**Dokumentvorlage, Version vom 18.04.2013**

# Dossier zur Nutzenbewertung gemäß § 35a SGB V

*Sarilumab (Kevzara®)*

Sanofi-Aventis Deutschland GmbH

**Modul 4 A**

*Rheumatoide Arthritis*

Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen

Stand: 15.08.2017

# Inhaltsverzeichnis

**Seite**

**Tabellenverzeichnis ... 4**
**Abbildungsverzeichnis ... 13**
**Abkürzungsverzeichnis ... 14**
**4 Modul 4 – allgemeine Informationen ... 17**
4.1 Zusammenfassung der Inhalte von Modul 4 ... 18
4.2 Methodik ... 34
4.2.1 Fragestellung ... 34
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung ... 37
4.2.3 Informationsbeschaffung ... 42
4.2.3.1 Studien des pharmazeutischen Unternehmers ... 42
4.2.3.2 Bibliografische Literaturrecherche ... 42
4.2.3.3 Suche in Studienregistern ... 43
4.2.3.4 Selektion relevanter Studien ... 44
4.2.4 Bewertung der Aussagekraft der Nachweise ... 45
4.2.5 Informationssynthese und -analyse ... 46
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien ... 46
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien ... 47
4.2.5.3 Meta-Analysen ... 61
4.2.5.4 Sensitivitätsanalysen ... 62
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren ... 62
4.2.5.6 Indirekte Vergleiche ... 65
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen ... 68
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel ... 68
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel ... 68
4.3.1.1.1 Studien des pharmazeutischen Unternehmers ... 68
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche ... 73
4.3.1.1.3 Studien aus der Suche in Studienregistern ... 74
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel ... 75
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel ... 76
4.3.1.2.1 Studiendesign und Studienpopulationen ... 76
4.3.1.2.2 Verzerrungspotenzial auf Studienebene ... 85
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien ... 87
4.3.1.3.1 Morbidität: Änderung der Krankheitsaktivität – RCT ... 90
4.3.1.3.1.1 Änderung des DAS28-ESR ... 90
4.3.1.3.1.2 Änderung des CDAI ... 95
4.3.1.3.1.3 Änderung des SDAI ... 99
4.3.1.3.2 Morbidität: Remission – RCT ... 103
4.3.1.3.2.1 DAS28-Remission ... 103

4.3.1.3.2.2 CDAI-Remission ... 106
4.3.1.3.2.3 SDAI-Remission ... 110
4.3.1.3.2.4 Boolesche ACR/EULAR-Remission ... 115
4.3.1.3.3 Morbidität: Niedrige Krankheitsaktivität – RCT ... 119
4.3.1.3.3.1 Niedrige Krankheitsaktivität gemäß DAS28 ... 119
4.3.1.3.3.2 Niedrige Krankheitsaktivität gemäß CDAI ... 123
4.3.1.3.3.3 Niedrige Krankheitsaktivität gemäß SDAI ... 127
4.3.1.3.4 Therapieansprechen – RCT ... 131
4.3.1.3.4.1 Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR ... 131
4.3.1.3.4.2 Ansprechen gemäß ACR-Kriterien ... 135
4.3.1.3.5 Einzelkomponenten der Scores zur Bestimmung der Krankheitsaktivität –RCT ... 145
4.3.1.3.5.1 Gelenksymptome: Druckschmerzhafte Gelenke ... 146
4.3.1.3.5.2 Gelenksymptome: Geschwollene Gelenke ... 151
4.3.1.3.5.3 Schmerzen ... 156
4.3.1.3.5.4 Globale Einschätzung des Arztes (PhGA) ... 162
4.3.1.3.5.5 Globale Einschätzung des Patienten (PtGA) ... 164
4.3.1.3.5.6 Körperlicher Funktionsstatus - HAQ-DI ... 170
4.3.1.3.5.7 Laborparameter: CRP und ESR ... 177
4.3.1.3.6 Morbidität: Morgensteifigkeit – RCT ... 181
4.3.1.3.7 Morbidität: Fatigue – RCT ... 184
4.3.1.3.8 Morbidität: Änderung des allgemeinen Gesundheitszustandes – RCT ... 190
4.3.1.3.9 Lebensqualität – RCT ... 193
4.3.1.3.9.1 SF-36 ... 193
4.3.1.3.9.2 RAID ... 213
4.3.1.3.10 Sicherheit/Verträglichkeit – RCT ... 219
4.3.1.3.11 Subgruppenanalysen – RCT ... 249
4.3.1.3.12 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien ... 255
4.3.2 Weitere Unterlagen ... 265
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien ... 265
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche ... 265
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche ... 265
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen ... 266
4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT ... 266
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT ... 268
4.3.2.2 Nicht randomisierte vergleichende Studien ... 269
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien ... 269
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien ... 269
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien ... 270
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien ... 270
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien ... 271
4.3.2.3 Weitere Untersuchungen ... 271
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen ... 272

4.3.2.3.2 Charakteristika der weiteren Untersuchungen ..... 272
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen ..... 272
4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen ..... 272
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen ..... 273
4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen ..... 273
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens ..... 274
4.4.1 Beurteilung der Aussagekraft der Nachweise ..... 274
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß ..... 274
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht ..... 283
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte ..... 284
4.5.1 Begründung für die Vorlage indirekter Vergleiche ..... 284
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen ..... 284
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen ..... 285
4.5.4 Verwendung von Surrogatendpunkten ..... 285
4.6 Liste der eingeschlossenen Studien ..... 286
4.7 Referenzliste ..... 286
**Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche ..... 294**
**Anhang 4-B : Suchstrategien – Suche in Studienregistern ..... 298**
**Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche) ..... 301**
**Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern) ..... 302**
**Anhang 4-E : Methodik der eingeschlossenen Studien – RCT ..... 317**
**Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten ..... 330**

**Tabellenverzeichnis**

**Seite**

Tabelle 4-1: Teilanwendungsgebiete 1 bis 4 und zweckmäßige Vergleichstherapie ..... 18
Tabelle 4-2: Teilanwendungsgebiete und zweckmäßige Vergleichstherapie ..... 35
Tabelle 4-3: Untergliederung der Teilanwendungsgebiete 2 und 3 anhand der MTX-Eignung ..... 36
Tabelle 4-4: Ein- und Ausschlusskriterien ..... 38
Tabelle 4-5: Patientenrelevante und ergänzend dargestellte Endpunkte der Studie MONARCH ..... 49
Tabelle 4-6: Einzelkomponenten der Scores zur Bestimmung der Krankheitsaktivität ..... 50
Tabelle 4-7: Bestimmung der Krankheitsaktivität mittels DAS28, CDAI und SDAI [34] ..... 55
Tabelle 4-8: Kriterien für ein Therapieansprechen gemäß EULAR [35] ..... 56
Tabelle 4-9: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel ..... 69
Tabelle 4-10: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel ..... 72
Tabelle 4-11: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel ..... 75
Tabelle 4-12: Studienpool – RCT mit dem zu bewertenden Arzneimittel ..... 76
Tabelle 4-13: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel ..... 77
Tabelle 4-14: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel ..... 78
Tabelle 4-15: Charakterisierung der Studienpopulation (demographische Charakteristika) – RCT mit dem zu bewertenden Arzneimittel ..... 78
Tabelle 4-16: Charakterisierung der Studienpopulation (Basischarakteristika der Erkrankung I) – RCT mit dem zu bewertenden Arzneimittel ..... 79
Tabelle 4-17: Charakterisierung der Studienpopulation (Basischarakteristika der Erkrankung II) – RCT mit dem zu bewertenden Arzneimittel ..... 80
Tabelle 4-18: Charakterisierung der Studienpopulation (MTX-Historie) – RCT mit dem zu bewertenden Arzneimittel ..... 83
Tabelle 4-19: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel ..... 85
Tabelle 4-20: Patientenrelevante und ergänzend dargestellte Endpunkte der Studie MONARCH ..... 87
Tabelle 4-21: Operationalisierung von „Änderung des DAS28-ESR“ ..... 90
Tabelle 4-22: Bewertung des Verzerrungspotenzials für „Änderung des DAS28-ESR“ in RCT mit dem zu bewertenden Arzneimittel ..... 91

