließlich eines Lymphoms, oder Anzeichen und Symptome, die auf eine mögliche lymphoproliferative Erkrankung hindeuten, wie Lymphadenopathie von ungewöhnlicher Größe oder Lage, z. B. Knoten im hinteren Dreieck des Halses, infraklavikulär, epitrochleäre oder periaortale Bereiche, oder klinisch signifikante Hepatomegalie oder Splenomegalie, oder monoklonale Gammopathie von undefinierter Bedeutung.

Item[a]	Charakteristikum	Studieninformation
		25. Bekannte Allergien, Überempfindlichkeiten oder Unverträglichkeiten gegen Ustekinumab oder einen seiner Hilfsstoffe. 26. Schwere, fortschreitende oder unkontrollierte Nieren-, Leber-, hämatologische, endokrine, pulmonale, kardiale, neurologische, psychiatrische oder zerebrale Erkrankungen oder Anzeichen oder Symptome davon. 27. Der/die Patient/Patientin hat ein transplantiertes Organ, mit Ausnahme einer Hornhauttransplantation, die vor mehr als zwölf Wochen vor dem Screening durchgeführt wurde. 28. Der/die Patient/Patientin hat sich zuvor einer Allergie-Immuntherapie unterzogen, um anaphylaktische Reaktionen zu verhindern. 29. Ein früherer Drogen- oder Alkoholmissbrauch gemäß dem Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV), innerhalb eines Jahres vor dem Screening. 30. Der/die Patient/Patientin leidet unter schlechter Verträglichkeit der Venenpunktion oder hat keinen adäquaten venösen Zugang für die erforderliche Blutentnahme während des Studienzeitraums. 31. Patientinnen, die schwanger sind, stillen oder planen, schwanger zu werden, oder Patienten, die planen, während der Teilnahme an der Studie oder innerhalb von 20 Wochen nach Verabreichung der letzten Dosis der Studienmedikation ein Kind zu zeugen. 32. Vorliegen eines Zustands, bei dem nach Ansicht des/der Prüfarztes/Prüfärztin die Teilnahme nicht im besten Interesse des/der Patienten/Patientin liegt, wie z. B. Beeinträchtigung des Wohlbefindens, oder der die im Prüfplan vorgesehenen Bewertungen verhindern, einschränken oder verzerren könnte. 33. Patienten, die derzeit an einer anderen Studie teilnehmen oder beabsichtigen, an einer anderen Studie teilzunehmen, bei der ein Prüfpräparat oder ein Verfahren verwendet wird, während sie an der Studie teilnehmen. 34. Mitarbeiter/Mitarbeiterinnen des/der Prüfarztes/Prüfärztin oder des Prüfzentrums, die direkt an der vorgeschlagenen Studie oder anderen Studien unter der Leitung dieses/dieser Prüfarztes/Prüfärztin oder des Prüfzentrums beteiligt sind, sowie Familienangehörige der Mitarbeiter/Mitarbeiterinnen oder des/der Prüfarztes/Prüfärztin.
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie

Item[a]	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	In der Induktionsstudie erhielten die Patienten eine einzelne intravenöse Infusion von Ustekinumab (130 mg oder 6 mg pro Kilogramm Körpergewicht) oder Placebo. In der Maintenance-Studie erhielten die Patienten subkutane Injektionen von Ustekinumab 90 mg (alle 12 Wochen) oder Ustekinumab 90 mg (alle 8 Wochen) oder Placebo bis Woche 40.
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	Induktionsstudie: • Primäre Endpunkte: a. Klinische Remission, definiert als ein Mayo Gesamtscore von ≤ 2 und kein Subscore > 1 oder eine absolute Stuhlzahl von 3 oder weniger (durchschnittliche Stuhlzahl pro Tag in den drei Tagen vor einer Visite), einen Mayo-Rektalblutung-Subscore von 0 und einen endoskopischen Mayo Subscore von 0 oder 1) zu Woche 8 • Wichtige sekundäre Endpunkte: a. Endoskopische Verbesserung (definiert als ein endoskopischer Mayo-Subscore von 0 oder 1) zu Woche 8 b. Klinisches Ansprechen zu Woche 8 c. Veränderung des IBDQ-Score im Vergleich zur Baseline (mit Ausnahme der Vereinigten Staaten) • Sekundäre Endpunkte: a. Histo-endoskopische Schleimhautheilung zu Woche 8, die sowohl eine histologische Verbesserung, definiert als neutrophile Infiltrationin < 5 % der Krypten, keine Zerstörung der Krypta und keine Erosionen, Ulzerationen oder Granulationsgewebe, als auch eine endoskopische Verbesserung erforderte Maintenance-Studie: • Primäre Endpunkte: a. Klinische Remission (beide Definitionen) zu Woche 44 • Wichtige sekundäre Endpunkte: a. Aufrechterhaltung des klinischen Ansprechens bis Woche 44 b. Endoskopische Verbesserung zu Woche 44 c. Kortikosteroid-freie klinische Remission zu Woche 44 d. Aufrechterhaltung der klinischen Remission bis Woche 44 bei Patienten, die zur Baseline der Maintenance-Studie eine klinische Remission hatten • Weitere sekundäre Endpunkte: a. Veränderungen des Mayo Score b. Mayo-Subscore c. Veränderungen des partiellen Mayo Score d. Stuhlfrequenz-Subscore