Tabelle 4-23: Ergebnisse für „Änderung des DAS28-ESR (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 92
Tabelle 4-24: Ergebnisse für „Änderung des DAS28-ESR (Woche 48)“ ..... 94
Tabelle 4-25: Operationalisierung von „Änderung des CDAI“ ..... 95
Tabelle 4-26: Bewertung des Verzerrungspotenzials für „Änderung des CDAI“ in RCT mit dem zu bewertenden Arzneimittel ..... 96
Tabelle 4-27: Ergebnisse für „Änderung des CDAI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 97
Tabelle 4-28: Operationalisierung von „Änderung des SDAI“ ..... 99
Tabelle 4-29: Bewertung des Verzerrungspotenzials für „Änderung des SDAI“ in RCT mit dem zu bewertenden Arzneimittel ..... 99
Tabelle 4-30: Ergebnisse für „Änderung des SDAI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 101
Tabelle 4-31: Operationalisierung von „DAS28-Remission“ ..... 103
Tabelle 4-32: Bewertung des Verzerrungspotenzials für „DAS28-Remission“ in RCT mit dem zu bewertenden Arzneimittel ..... 103
Tabelle 4-33: Ergebnisse für „DAS28-Remission (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 104
Tabelle 4-34: Ergebnisse für „DAS28-Remission (Woche 48)“ ..... 105
Tabelle 4-35: Operationalisierung von „CDAI-Remission“ ..... 106
Tabelle 4-36: Bewertung des Verzerrungspotenzials für „CDAI-Remission“ in RCT mit dem zu bewertenden Arzneimittel ..... 106
Tabelle 4-37: Ergebnisse für „CDAI-Remission (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 108
Tabelle 4-38: Ergebnisse für „CDAI-Remission (Woche 48)“ ..... 109
Tabelle 4-39: Operationalisierung von „SDAI-Remission“ ..... 110
Tabelle 4-40: Bewertung des Verzerrungspotenzials für „SDAI-Remission“ in RCT mit dem zu bewertenden Arzneimittel ..... 111
Tabelle 4-41: Ergebnisse für „SDAI-Remission (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 112
Tabelle 4-42: Ergebnisse für „SDAI-Remission (Woche 48)“ ..... 113
Tabelle 4-43: Operationalisierung von „Boolesche ACR/EULAR-Remission“ ..... 115
Tabelle 4-44: Bewertung des Verzerrungspotenzials für „Boolesche ACR/EULAR-Remission“ in RCT mit dem zu bewertenden Arzneimittel ..... 116
Tabelle 4-45: Ergebnisse für „Boolesche ACR/EULAR-Remission (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 117
Tabelle 4-46: Ergebnisse für „Boolesche ACR/EULAR-Remission (Woche 48)“ ..... 118
Tabelle 4-47: Operationalisierung von „Niedrige Krankheitsaktivität gemäß DAS28“ ..... 119

Tabelle 4-48: Bewertung des Verzerrungspotenzials für „Niedrige Krankheitsaktivität gemäß DAS28“ in RCT mit dem zu bewertenden Arzneimittel ..... 120
Tabelle 4-49: Ergebnisse für „Niedrige Krankheitsaktivität gemäß DAS28 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 121
Tabelle 4-50: Ergebnisse für „Niedrige Krankheitsaktivität gemäß DAS28 (Woche 48)“ ... 122
Tabelle 4-51: Operationalisierung von „Niedrige Krankheitsaktivität gemäß CDAI“ ..... 123
Tabelle 4-52: Bewertung des Verzerrungspotenzials für „Niedrige Krankheitsaktivität gemäß CDAI“ in RCT mit dem zu bewertenden Arzneimittel ..... 124
Tabelle 4-53: Ergebnisse für „Niedrige Krankheitsaktivität gemäß CDAI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 125
Tabelle 4-54: Ergebnisse für „Niedrige Krankheitsaktivität gemäß CDAI (Woche 48)“..... 126
Tabelle 4-55: Operationalisierung von „Niedrige Krankheitsaktivität gemäß SDAI“ ..... 127
Tabelle 4-56: Bewertung des Verzerrungspotenzials für „Niedrige Krankheitsaktivität gemäß SDAI“ in RCT mit dem zu bewertenden Arzneimittel..... 128
Tabelle 4-57: Ergebnisse für „Niedrige Krankheitsaktivität gemäß SDAI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 129
Tabelle 4-58: Ergebnisse für „Niedrige Krankheitsaktivität gemäß SDAI (Woche 48)“ ..... 130
Tabelle 4-59: Operationalisierung von „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR“ ..... 131
Tabelle 4-60: Bewertung des Verzerrungspotenzials für „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR“ in RCT mit dem zu bewertenden Arzneimittel ..... 132
Tabelle 4-61: Ergebnisse für „Ansprechen gemäß EULAR-Kriterien (Woche 24) basierend auf DAS28-ESR“ aus RCT mit dem zu bewertenden Arzneimittel ..... 133
Tabelle 4-62: Ergebnisse für „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR (Woche 48)“ ..... 134
Tabelle 4-63: Operationalisierung von „Ansprechen gemäß ACR-Kriterien“ ..... 136
Tabelle 4-64: Bewertung des Verzerrungspotenzials für „Ansprechen gemäß ACR-Kriterien“ in RCT mit dem zu bewertenden Arzneimittel ..... 137
Tabelle 4-65: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR20 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 138
Tabelle 4-66: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR20 (Woche 48)“.. 139
Tabelle 4-67: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR50 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 140
Tabelle 4-68: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR50 (Woche 48)“.. 141
Tabelle 4-69: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR70 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 142
Tabelle 4-70: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR70 (Woche 48)“.. 143
Tabelle 4-71: Einzelkomponenten der Scores zur Bestimmung der Krankheitsaktivität ..... 145
Tabelle 4-72: Operationalisierung von „Druckschmerzhafte Gelenke“ ..... 146

Tabelle 4-73: Bewertung des Verzerrungspotenzials für „Druckschmerzhafte Gelenke“ in RCT mit dem zu bewertenden Arzneimittel ..... 147
Tabelle 4-74: Ergebnisse für „Druckschmerzhafte Gelenke – 68TJC (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 148
Tabelle 4-75: Ergebnisse für „Druckschmerzhafte Gelenke – 28TJC (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 149
Tabelle 4-76: Operationalisierung von „Geschwollene Gelenke “ ..... 151
Tabelle 4-77: Bewertung des Verzerrungspotenzials für „Geschwollene Gelenke “ in RCT mit dem zu bewertenden Arzneimittel ..... 152
Tabelle 4-78: Ergebnisse für „Geschwollene Gelenke – 66SJC (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 153
Tabelle 4-79: Ergebnisse für „Geschwollene Gelenke – 28SJC (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 154
Tabelle 4-80: Operationalisierung von „Schmerzen“ ..... 156
Tabelle 4-81: Bewertung des Verzerrungspotenzials für „Schmerzen“ in RCT mit dem zu bewertenden Arzneimittel ..... 156
Tabelle 4-82: Ergebnisse für „Schmerz (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 158
Tabelle 4-83: Ergebnisse für „Schmerz-Responder (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 160
Tabelle 4-84: Ergebnisse für „Schmerz-Responder (Woche 48)“ ..... 161
Tabelle 4-85: Operationalisierung von „PhGA“ ..... 162
Tabelle 4-86: Bewertung des Verzerrungspotenzials für „PhGA“ in RCT mit dem zu bewertenden Arzneimittel ..... 162
Tabelle 4-87: Ergebnisse für „PhGA (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 163
Tabelle 4-88: Operationalisierung von „PtGA“ ..... 164
Tabelle 4-89: Bewertung des Verzerrungspotenzials für „PtGA“ in RCT mit dem zu bewertenden Arzneimittel ..... 165
Tabelle 4-90: Ergebnisse für „PtGA (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 166
Tabelle 4-91: Ergebnisse für „PtGA-Responder (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 168
Tabelle 4-92: Ergebnisse für „PtGA-Responder (Woche 48)“ ..... 169
Tabelle 4-93: Operationalisierung von „Körperlicher Funktionsstatus – HAQ-DI“ ..... 170
Tabelle 4-94: Bewertung des Verzerrungspotenzials für „Körperlicher Funktionsstatus – HAQ-DI“ in RCT mit dem zu bewertenden Arzneimittel ..... 170
Tabelle 4-95: Ergebnisse für „Änderung des HAQ-DI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 172

Tabelle 4-96: Ergebnisse für „HAQ-DI-Responder – MCID 0,22 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 174
Tabelle 4-97: Ergebnisse für „HAQ-DI-Responder – MCID 0,375 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 175
Tabelle 4-98: Ergebnisse für „HAQ-DI-Responder – MCID 0,22 (Woche 48)“..... 176
Tabelle 4-99: Ergebnisse für „HAQ-DI-Responder – MCID 0,375 (Woche 48)“..... 176
Tabelle 4-100: Operationalisierung von „Laborparameter: CRP und ESR“..... 177
Tabelle 4-101: Bewertung des Verzerrungspotenzials für „Laborparameter: CRP und ESR“ in RCT mit dem zu bewertenden Arzneimittel ..... 178
Tabelle 4-102: Ergebnisse für „CRP (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel..... 179
Tabelle 4-103: Ergebnisse für „ESR (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel..... 180
Tabelle 4-104: Operationalisierung von „Körperlicher Funktionsstatus – Morgensteifigkeit“ ..... 181
Tabelle 4-105: Bewertung des Verzerrungspotenzials für „Körperlicher Funktionsstatus – Morgensteifigkeit“ in RCT mit dem zu bewertenden Arzneimittel..... 182
Tabelle 4-106: Ergebnisse für „Körperlicher Funktionsstatus – Morgensteifigkeit (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel..... 183
Tabelle 4-107: Operationalisierung von „Fatigue“ ..... 185
Tabelle 4-108: Bewertung des Verzerrungspotenzials für „Fatigue“ in RCT mit dem zu bewertenden Arzneimittel ..... 185
Tabelle 4-109: Ergebnisse für „Fatigue (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel..... 186
Tabelle 4-110: Ergebnisse für „Fatigue-Responder (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 188
Tabelle 4-111: Ergebnisse für „Fatigue-Responder (Woche 48)“ ..... 189
Tabelle 4-112: Operationalisierung von „Änderung des allgemeinen Gesundheitszustandes“..... 190
Tabelle 4-113: Bewertung des Verzerrungspotenzials für „Änderung des allgemeinen Gesundheitszustandes“ in RCT mit dem zu bewertenden Arzneimittel ..... 190
Tabelle 4-114: Ergebnisse für „Änderung des allgemeinen Gesundheitszustandes (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel..... 191
Tabelle 4-115: Operationalisierung von „Lebensqualität – SF-36“..... 193
Tabelle 4-116: Bewertung des Verzerrungspotenzials für „Lebensqualität – SF-36“ in RCT mit dem zu bewertenden Arzneimittel ..... 194
Tabelle 4-117: Ergebnisse für „Änderung des SF-36-PCS (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 195
Tabelle 4-118: Ergebnisse für „SF-36-PCS-Responder – MCID 2,5 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 197

Tabelle 4-119: Ergebnisse für „SF-36-PCS-Responder – MCID 5,1 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 198
Tabelle 4-120: Ergebnisse für „SF-36-PCS-Responder – MCID 7,2 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 200
Tabelle 4-121: Ergebnisse für „SF-36-PCS-Responder – MCID 2,5 (Woche 48)“ ..... 201
Tabelle 4-122: Ergebnisse für „SF-36-PCS-Responder – MCID 5,1 (Woche 48)“ ..... 202
Tabelle 4-123: Ergebnisse für „SF-36-PCS-Responder – MCID 7,2 (Woche 48)“ ..... 202
Tabelle 4-124: Ergebnisse für „Änderung des SF-36-MCS (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 203
Tabelle 4-125: Ergebnisse für „SF-36 - Körperliche Funktionsfähigkeit (*Physical Function*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 205
Tabelle 4-126: Ergebnisse für „SF-36 - Körperliche Rollenfunktion (*Role Physical*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 206
Tabelle 4-127: Ergebnisse für „SF-36 - Körperlicher Schmerz (*Bodily Pain*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 207
Tabelle 4-128: Ergebnisse für „SF-36 - Allgemeine Gesundheitswahrnehmung (*Global Health*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 208
Tabelle 4-129: Ergebnisse für „SF-36 - Vitalität (*Vitality*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 209
Tabelle 4-130: Ergebnisse für „SF-36 - Soziale Funktionsfähigkeit (*Social Function*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 210
Tabelle 4-131: Ergebnisse für „SF-36 - Emotionale Rollenfunktion (*Role Emotional*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 211
Tabelle 4-132: Ergebnisse für „SF-36 - Psychisches Wohlbefinden (*Mental Health*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 212
Tabelle 4-133: Operationalisierung von „Lebensqualität – RAID“ ..... 213
Tabelle 4-134: Bewertung des Verzerrungspotenzials für „Lebensqualität – RAID“ in RCT mit dem zu bewertenden Arzneimittel ..... 214
Tabelle 4-135: Ergebnisse für „Änderung des RAID (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 215
Tabelle 4-136: Ergebnisse für „RAID-Responder (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 217
Tabelle 4-137: Ergebnisse für „RAID-Responder (Woche 48)“ ..... 218
Tabelle 4-138: Operationalisierung von „Sicherheit/Verträglichkeit“ ..... 219
Tabelle 4-139: Bewertung des Verzerrungspotenzials für „Sicherheit/Verträglichkeit“ in RCT mit dem zu bewertenden Arzneimittel ..... 220
Tabelle 4-140: Ergebnisse für „Gesamtrate unerwünschter Ereignisse (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 221
Tabelle 4-141: Deskriptive Darstellung ausgewählter unerwünschter Ereignisse in der relevanten Studienpopulation nach MedDRA SOC und PT (Woche 24) ..... 222