Item[a]	Charakteristikum	Studieninformation
		e. Rektalblutung-Subscore f. Endoskopie-Subscore g. Symptomatische Remission h. Klinische Remission i. Aufrechterhaltung des klinischen Ansprechens j. Endoskopische Verbesserung k. Vorliegen einer normalen oder inaktiven Schleimhauterkrankung l. Veränderungen bzgl. der Einnahme von begleitenden Kortikosteroiden m. Verbesserung des IBDQ-Score n. Veränderungen des SF-36-Score bzgl. physischer sowie mentaler Komponenten o. Veränderungen der EQ-5D-Dimension, des EQ-5D-Index sowie des VAS-Gesundheitszustand-Score p. Schleimhautheilung q. Veränderungen der CRP-Konzentration r. Veränderungen der fäkalen Laktoferrin-Konzentration s. Veränderungen der fäkalen Calprotectin-Konzentration t. Dauer bis zum Verlust des klinischen Ansprechens u. Dauer bis zum Verlust der klinischen Remission Die Follow-Up-Assessment für die Sicherheitsendpunkte a. Begleitmedikation, b. unerwünschte Ereignisse, c. schwerwiegende unerwünschte Ereignisse, Krankenhausaufenthalte im Zusammenhang mit Colitis Ulcerosa und d. chirurgische Eingriffe erfolgte während der Induktionsstudie bis Woche 8 oder Woche 16, als die Patienten die Maintenance-Studie begonnen haben, oder 20 Wochen nach Verabreichung der letzten Induktionsdosis für Patienten, die die Studie abbrachen, sowie während der gesamten Maintenance-Studie bis Woche 44.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<u>Wesentliche Änderungen aus Amendment 1 (14. Juli 2015):</u> 1. Eine zusätzliche Erläuterung wurde ergänzt, um zu verdeutlichen, dass die globale Definition der klinischen Remission ebenfalls zur Stratifizierung der Patienten, die in die Maintenance-Studie randomisiert werden, verwendet wird.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Unter der Annahme einer Inzidenz der klinischen Remission von 7 % in der Placebogruppe und 19 % in jeder Ustekinumab-Gruppe, wurde unter Verwendung eines Step-Up-Hochberg-Testverfahrens auf einem zweiseitigen Signifikanzniveau von 0,05 errechnet, dass 317 Patienten pro Induktionsgruppe ein Power von mehr als 90 % für den primären Endpunkt zu Woche 8

Item[a]	Charakteristikum	Studieninformation
		erreichen würden. Diese Fallzahl würde auch genügend Patienten für die primäre Analysepopulation der Maintenance-Studie bereitstellen. Unter der Annahme einer Inzidenz der klinischen Remission von 20 % in der Placebogruppe und 40 % in der Gruppe, die Ustekinumab 90 mg alle 8 Wochen (subkutan) erhielt, wurde errechnet, dass 109 Patienten pro Behandlungsgruppe der Maintenance-Studie eine 90%ige Power für den primären Endpunkt zu Woche 44 erreichen würden. Die Berechnung basierte auf dem Testverfahren mit fester Reihenfolge, beginnend mit der hochdosierten Ustekinumab-Gruppe (alle 8 Wochen), auf einem zweiseitigen Signifikanzniveau von 0,05.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Für die Induktionsstudie wurde eine Zwischenanalyse basierend auf dem primären Endpunkt der klinischen Remission zu Woche 8 geplant, nachdem die ersten 30 % der Patienten entweder die Induktionsuntersuchung zu Woche 8 abgeschlossen oder die Teilnahme an der Studie vor Woche 8 abgebrochen haben. Für die Maintenance-Studie wurden keine Zwischenanalysen geplant.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Zuteilung der Patienten in die Behandlungsgruppen der Induktions- und der Maintenance-Studie wurde zentral randomisiert. Dafür wurde ein computergenerierter Randomisierungsplan unter Aufsicht des Sponsors erstellt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	In der Induktionsphase wurde die Randomisierung mittels permutierter Blöcke durchgeführt, mit Stratifizierung nach dem Patientenstatus bezüglich eines früheren Behandlungsversagens mit biologischen Medikamenten (ja oder nein) sowie der geografischen Region (Osteuropa, Asien oder Rest der Welt). In der Maintenance-Studie erfolgte die Randomisierung der Patienten, die intravenöses Ustekinumab erhielten und zu Woche 8 der Induktionsphase ein klinisches Ansprechen erzielten, sowie derjenigen, die zu Woche 8 nach Erhalt von Placebo (intravenös) kein klinisches Ansprechen erzielten, dies jedoch zu Woche 16 nach der intraveösen Verabreichung von Ustekinumab (6 mg pro Kilogramm) erreichen konnten, ebenfalls mittels permutierter Blöcke, mit Stratifizierung nach intravenöser Induktionsphase (Ustekinumab 130 mg, Ustekinumab 6 mg pro Kilogramm oder Placebo, gefolgt von Ustekinumab 6 mg pro Kilogramm), dem Patientenstatus bezüglich klinischer Remission zur Baseline der Maintenance-Studie (ja oder nein) sowie der oralen Kortikosteroideinnahme (ja oder nein). Diese Patienten bildeten die randomisierte Maintenance-Population bzw. die primäre Analysepopulation.

Item[a]	Charakteristikum	Studieninformation
		Patienten, die in der Induktionsphase auf intravenöses Ustekinumab nicht ansprachen, erhielten in Woche 8 Ustekinumab 90 mg (subkutan) in verblindeter Form und wurden in Woche 16 erneut untersucht; diejenigen, die ein Ansprechen zeigten, nahmen an der Maintenance-Studie teil und erhielten subkutan Ustekinumab 90 mg alle 8 Wochen. Patienten, die zu Woche 8 der Induktionsphase auf intravenöses Placebo ansprachen, erhielten erneut Placebo (subkutan). Diese Patienten sowie diejenigen Patienten mit verzögertem Ansprechen auf Ustekinumab umfasste die nicht randomisierte Maintenance-Population.
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	Die Behälter der Studienmedikation wurden mit einem Abreißetikett versehen, das die Studienbezeichnung, die Medikamentennummer sowie die Referenznummer enthielt. Die Medikamentennummer wurde in den Fallbericht eingetragen, wenn die Studienmedikation verabreicht wurde. Das Etikett wurde dann abgetrennt und den Unterlagen des/der Patienten /Patientin beigelegt. Dabei konnte die Studienmedikation im Behälter nicht durch das Etikett eindeutig identifiziert werden.
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?	Die Zuteilung der Patienten in die Behandlungsgruppen der Induktions- und der Maintenance-Studie wurde zentral randomisiert. Dafür wurde ein computergenerierter Randomisierungsplan unter Aufsicht des Sponsors erstellt. Für eine detaillierte Beschreibung des Randomisierungsprozesses siehe Abschnitt 8b.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Ja, mit Ausnahme von 3 Patienten in der Induktionsstudie und 5 Patienten in der Maintenance-Studie b) Ja c) Ja Um die Verblindung während der Induktions- sowie der Maintenance-Studie aufrechtzuerhalten, wurden die Behälter der Studienmedikation mit einem Abreißetikett versehen, das die Studienbezeichnung, die Medikamentennummer sowie die Referenznummer enthielt. Die Medikamentennummer wurde in den Fallbericht eingetragen, wenn die Studienmedikation verabreicht wurde. Das Etikett wurde dann abgetrennt und den Unterlagen des/der Patienten /Patientin beigelegt. Dabei konnte die Studienmedikation im Behälter nicht durch das Etikett eindeutig identifiziert werden.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Studienmedikationen waren identisch in Aussehen und Verpackung
12	Statistische Methoden	

Item[a]	Charakteristikum	Studieninformation
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die primären und wichtigsten sekundären Endpunkte der Induktionsstudie (einschließlich histo-endoskopischer Schleimhautheilung) sowie der Maintenance-Studie wurden für Multiple Vergleiche kontrolliert. Der Fehler I Art wurde in jeder Studie auf einem Signifikanzniveau von 0,05 über die Endpunkte kontrolliert, die für die multiplen Vergleiche mithilfe eines vordefinierten Verfahrens für Multiples Testen kontrolliert. Ein weiteres vordefiniertes Verfahren für Multiples Testen wurde außerdem zur Unterstützung des FDA-Antrags verwendet. Dichotome Endpunkte wurden zwischen jeder Ustekinumab-Gruppe und der Placebogruppe mithilfe eines zweiseitigen Cochran-Mantel-Haenszel-Chi-Quadrat-Test mit Anpassung für Stratifikationsvariablen verglichen. Stetige Endpunkte wurden mithilfe der arithmetischen Mittel der Kovarianzanalyse oder der Kovarianzanalyse auf van-der-Waerden-Normalscores mit Anpassung für den Baseline-Wert und Stratifikationsvariablen. Die Analysen der anderen Endpunkte wurden nicht für Multiple Vergleiche angepasst. Die Ergebnisse wurden mit 95 %-Konfidenzintervallen angegeben, ohne p-Werte und Anpassung für Multiple Vergleiche. Schlussfolgerungen aus diesen Ergebnissen sind möglicherweise nicht reproduzierbar. Sofern nicht anders angegeben, basierten alle Wirksamkeitsanalysen auf dem Intention-to-Treat-Prinzip. Die Datensätze für die primären Wirksamkeitsanalysen umfassen die Patienten, die in der Induktionsstudie oder der Maintenance-Studie randomisiert wurden. Vorab spezifizierte Wirksamkeitsanalysen wurden auch für Patienten durchgeführt, die an der Maintenance-Studie teilnahmen, nachdem sie mit Verzögerung auf Ustekinumab angesprochen hatten. Zur Bewertung der Konsistenz des Behandlungseffekts für den primären Endpunkt, wurde die klinische Remission in vordefinierten Subgruppen analysiert. Die Patienten hatten dichotome Endpunkte nicht erreicht, wenn sie eine verbotene Änderung der Begleitmedikation bei Colitis Ulcerosa hatten, sich einem Stoma oder einer Kolektomie vor Woche 8 (während der Induktionsstudie) oder Woche 44 (während der Maintenance-Studie) unterzogen hatten, nach einem klinischen Schub eine Notfallmedikation verwendet hatten (während der Maintenance-Studie) oder Ustekinumab oder Placebo wegen mangelnder Wirksamkeit oder eines unerwünschten Ereignisses aufgrund einer schlimmer werdenden Krankheit (während der Maintenance-Studie) abgesetzt hatten. Für stetige Endpunkte wurden die Baseline-Werte der Induktionsstudie bei Patienten, bei denen ein Behandlungsversagen festgestellt wurde, ab dem Zeitpunkt des Ereignisses übernommen. Für dichotome Endpunkte, einschließlich aller Endpunkte die für Multiple Vergleiche kontrolliert wurden, wurden Patienten mit fehlenden Werten als Patienten, die die Endpunkte nicht erreicht haben, betrachtet. Vorab spezifizierte Sensitivitätsanalysen