Tabelle 4-142: Ergebnisse für „Gesamtrate unerwünschter Ereignisse (Woche 48)“ ..... 223
Tabelle 4-143: Deskriptive Darstellung ausgewählter unerwünschter Ereignisse in der relevanten Studienpopulation nach MedDRA SOC und PT (Woche 25 - 48) ..... 224
Tabelle 4-144: Ergebnisse für „Gesamtrate schwerwiegender unerwünschter Ereignisse (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 225
Tabelle 4-145: Ergebnisse für „Gesamtrate schwerwiegender unerwünschter Ereignisse (Woche 48)“ ..... 226
Tabelle 4-146: Ergebnisse für „Therapieabbrüche aufgrund unerwünschter Ereignisse (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 227
Tabelle 4-147: Ergebnisse für „Therapieabbrüche aufgrund unerwünschter Ereignisse (Woche 48)“ ..... 228
Tabelle 4-148: Ergebnisse für „Unerwünschte Ereignisse mit Todesfolge (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 229
Tabelle 4-149: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Infektionen (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 230
Tabelle 4-150: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Infektionen (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 232
Tabelle 4-151: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Infektionen (Woche 48)“ ..... 233
Tabelle 4-152: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Infektionen (Woche 48)“ ..... 233
Tabelle 4-153: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Leukopenien (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 234
Tabelle 4-154: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Leukopenien (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 236
Tabelle 4-155: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Leukopenien (Woche 48)“ ..... 237
Tabelle 4-156: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Leukopenien (Woche 48)“ ..... 237
Tabelle 4-157: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Leberwerterhöhungen (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 238
Tabelle 4-158: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Leberwerterhöhungen (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 240
Tabelle 4-159: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Leberwerterhöhungen (Woche 48)“ ..... 241
Tabelle 4-160: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Leberwerterhöhungen (Woche 48)“ ..... 241

Tabelle 4-161: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Lipidwerterhöhung (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 242
Tabelle 4-162: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Lipidwerterhöhung (Woche 48)“ ..... 243
Tabelle 4-163: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – Überempfindlichkeit (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 244
Tabelle 4-164: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – Überempfindlichkeit (Woche 48)“ ..... 245
Tabelle 4-165: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – Reaktionen an der Injektionsstelle (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel ..... 246
Tabelle 4-166: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – Reaktionen der Injektionsstelle (Woche 48)“ ..... 247
Tabelle 4-167: In der Studie MONARCH a priori spezifizierte Subgruppenanalysen, post-hoc-Auswertung für die relevante Studienpopulation (Woche 24) ..... 251
Tabelle 4-168: Post hoc geplante Subgruppenanalysen für die relevante Studienpopulation der Studie MONARCH, Morbidität (Woche 24) ..... 252
Tabelle 4-169: Post hoc geplante Subgruppenanalysen für die relevante Studienpopulation der Studie MONARCH, Lebensqualität (Woche 24) ..... 253
Tabelle 4-170: Post hoc geplante Subgruppenanalysen für die relevante Studienpopulation der Studie MONARCH, Sicherheit (Woche 24) ..... 253
Tabelle 4-171: Subgruppenanalyse nach Baseline-CRP für den Endpunkt Änderung des DAS28-ESR (Woche 24) ..... 254
Tabelle 4-172: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche ..... 266
Tabelle 4-173: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden ..... 266
Tabelle 4-174: Operationalisierung von <Endpunkt xxx> ..... 267
Tabelle 4-175: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche ..... 267
Tabelle 4-176: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche ..... 268
Tabelle 4-177: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien ..... 270
Tabelle 4-178: Operationalisierung von <Endpunkt xxx> ..... 270
Tabelle 4-179: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien ..... 271
Tabelle 4-180: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen ..... 272
Tabelle 4-181: Bewertung des Zusatznutzens von Sarilumab gegenüber Adalimumab als Monotherapien für Patienten mit rheumatoider Arthritis und einer MTX-Intoleranz ..... 277
Tabelle 4-182: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens ..... 284

Tabelle 4-183: Bibliografische Literaturrecherche in EMBASE ..... 295
Tabelle 4-184: Bibliografische Literaturrecherche in MEDLINE ..... 296
Tabelle 4-185: Bibliografische Literaturrecherche in der Cochrane Library ..... 297
Tabelle 4-186: Suchstrategie der Recherche in Studienregistern ..... 299
Tabelle 4-187: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – ClinicalTrials.gov ..... 302
Tabelle 4-188: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – EU-CTR ..... 306
Tabelle 4-189: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – ICTRP ..... 308
Tabelle 4-190: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – PharmNet.Bund ..... 314
Tabelle 4-191 (Anhang): Studiendesign und -methodik für Studie MONARCH ..... 318
Tabelle 4-192 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MONARCH ..... 331

## Abbildungsverzeichnis

**Seite**

Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel ..... 74
Abbildung 4-2: Schematische Darstellung des Studiendesigns der Studie MONARCH ..... 81
Abbildung 4-3: Patientenfluss der Studie MONARCH ..... 329

**Abkürzungsverzeichnis**

| Abkürzung | Bedeutung |
|---|---|
| ACPA | *Anti Citrullinated Peptide/Protein Antibodies* |
| ACR | *American College of Rheumatology* |
| ACR20/50/70 | 20%iges / 50%iges / 70%iges Ansprechen gemäß ACR-Kriterien |
| ARA | *American Rheumatism Association* |
| bDMARD | biologisches DMARD |
| BMI | *Body Mass Index* |
| CAC | *Cardiovascular Adjudication Committee* |
| CCP | cyclisches citrulliniertes Peptid |
| CDAI | *Clinical Disease Activity Index* |
| CONSORT | *Consolidated Standards of Reporting Trials* |
| CRP | C-reaktives Protein |
| csDMARD | konventionelles synthetisches DMARD |
| DAS28 | *Disease Activity Score with 28 Joint Count* |
| DAS28-CRP | DAS28 mit C-reaktivem Protein |
| DAS28-ESR | DAS28 mit Erythrozytensedimentationsrate |
| DMARD | *Disease Modifying Antirheumatic Drug* |
| DMC | Daten-Monitoring-Komitee |
| DIMDI | Deutsches Institut für Medizinische Dokumentation |
| EG | Europäische Gemeinschaft |
| EMA | *European Medicines Agency* |
| EOT | *End of Treatment* |
| EQ-5D | EuroQol-5D |
| ESR | Erythrozytensedimentationsrate |
| EULAR | European League Against Rheumatism |
| FACIT | *Functional Assessment of Chronic Illness Therapy* |
| G-BA | Gemeinsamer Bundesausschuss |
| HAQ-DI | *Health Assessment Questionnaire – Disability Index* |
| HLT | *High Level Term* |
| HrQoL | *Health related Quality of Life* |
| ICH | *International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use* |