Item[a]	Charakteristikum	Studieninformation
		einschließlich Methoden zur Berücksichtigung fehlender Werte wurden durchgeführt, um die Robustheit der Analysen des primären Endpunkts für beide Definitionen der klinischen Remission zu testen. In den Sicherheitsanalysen wurden die Daten von Patienten, die mindestens eine Dosis von Ustekinumab oder Placebo in der Induktionsstudie erhielten, nach der erhaltenen Substanz analysiert. Die Daten der Patienten, die mindestens eine Dosis von Ustekinumab oder Placebo in der Maintenance-Studie erhielten, wurden ebenfalls nach der zugewiesenen Studiengruppe bezüglich der Sicherheit analysiert. Die Häufigkeit und Art der unerwünschten Ereignisse wurden zusammengefasst. Analysen der Immunogenität schlossen Patienten ein, bei denen mindestens eine Blutprobe nach der Verabreichung von Ustekinumab entnommen wurde.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Der primäre Endpunkt wird sowohl in der Induktionsstudie als auch in der Maintenance-Studie für folgende Subgruppen analysiert: • Geschlecht (männlich, weiblich) • Abstammung (kaukasisch, nicht-kaukasisch) • Alter (≤ Median, > Median) • Gewicht (≤ Q1, > Q1 ≤ Q2, > Q2 ≤ Q3, > Q3) • Raucherstatus (Nichtraucher, früherer Raucher, Raucher) • Region (Asien, Osteuropa, Rest der Welt) • Krankheitsdauer (≤ 5, > 5 ≤ 15, > 15) • Ausmaß der Krankheit (limitiert, extensiv) • Schwere der Krankheit Colitis Ulcerosa (moderat: 6 ≤ Mayo Score ≤ 10, schwerwiegend: > 10) • Extraintestinale Anzeichen (abwesend, vorhanden) • CRP (≤ 3 mg/l, > 3 mg/l) • CRP (≤ Median mg/l, > Median mg/l) • CRP (≤ 10 mg/l, > 10 mg/l) • Fäkales Calprotectin (≤ 250 mg/kg, > 250 mg/kg) • Fäkales Lactoferrin (≤ 7,24 µg/g, > 7,24 µg/g) • Einnahme von oralen 5-ASA-Präparaten (ja, nein) • Einnahme von oralen Kortikosteroiden (ja, nein) • Einnahme von 6-MP/AZA/MTX (ja, nein) • Einnahme von oralen Kortikosteroiden und 6-MP/AZA/MTX (ja, nein) • Einnahme von oralen Kortikosteroiden oder 6-MP/AZA/MTX (ja, nein) • Status des biologischen Versagens (ja, nein) • Patienten mit biologischem Versagen bei: ▪ Nur Anti-TNF (Nicht zu Vedolizumab: ja, nein) ▪ Mindestens Anti-TNF (Unabhängig von Vedolizumab: ja, nein) ▪ Jeglichem Anti-TNF und Vedolizumab (ja, nein) ▪ Vedolizumab (ja, nein) • Patienten mit biologischem Versagen nur bei Anti-TNF: ▪ Primäres Nicht-Ansprechen (ja, nein) ▪ Sekundäres Nicht-Ansprechen (ja, nein) ▪ Intoleranz (ja, nein) • Patienten mit biologischem Versagen bei mindestens einem Anti-TNF:

Item[a]	Charakteristikum	Studieninformation
		▪ Primäres Nicht-Ansprechen (ja, nein) ▪ Sekundäres Nicht-Ansprechen (ja, nein) ▪ Intoleranz (ja, nein) • Patienten mit biologischem Versagen bei jeglichem Anti-TNF und Vedolizumab: ▪ Primäres Nicht-Ansprechen (ja, nein) ▪ Sekundäres Nicht-Ansprechen (ja, nein) ▪ Intoleranz (ja, nein) • Patienten mit biologischem Versagen bei Vedolizumab: ▪ Primäres Nicht-Ansprechen (ja, nein) ▪ Sekundäres Nicht-Ansprechen (ja, nein) ▪ Intoleranz (ja, nein) • Patienten ohne biologisches Versagen (Bio-naiv, Bio-erfahren)
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
		<u>Induktionsstudie:</u>
		d) Placebo: N = 319; Ustekinumab 130 mg: N = 320; Ustekinumab 6 mg/kg: N = 322;
		e) Placebo: N = 319; Ustekinumab 130 mg: N = 319; Ustekinumab 6 mg/kg: N = 322;
		f) Placebo: N = 319; Ustekinumab 130 mg: N = 320; Ustekinumab 6 mg/kg: N = 322;
		<u>Maintenance-Studie:</u> (Nur randomisierte Maintenance-Population mit N = 523)
		a) Placebo: N = 175; Ustekinumab 90 mg (alle 12 Wochen): N = 172; Ustekinumab 90 mg (alle 8 Wochen): N = 176;
		b) Keine Angaben
		c) Placebo: N = 175; Ustekinumab 90 mg (alle 12 Wochen): N = 172; Ustekinumab 90 mg (alle 8 Wochen): N = 176;

13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studienbeginn: August 2015 Studienende: August 2018

Item[a]	Charakteristikum	Studieninformation
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete regulär

a: nach CONSORT 2010.

6-MP: 6-Mercaptopurin; 6-TG: 6-Thioguanin; ASA: Aminosalicylsäure; AZA: Azathioprin ; Anti-HBc: Hepatitis B-Core-Antigen; Anti-HBs: Antikörper gegen Hepatitis B Surface Antigen; BCG: Bacille-Calmette-Guerin; CONSORT: Consolidated Standards of Reporting Trials; CRP: C-reaktives Protein; CTCAE: Common Terminology Criteria for Adverse Events; DNA: Desoxyribonukleinsäure; DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; EQ-5D: EuroQol-5 dimensions; FDA: Food and Drug Administration; HB: Hämoglobin; HBV: Hepatitis B Virus; HBsAg: HBV-Oberflächenantigen; HCV: Hepatitis C Virus; HIV: Humanes Immundefizienz-Virus; IBDQ: Inflammatory Bowel Disease Questionnaire; MTX: Methotrexat; N: Anzahl Patienten; RNA: Ribonukleinsäure; SF-36: Short Form – 36; SI: Internationales Einheitensystem; TB: Tuberkulose; TNF: Tumornekrosefaktor; TPN: Totale parenterale Ernährung; VAS: Visuelle Analogskala