| | |
|---|---|
| IL-6 | Interleukin-6 |
| IL-6R | Interleukin-6-Rezeptor |
| IQWiG | Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen |
| ITT | *Intention to treat* |
| IVRS | *Interactive Voice Response System* |
| JAK | Januskinase |
| KI | Konfidenzintervall |
| LDA | *Low Disease Activity* |
| LSMD | *Least-squares mean difference* |
| MCID | *Minimal Clinically Important Difference* |
| MD | Mittelwertdifferenz |
| MedDRA | *Medical Dictionary for Regulatory Activities* |
| MMRM | *Mixed-Effects Model with Repeated Measures* |
| MTC | *Mixed Treatment Comparison* |
| MTX | Methotrexat |
| MW | Mittelwert |
| OLE | *Open Label Extension* |
| OR | *Odds Ratio* |
| PhGA | *Physicians Global Assessment* |
| PT | *Preferred Term* |
| PtGA | *Patients Global Assessment* |
| qw/q2w | jede Woche / alle 2 Wochen |
| RAID | *Rheumatoid Arthritis Impact of Disease* |
| RCT | *Randomized Controlled Trial* |
| RD | *Risk Difference* |
| RF | Rheumafaktor |
| RP | randomisierte Population |
| RR | *Risk Ratio* |
| SAP | Statistischer Analyseplan |
| s.c. | subkutan |
| SD | *Standard Deviation* |
| SDAI | *Simplified Disease Activity Index* |
| SF-36 | *Short Form*-36 |

| | |
|---|---|
| SF-36-MCS | *Mental Component Summary Score* des SF-36 |
| SF-36-PCS | *Physical Component Summary Score* des SF-36 |
| SGB | Sozialgesetzbuch |
| SJC | *Swollen Joint Count* |
| SMQ | *Standardised MedDRA Queries* |
| SOC | *System Organ Class* |
| STE | *Surrogate Threshold Effects* |
| STROBE | *Strengthening the Reporting of Observational Studies in Epidemiology* |
| SUE | schwerwiegendes unerwünschtes Ereignis |
| TJC | *Tender Joint Count* |
| TNF | Tumornekrosefaktor |
| TREND | *Transparent Reporting of Evaluations with Non-Randomized Design* |
| UE | unerwünschtes Ereignis |
| UEBI | unerwünschtes Ereignis von besonderem Interesse |
| VAS | visuelle Analogskala |
| WHO | *World Health Organization* |
| zVT | zweckmäßige Vergleichstherapie |

# 4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

– Zusammenfassung (Abschnitt 4.1)

– Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)

– Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)

– eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)

– ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

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

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

## 4.1 Zusammenfassung der Inhalte von Modul 4

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

### Fragestellung

Das vorliegende Modul 4A dient der frühen Nutzenbewertung von Sarilumab (Kevzara®).

Bei der **Intervention** Sarilumab handelt es sich um einen monoklonaler Antikörper gegen den Interleukin-6-Rezeptor zur Behandlung der mittelschweren bis schweren aktiven rheumatoiden Arthritis bei erwachsenen Patienten, die auf ein oder mehrere krankheitsmodifizierende antirheumatische Arzneimittel (DMARDs) unzureichend angesprochen oder diese nicht vertragen haben. Gemäß Fachinformation wird Sarilumab in Kombination mit MTX verabreicht. Bei Patienten, bei denen eine MTX-Unverträglichkeit vorliegt oder eine Fortsetzung der Therapie mit MTX ungeeignet erscheint, kann Sarilumab als Monotherapie gegeben werden. Aus diesem Anwendungsgebiet ergeben sich gemäß Beratungsgespräch mit dem Gemeinsamen Bundesausschuss (G-BA) **vier Teilanwendungsgebiete** (siehe Tabelle 4-1).

Die Teilanwendungsgebiete 1 bis 4 werden gemeinsam im vorliegenden Modul 4A dargestellt.

Tabelle 4-1: Teilanwendungsgebiete 1 bis 4 und zweckmäßige Vergleichstherapie

| **Benennung der Teilanwendungsgebiete im Dossier** | **zweckmäßige Vergleichstherapie** |
|---|---|
| 1) [1] Patienten, die mit **einem csDMARD** vorbehandelt sind und **keine ungünstigen Prognosefaktoren**[2] aufweisen | – Alternative csDMARDs (z.B. MTX, Leflunomid) in Mono- oder Kombinationstherapie |
| 2) Patienten, die mit **einem csDMARD** vorbehandelt sind und **ungünstige Prognosefaktoren**[2] aufweisen | – bDMARD (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab) in Kombination mit MTX bzw.<br>– bDMARD (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Tocilizumab) als Monotherapie bei MTX-Unverträglichkeit |
| 3) Patienten, die mit **mehreren csDMARDs** vorbehandelt sind | – bDMARD (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab) in Kombination mit MTX bzw.<br>– bDMARD (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Tocilizumab) als Monotherapie bei MTX-Unverträglichkeit |

| **Benennung der Teilanwendungsgebiete im Dossier** | **zweckmäßige Vergleichstherapie** |
|---|---|
| 4) Patienten, die mit einem oder mehreren **bDMARDs** vorbehandelt sind | in Abhängigkeit von der Vortherapie:<br>– Wechsel des bDMARD (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab) in Kombination mit MTX bzw.<br>– Wechsel des bDMARD (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Tocilizumab) als Monotherapie bei MTX-Unverträglichkeit bzw.<br>– Rituximab bei Patienten mit schwerer rheumatoider Arthritis in Kombination mit MTX |
| [1] Anmerkung bezüglich Teilanwendungsgebiet 1: Es handelt sich dabei um Patienten, welche zwar formal von der Zulassung für Sarilumab umfasst sind, die aber im klinischen Alltag noch nicht mit einem Biologikum behandelt werden würden. Das Teilanwendungsgebiet 1 wird somit für die Nutzenbewertung von Sarilumab als nicht relevant erachtet.<br>[2] Ungünstige Prognosefaktoren wie z.B.: Nachweis von Autoantikörpern (z.B. Rheumafaktoren, hohe Spiegel von Antikörpern gegen citrullinierte Peptid-Antigene), hohe Krankheitsaktivität (nachgewiesen durch DAS bzw. DAS28-Bewertungssystem, geschwollene Gelenke, Parameter der Akute-Phase-Reaktion wie z.B. CRP, ESR), frühes Auftreten von Gelenkerosionen | |

Zur Bewertung des medizinischen Nutzens und Zusatznutzens von Sarilumab wurde die Zulassungsstudie MONARCH herangezogen, in welche Patienten der Teilanwendungsgebiete 2 und 3 eingeschlossen wurden. Da in der Studie ausschließlich die Monotherapie mit Sarilumab und Adalimumab untersucht wurde, liegt für diejenigen Patienten der Teilanwendungsgebiete 2 und 3, die für eine Kombinationstherapie mit MTX geeignet wären, keine Evidenz vor. Dementsprechend wird für die Ableitung des Zusatznutzens im Rahmen des Dossiers eine weitere Unterscheidung zwischen Patienten, die für MTX geeignet sind (Teilanwendungsgebiete 2a und 3a), und Patienten mit MTX-Intoleranz (Teilanwendungsgebiete 2b und 3b) vorgenommen.