Induktion

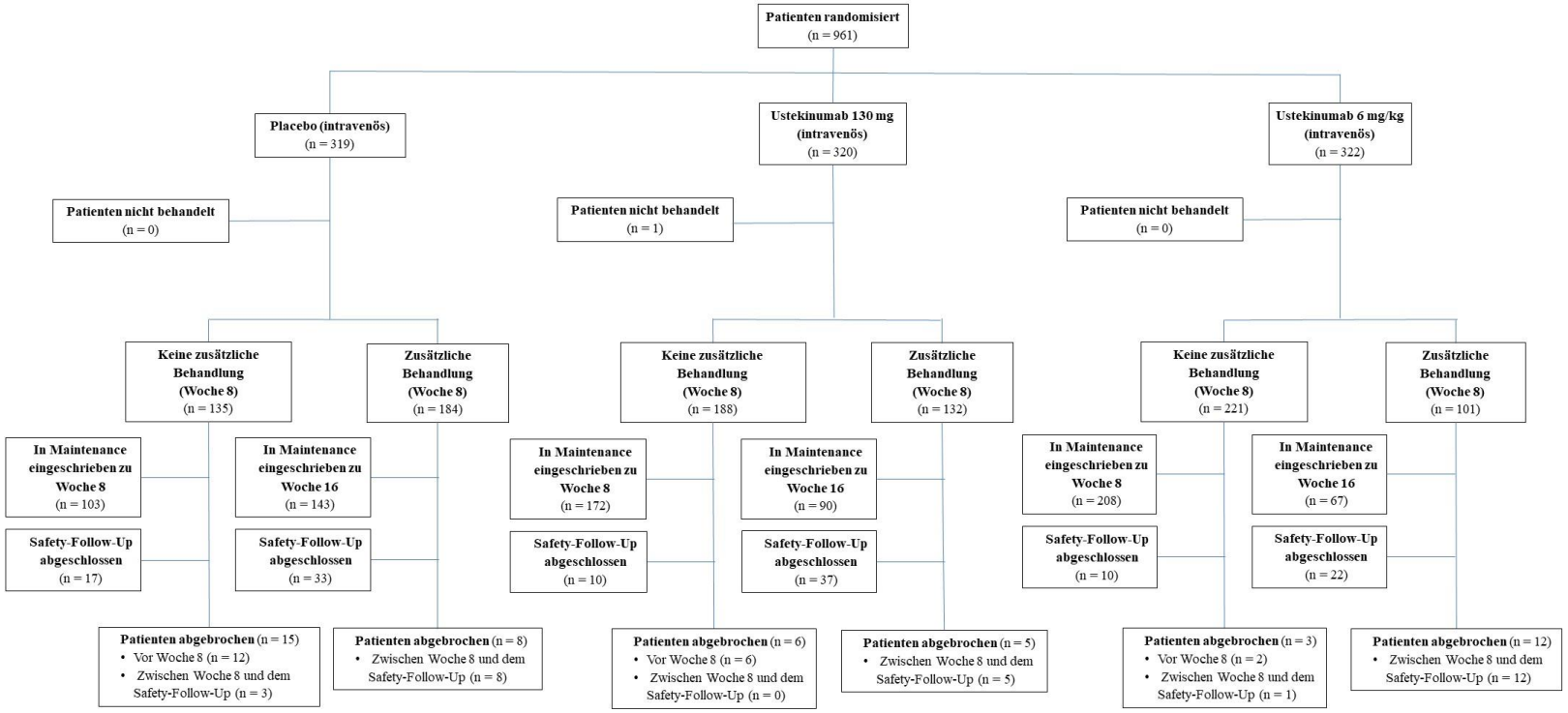

Abbildung 35: Flussdiagramm für Studie UNIFI (Induktionsphase)

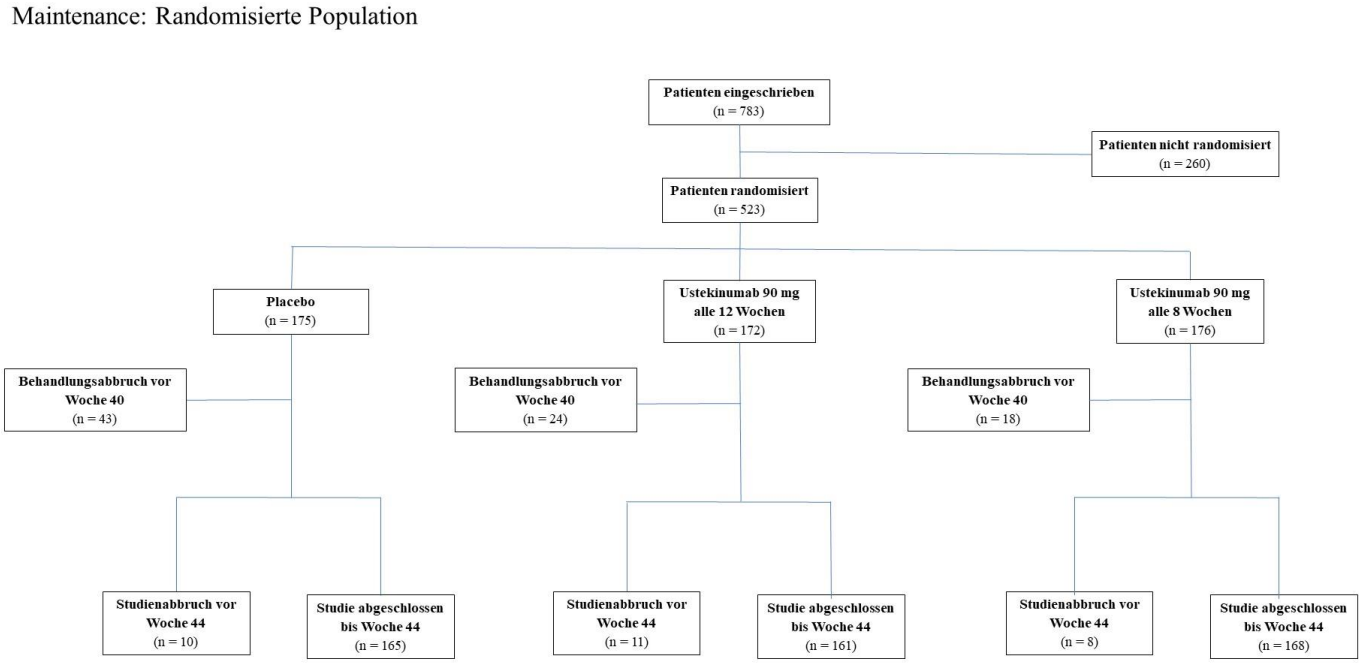

Abbildung 36: Flussdiagramm für Studie UNIFI (Erhaltungsphase; randomisierte Population)