Zusammenfassend leitet sich folgende **Fragestellung** für Modul 4A ab: Welchen Zusatznutzen hat Sarilumab als Monotherapie oder in Kombination mit MTX im Vergleich zur gewählten zweckmäßigen Vergleichstherapie bei der Behandlung der mittelschweren bis schweren aktiven rheumatoiden Arthritis bei erwachsenen Patienten, die auf ein oder mehrere DMARDs unzureichend angesprochen oder diese nicht vertragen haben, hinsichtlich der patientenrelevanten Endpunkte Mortalität, Morbidität, Lebensqualität und Sicherheit im Rahmen von RCT?

### Datenquellen

Zur Bewertung des medizinischen Nutzens und Zusatznutzens von Sarilumab wurde die Zulassungsstudie MONARCH herangezogen. Es handelt sich um eine randomisierte, aktiv kontrollierte, doppelblinde, multizentrische Studie, in der die Monotherapie mit Sarilumab im Vergleich zur Monotherapie mit Adalimumab untersucht wurde.

Die bibliographische Literaturrecherche und die Suche in Studienregistern ergaben keine weiteren Studien, die zur Bewertung des Ausmaßes des Zusatznutzens von Sarilumab herangezogen werden können.

**Ein-/Ausschlusskriterien für Studien**

Zur Selektion der Studien wurden Einschlusskriterien bezüglich Patientenpopulation, Intervention, Vergleichstherapie, Endpunkten, Studientyp, Studiendauer und Publikationstyp formuliert. Studien, die ein oder mehrere Einschlusskriterien nicht erfüllten, wurden ausgeschlossen.

Eingeschlossen wurden randomisierte, kontrollierte, doppelblinde Studien mit erwachsenen Patienten mit mittelschwerer bis schwerer aktiver rheumatoider Arthritis, die auf ein oder mehrere DMARDs unzureichend angesprochen oder diese nicht vertragen haben, und die mit Sarilumab entsprechend der Fachinformation behandelt wurden. Die Vergleichstherapie sollte ein csDMARD oder bDMARD entsprechend der Festlegung des G-BA sein. In den Studien sollte mindestens ein patientenrelevanter Endpunkt berichtet sein. Die Studiendauer sollte mindestens 24 Wochen betragen. Die Einschlusskriterien sind ausführlich in Tabelle 4-4 dargestellt.

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

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise wurden Verzerrungsaspekte auf Studienebene und für jeden patientenrelevanten Endpunkt beschrieben (detaillierte Darstellung im Anhang 4-F).

Das Design und die Methodik der eingeschlossenen Studie von Sarilumab wurden ausführlich in Abschnitt 4.3.1 und Anhang 4-E dokumentiert. Die Darstellung und Bewertung der Studien erfolgte entsprechend der CONSORT-Vorgaben (Items 2b bis 14) sowie dem CONSORT-Flow-Chart.

Die Ergebnisse hinsichtlich Wirksamkeit und Sicherheit von Sarilumab wurden in Abschnitt 4.3.1 in den entsprechenden Unterabschnitten für jeden patientenrelevanten Endpunkt tabellarisch dargestellt und zusammenfassend beschrieben. Abschließend wurde eine Zusammenfassung und Bewertung der vorliegenden Ergebnisse unter Berücksichtigung der Validität der herangezogenen Endpunkte, der Studienqualität sowie der Evidenzstufe vorgenommen. Anhand dessen wurde der medizinische Zusatznutzen sowie dessen Ausmaß und Wahrscheinlichkeit bestimmt und beschrieben.

**Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen**

Für das vorliegende Nutzenbewertungsdossier wurden drei Patientenpopulationen der Studie MONARCH ausgewertet. Die relevante Studienpopulation umfasst alle Patienten der Studie mit einer ärztlich bestätigten MTX-Intoleranz. Die Patienten der relevanten Studienpopulation gehören zu den Teilanwendungsgebieten 2b und 3b. Aus diesem Grund wurde eine Subgruppenanalyse nach Anzahl der csDMARD-Vortherapien durchgeführt, um zwischen Patienten zu unterscheiden, die vor Eintritt in die Studie mit einem csDMARD oder

mindestens zwei csDMARDs behandelt wurden. Da für keinen der Endpunkte eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien bestand, können die Ergebnisse der relevanten Studienpopulation für die Ableitung des Zusatznutzens für die Teilanwendungsgebiete 2b und 3b gleichermaßen herangezogen werden. Im Folgenden sind somit die Ergebnisse für die patientenrelevanten Endpunkte der relevanten Studienpopulation zusammengefasst.

***Morbidität***

*Änderung der Krankheitsaktivität*

In der Studie MONARCH wurde die Krankheitsaktivität der Patienten mittels DAS28-ESR, und CDAI ermittelt. Post hoc wurde zusätzlich eine Auswertung des SDAI vorgenommen. Die Änderung der Krankheitsaktivität nach 24 Wochen gemessen mittels DAS28-ESR war der primäre Endpunkt der Studie.

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verminderung der Krankheitsaktivität unabhängig vom verwendeten Messinstrument.

Bezüglich des DAS28-ESR zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung der Krankheitsaktivität (LSMD [95% KI] p-Wert: -1,253 [-1,660; -0,846] < 0,0001). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g belegt, dass es sich dabei um einen klinisch relevanten Effekt mit mittlerer Effektstärke handelt (Hedges´g [95% KI]: -0,990 [-1,311; -0,668]).

Auch bezüglich des CDAI zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung (LSMD [95% KI] p-Wert: -5,005 [-7,998; -2,011] 0,0012). Dabei handelt es sich ebenfalls um einen klinisch relevanten Effekt mit kleiner Effektstärke (Hedges´g [95% KI]: -0,536 [-0,856; -0,215]).

Des Weiteren war bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung des SDAI zu beobachten (LSMD [95% KI] p-Wert: -5,922 [-9,116; -2,727] 0,0003). Dabei handelt es sich ebenfalls um einen klinisch relevanten Effekt mit kleiner Effektstärke (Hedges´g [95% KI]: -0,596 [-0,917; -0,274]).

*Remission*

In der Studie MONARCH wurden verschiedene Operationalisierungen für den Endpunkt Remission ausgewertet.

Das Kriterium DAS28-ESR < 2,6 erfüllten 29,9 % der Patienten unter Sarilumab, wohingegen nur 6,1 % der Patienten unter Adalimumab diesen Schwellenwert unterschritten. Damit war unter Sarilumab die Wahrscheinlichkeit, eine Remission gemäß DAS28-ESR zu erreichen,

gegenüber Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 4,39 [1,78; 10,85] 0,0014).

Bei Anwendung der strengeren CDAI- und SDAI- sowie der booleschen ACR/EULAR-Kriterien für eine Remission erreichten in beiden Behandlungsarmen deutlich weniger Patienten die Remission als bei Anwendung des DAS28. Für alle drei Remissions-Kriterien zeigte sich nach 24 Wochen Behandlung ein numerischer Vorteil für Sarilumab gegenüber Adalimumab. Der Behandlungsunterschied war jedoch in keinem Fall signifikant (RR [95 % KI] p-Wert für CDAI-Remission: 3,37 [0,73; 15,45] 0,1182; für SDAI-Remission: 3,75 [0,83; 16,88] 0,0844; für ACR/EULAR-Remission: 2,20 [0,43; 11,16] 0,3423).

*Niedrige Krankheitsaktivität*

Das Erreichen einer niedrigen Krankheitsaktivität stellt insbesondere für Patienten mit einer langjährigen schweren RA, wie sie in die MONARCH Studie eingeschlossen wurden, ein relevantes Therapieziel dar. Nach 24-wöchiger Behandlung zeigte sich mit allen drei Messinstrumenten ein deutlicher Vorteil für Sarilumab gegenüber Adalimumab bei der Zahl der Patienten, die eine niedrige Krankheitsaktivität erreichten.

Nach 24 Wochen Behandlung erzielten 44,8 % der mit Sarilumab behandelten Patienten und 9,8 % der mit Adalimumab behandelten Patienten einen DAS28-ESR < 3,2. Dieser Behandlungsunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 4,22 [2,10; 8,46] < 0,0001).

Die Wahrscheinlichkeit, nach 24 Wochen einen CDAI ≤ 10 oder SDAI ≤ 11 zu erreichen, war unter der Behandlung mit Sarilumab ebenfalls signifikant größer als unter Behandlung mit Adalimumab (RR [95 % KI] p-Wert für CDAI: 2,39 [1,44; 3,97] 0,0008 und RR [95 % KI] p-Wert für SDAI: 2,29 [1,40; 3,74] 0,0010). Die Behandlungseffekte sind über alle drei Messinstrumente konsistent.

*Therapieansprechen*

Zur Quantifizierung des Therapieansprechens wurden die Kriterien der EULAR, basierend auf dem DAS28-ESR, und die Kriterien des ACR angewendet. Nach 24-wöchiger Behandlung zeigte sich über alle Auswertungen ein Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich des Anteils der Patienten mit einem Ansprechen auf die Behandlung.

Die Wahrscheinlichkeit für ein gutes bis moderates Ansprechen gemäß EULAR-Kriterien war unter Sarilumab im Vergleich zu Adalimumab signifikant erhöht (RR [95 % KI] p-Wert: 1,22 [1,02; 1,46] 0,0328). Dabei erreichten 44,8 % der mit Sarilumab behandelten Patienten und 9,8 % der mit Adalimumab behandelten Patienten ein gutes Ansprechen. Demgegenüber zeigten 32,9 % der mit Adalimumab behandelten Patienten kein Ansprechen gemäß EULAR-Kriterien, während es unter Sarilumab nur 18,4 % waren.

Die ACR-Kriterien für das Therapieansprechen definieren Schwellenwerte für eine 20%ige, 50%ige und 70%ige Verbesserung der Krankheitsaktivität (ACR20, ACR50, ACR70). Für die

Patienten der Studie MONARCH, die bei Studieneinschluss sämtlich eine hohe Krankheitsaktivität aufwiesen, stellt eine Verbesserung um 20 % jedoch vermutlich keinen ausreichenden Therapieerfolg dar. Aus diesem Grund wird die Auswertung zum ACR20 als nicht patientenrelevant eingestuft, aber aus Transparenzgründen mit dargestellt. Eine 50%ige oder gar 70%ige Verringerung der Krankheitsaktivität ist hingegen als klinisch relevantes Therapieansprechen einzuschätzen.

Die Wahrscheinlichkeit, ACR20 zu erreichen, war unter Sarilumab im Vergleich zu Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 1,29 [1,03; 1,61] 0,0296). Die Wahrscheinlichkeit, ACR50 zu erreichen, war ebenfalls unter Sarilumab im Vergleich zu Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 1,73 [1,14; 2,63] 0,0104). 47,1 % der Patienten unter Sarilumab erreichten ACR50 gegenüber 26,8 % der Patienten unter Adalimumab. Auch die Wahrscheinlichkeit, ACR70 zu erreichen, war unter Sarilumab statistisch signifikant größer als unter Adalimumab (RR [95 % KI] p-Wert: 2,62 [1,17; 5,85] 0,0194). 24,1 % der Patienten unter Sarilumab und 8,5 % der Patienten unter Adalimumab erreichten ACR70.

*Gelenksymptome*

Die Anzahl der druckschmerzhaften Gelenke (TJC, *Tender Joint Counts*) wurde basierend auf 68 und 28 definierten Gelenken bestimmt. Bei der Auswertung der druckschmerzhaften Gelenke ausgehend sowohl von 68 als auch 28 Gelenken zeigte sich nach 24 Wochen Behandlung jeweils ein numerischer Vorteil für Sarilumab gegenüber Adalimumab, jedoch keine statistisch signifikanten Unterschiede (LSMD [95 % KI] p-Wert für 68TJC: -1,281 [-4,314; 1,752] 0,4056 und LSMD [95 % KI] p-Wert für 28TJC: -1,452 [-3,126; 0,222] 0,0887).

Die Anzahl der geschwollenen Gelenke (SJC, *Swollen Joint Counts*) wurde basierend auf 66 und 28 definierten Gelenken bestimmt. Bei der Auswertung der geschwollenen Gelenke ausgehend von 66 Gelenken zeigte sich nach 24 Wochen Behandlung ein numerischer Vorteil für Sarilumab gegenüber Adalimumab, jedoch kein statistisch signifikanter Unterschied (LSMD [95 % KI] p-Wert: -1,513 [-3,080; 0,054] 0,0584). Bei der Auswertung der geschwollenen Gelenke ausgehend von 28 Gelenken zeigte sich jedoch ein statistisch signifikanter Vorteil für Sarilumab gegenüber Adalimumab (LSMD [95 % KI] p-Wert: -1,333 [-2,336; -0,329] 0,0095). Anhand der standardisierten Mittelwertdifferenz Hedges`g kann allerdings keine Aussage zur klinischen Relevanz dieses Effekts abgeleitet werden, da das Konfidenzintervall die -0,2 einschließt (Hedges´g [95% KI]: -0,431 [-0,755; -0,107]).

*Schmerzen*

Nach 24 Wochen Behandlung zeigte sich eine Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verringerung von Schmerzen.

Die Schmerzintensität zu Beginn der Studie wurde von den Patienten durchschnittlich mit etwa 71 mm auf der 100 mm-VAS angegeben, so dass man davon ausgehen kann, dass die Schmerzen zu Beginn der Studie als starke Schmerzen empfunden wurden. Nach 24-wöchiger Behandlung zeigte sich bei den Patienten unter Sarilumab im Vergleich zu den Patienten

unter Adalimumab eine statistisch signifikant stärkere Verringerung der Schmerzen (LSMD [95% KI] p-Wert: -11,437 [-18,456; -4,419] 0,0016). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g belegt, dass es sich dabei um einen klinisch relevanten Effekt mit kleiner Effektstärke handelt (Hedges´g [95% KI]: -0,524 [-0,846; -0,202]).

In einer Responderanalyse wurde des Weiteren untersucht, wie viele Patienten eine minimale klinische relevante Verringerung ihrer Schmerzen um mindestens 20,4 mm erreichten. Dieses Kriterium erfüllten 52,4 % der Patienten unter Adalimumab und 58,6 % der Patienten unter Sarilumab. Dieser Behandlungsunterschied für die minimale klinische relevante Verbesserung ist statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,07 [0,83; 1,39] 0,5931).

*Globale Einschätzung des Patienten (PtGA)*

Nach 24-wöchiger Behandlung zeigte sich hinsichtlich der globalen Einschätzung der Patienten zu ihrer Krankheitsaktivität (PtGA) ein Vorteil von Sarilumab gegenüber Adalimumab.

In der relevanten Studienpopulation ergab sich ein statistisch signifikanter Behandlungsunterschied zugunsten von Sarilumab (LSMD [95 % KI] p-Wert: -12,677 [-19,486; -5,868] 0,0003). Die Betrachtung der standardisierten Mittelwertdifferenz belegt, dass es sich dabei um einen klinisch relevanten Effekt mit kleiner Effektstärke handelt (Hedges´g [95% KI]: -0,597 [-0,917; -0,276]).

In einer Responderanalyse wurde des Weiteren untersucht, wie viele Patienten eine minimale klinisch relevante Verbesserung beim PtGA um mindestens 18,4 mm erreichten. Dieses Kriterium erfüllten 50,0 % der Patienten unter Adalimumab und 60,9 % der Patienten unter Sarilumab. Dieser Behandlungsunterschied für eine minimale klinisch relevante Verbesserung ist statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,18 [0,90; 1,55] 0,2196).

*Körperlicher Funktionsstatus – HAQ-DI*

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verbesserung der körperlichen Funktion gemessen mittels HAQ-DI.

Die Patienten unter Sarilumab erreichten im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Senkung des HAQ-DI nach 24 Wochen (LSMD [95% KI] p-Wert: -0,214 [-0,406; -0,022] 0,0293). Die klinische Relevanz dieser mittleren Verbesserung kann jedoch nicht abgeleitet werden kann, da das Konfidenzintervall der standardisierten Mittelwertdifferenz die -0,2 einschließt (Hedges´g [95% KI]: -0,352 [-0,668; -0,036]).

Des Weiteren wurden zwei Responderanalysen bezüglich einer minimalen klinisch relevanten Verbesserung um 0,22 und 0,375 Punkte durchgeführt.

66,7 % der Patienten unter Sarilumab und 48,8 % der Patienten unter Adalimumab erreichten eine Verbesserung im HAQ-DI um mindestens 0,22 Punkte. Die Wahrscheinlichkeit, eine

Verbesserung um 0,22 Punkte zu erreichen, war damit in der Sarilumab-Gruppe signifikant höher als in der Adalimumab-Gruppe (RR [95% KI] p-Wert: 1,37 [1,05; 1,78] 0,0213).

Eine Verbesserung im HAQ-DI um mindestens 0,375 Punkte erreichten 62,1 % der Patienten unter Sarilumab gegenüber 41,5 % der Patienten unter Adalimumab. Dieser Gruppenunterschied war ebenfalls statistisch signifikant (RR [95% KI] p-Wert: 1,49 [1,11; 2,01] 0,0090).

*Morgensteifigkeit*

Nach 24-wöchiger Behandlung zeigte sich ein Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verminderung der Morgensteifigkeit.

Die Behandlung mit Sarilumab führte zu einer signifikant stärkeren Verringerung der Morgensteifigkeit als die Behandlung mit Adalimumab (LSMD [95 % KI] p-Wert: -10,967 [-18,841; -3,094] 0,0066). Die klinische Relevanz dieses Gruppenunterschieds kann jedoch nicht abgeleitet werden, da das Konfidenzintervall der standardisierten Mittelwertdifferenz Hedges´g die -0,2 einschließt (Hedges´g [95 % KI]: -0,456 [-0,783; -0,129]).

*Fatigue*

Nach 24 Wochen Behandlung zeigte sich eine Überlegenheit von Sarilumab gegenüber Adalimumab hinsichtlich der Verringerung der Fatigue gemessen mittels FACIT-Fatigue Fragebogen.

Bei Auswertung des FACIT-Fatigue zeigte sich eine signifikant stärkere Verbesserung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab (LSMD [95 % KI] p-Wert: 3,711 [0,814; 6,608] 0,0124). Die klinische Relevanz dieser mittleren Verbesserung kann jedoch nicht abgeleitet werden kann, da das Konfidenzintervall der standardisierten Mittelwertdifferenz die 0,2 einschließt (Hedges´g [95 % KI]: 0,415 [0,091; 0,740]).

Des Weiteren wurde eine Responderanalyse durchgeführt, um den Anteil der Patienten, die eine minimale klinisch relevante Verbesserung im FACIT-Fatigue um mindestens 4 Punkte erreicht haben, zu bestimmen. Dieses Kriterium erfüllten in Woche 24 66,7 % der Patienten unter Sarilumab, während es im Adalimumab-Arm 52,4 % der Patienten waren. Bei Betrachtung des RR ist dieser Gruppenunterschied statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,27 [0,99; 1,64] 0,0631).

*Änderung des allgemeinen Gesundheitszustandes*

Nach 24-wöchiger Behandlung zeigte sich ein Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verbesserung des allgemeinen Gesundheitszustandes gemessen mittels EQ-5D-VAS.

In der relevanten Studienpopulation erreichten die Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine signifikant stärkere Verbesserung ihres allgemeinen Gesundheitszustandes (LSMD [95 % KI] p-Wert: 9,244 [2,683; 15,806] 0,0061). Die

klinische Relevanz dieses Behandlungsunterschieds bleibt jedoch unklar, da das Konfidenzintervall der standardisierten Mittelwertdifferenz die 0,2 einschließt (Hedges´g [95 % KI]: 0,461 [0,134; 0,788]).

***Gesundheitsbezogene Lebensqualität***

*SF-36*

Die Auswertung der körperlichen Summenskala des SF-36 (SF-36-PCS) ergab nach 24-wöchiger Behandlung einen Vorteil von Sarilumab gegenüber Adalimumab.

Die Lebensqualität gemäß SF-36-PCS verbesserte sich unter Sarilumab signifikant stärker als unter Adalimumab (LSMD [95 % KI] p-Wert: 3,640 [1,399; 5,882] 0,0016). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges´g zeigt, dass es sich dabei um einen klinisch relevanten Effekt von kleiner Effektstärke handelt (Hedges´g [95 % KI]: 0,532 [0,204; 0,860]).

Für Patienten mit rheumatoider Arthritis wurden für die körperliche Summenskala des SF-36 verschiedene MCIDs (2,5, 5,1 und 7,2 Punkte) als klinisch relevante Verbesserung definiert. Der Anteil der Patienten, die diese Schwellenwerte nach 24-wöchiger Behandlung erreichten, wurde in Responderanalysen bestimmt.

63,2 % der Patienten unter Sarilumab erzielten nach 24 Wochen eine Verbesserung der SF-36-PCS um mindestens 2,5 Punkte, während es unter Adalimumab 51,2 % der Patienten waren. Bei Betrachtung des RR ist dieser Gruppenunterschied statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,23 [0,95; 1,59] 0,1138).

Eine Verbesserung der SF-36-PCS um mindestens 5,1 Punkte erreichten nach 24-wöchiger Behandlung 57,5 % der Patienten unter Sarilumab, während es unter Adalimumab 43,9 % der Patienten waren. Dieser Behandlungsunterschied ist statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,30 [0,97; 1,75] 0,0803).

Des Weiteren erzielten 49,4 % der Patienten unter Sarilumab eine Verbesserung der SF-36-PCS um mindestens 7,2 Punkte, während es unter Adalimumab 35,4 % der Patienten waren. Dieser Behandlungsunterschied ist bei Betrachtung des RR statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,38 [0,96; 1,97] 0,0788).

Die Auswertung der psychischen Summenskala des SF-36 (SF-36-MCS) nach 24 Wochen Behandlung zeigte in beiden Behandlungsarmen eine Verbesserung von vergleichbarem