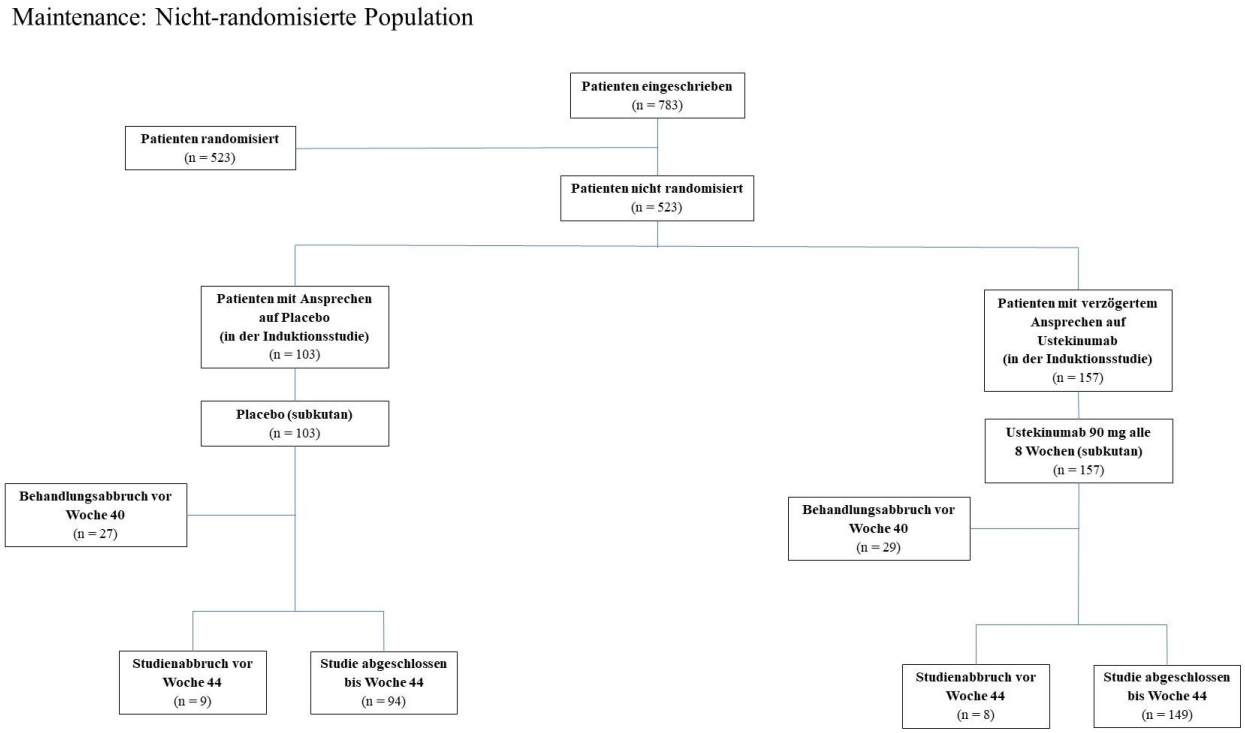

Abbildung 37: Flussdiagramm für Studie UNIFI (Erhaltungsphase; nicht randomisierte Population)

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 4-143 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie U-ACHIEVE Substudie 1

Studie: U-ACHIEVE Substudie 1

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienprotokoll	A
Studienbericht	B
Statistischer Analyseplan	C

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:
A, B, C

Es handelt sich um eine randomisierte Studie.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

<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, <u>obligate</u> Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:
A, B, C

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:
A, B, C

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:
A, B, C

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, <u>obligate</u> Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie erfolgte randomisiert mit verdeckter Gruppenzuteilung und sowohl die Patienten als auch die behandelnden Personen agierten verblindet. Es erfolgte eine ergebnisunabhängige Berichterstattung und weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Klinische Remission (nach Mayo Gesamtscore und Adaptiertem Mayo Score)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Klinisches Ansprechen (nach Mayo Gesamtscore)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Mukosaheilung

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Stuhlfrequenz (SFS) (Subscore des Mayo Gesamtscore)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Rektale Blutung (RBS) (Subscore des Mayo Gesamtscore)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: EuroQol -5 dimensions visuelle Analogskala (EQ-5D VAS)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Inflammatory Bowel Disease Questionnaire (IBDQ)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Short form – 36 (SF-36) physical health component summary score (PCS) und mental health component summary score (MCS)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Verträglichkeit

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Tabelle 4-144 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie U-ACHIEVE Substudie 2

Studie: U-ACHIEVE Substudie 2

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienprotokoll	A
Studienbericht	B
Statistischer Analyseplan	C

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:
A, B, C

Es handelt sich um eine randomisierte Studie.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

<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, <u>obligate</u> Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:
A, B, C

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:
A, B, C

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:
A, B, C

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, <u>obligate</u> Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie erfolgte randomisiert mit verdeckter Gruppenzuteilung und sowohl die Patienten als auch die behandelnden Personen agierten verblindet. Es erfolgte eine ergebnisunabhängige Berichterstattung und weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Klinische Remission (nach Mayo Gesamtscore und Adaptiertem Mayo Score)

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Klinisches Ansprechen (nach Mayo Gesamtscore)

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Mukosaheilung

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Stuhlfrequenz (SFS) (Subscore des Mayo Gesamtscore)

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Rektale Blutung (RBS) (Subscore des Mayo Gesamtscore)

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: EuroQol -5 dimensions visuelle Analogskala (EQ-5D VAS)

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Inflammatory Bowel Disease Questionnaire (IBDQ)

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Short form – 36 (SF-36) physical health component summary score (PCS) und mental health component summary score (MCS)

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Verträglichkeit

5. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

6. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Tabelle 4-145 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie U-ACCOMPLISH

Studie: U-ACCOMPLISH

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienprotokoll	A
Studienbericht	B
Statistischer Analyseplan	C

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:
A, B, C

Es handelt sich um eine randomisierte Studie.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

für nicht randomisierte Studien: 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:
A, B, C

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie erfolgte randomisiert mit verdeckter Gruppenzuteilung und sowohl die Patienten als auch die behandelnden Personen agierten verblindet. Es erfolgte eine ergebnisunabhängige Berichterstattung und weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Klinische Remission (nach Mayo Gesamtscore und Adaptiertem Mayo Score)

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Klinisches Ansprechen (nach Mayo Gesamtscore)

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Mukosaheilung

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Stuhlfrequenz (SFS) (Subscore des Mayo Gesamtscore)

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Rektale Blutung (RBS) (Subscore des Mayo Gesamtscore)

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: EuroQol -5 dimensions visuelle Analogskala (EQ-5D VAS)

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Inflammatory Bowel Disease Questionnaire (IBDQ)

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Short form – 36 (SF-36) physical health component summary score (PCS) und mental health component summary score (MCS)

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Verträglichkeit

9. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

10. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

11. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

12. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Tabelle 4-146 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UNIFI Induktionsstudie

Studie: UNIFI Induktionsstudie

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienprotokoll	A
Studienbericht	B
Statistischer Analyseplan	C

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:
A, B, C

Es handelt sich um eine randomisierte Studie.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

<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, <u>obligate</u> Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:
A, B, C

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:
A, B, C

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:
A, B, C

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, <u>obligate</u> Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie erfolgte randomisiert mit verdeckter Gruppenzuteilung und sowohl die Patienten als auch die behandelnden Personen agierten verblindet. Es erfolgte eine ergebnisunabhängige Berichterstattung und weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Klinische Remission (nach Mayo Gesamtscore und Adaptiertem Mayo Score)

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Klinisches Ansprechen (nach Mayo Gesamtscore)

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Mukosaheilung

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Stuhlfrequenz (SFS) (Subscore des Mayo Gesamtscore)

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Rektale Blutung (RBS) (Subscore des Mayo Gesamtscore)

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: EuroQol -5 dimensions visuelle Analogskala (EQ-5D VAS)

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Inflammatory Bowel Disease Questionnaire (IBDQ)

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Short form – 36 (SF-36) physical health component summary score (PCS) und mental health component summary score (MCS)

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Verträglichkeit

13. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

14. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

15. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

16. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Tabelle 4-147 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie U-ACHIEVE Substudie 3

Studie: U-ACHIEVE Substudie 3

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienprotokoll	A
Studienbericht	B
Statistischer Analyseplan	C

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:
A, B, C

Es handelt sich um eine randomisierte Studie.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

<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, <u>obligate</u> Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:
A, B, C

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:
A, B, C

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:
A, B, C

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, <u>obligate</u> Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie erfolgte randomisiert mit verdeckter Gruppenzuteilung und sowohl die Patienten als auch die behandelnden Personen agierten verblindet. Es erfolgte eine ergebnisunabhängige Berichterstattung und weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Klinische Remission (nach Mayo Gesamtscore und Adaptiertem Mayo Score)

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Klinisches Ansprechen (nach Mayo Gesamtscore)

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Mukosaheilung

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Stuhlfrequenz (SFS) (Subscore des Mayo Gesamtscore)

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Rektale Blutung (RBS) (Subscore des Mayo Gesamtscore)

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: EuroQol -5 dimensions visuelle Analogskala (EQ-5D VAS)

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Inflammatory Bowel Disease Questionnaire (IBDQ)

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Short form – 36 (SF-36) physical health component summary score (PCS) und mental health component summary score (MCS)

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Verträglichkeit

17. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A, B, C

18. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

19. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A, B, C

20. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A, B, C

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Tabelle 4-148 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UNIFI Erhaltungsstudie

Studie: UNIFI Erhaltungsstudie

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Publikation Sands et al. (2019); inklusive Anhang	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:
A
Es handelt sich um eine randomisierte Studie.

1.
für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.
für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

für nicht randomisierte Studien: 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:
A

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Studie erfolgte randomisiert mit verdeckter Gruppenzuteilung und sowohl die Patienten als auch die behandelnden Personen agierten verblindet. Es erfolgte eine ergebnisunabhängige Berichterstattung und weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Klinische Remission (nach Mayo Gesamtscore und Adaptiertem Mayo Score)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Klinisches Ansprechen (nach Mayo Gesamtscore)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Mukosaheilung

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Stuhlfrequenz (SFS) (Subscore des Mayo Gesamtscore)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Rektale Blutung (RBS) (Subscore des Mayo Gesamtscore)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: EuroQol -5 dimensions visuelle Analogskala (EQ-5D VAS)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Inflammatory Bowel Disease Questionnaire (IBDQ)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Short form – 36 (SF-36) physical health component summary score (PCS) und mental health component summary score (MCS)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

Endpunkt: Verträglichkeit

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:
A

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
A

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:
A

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Endpunkterheber agierten verblindet, es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Das Verzerrungspotenzial für diesen Endpunkt wird demzufolge als niedrig eingestuft.

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.*
  - *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: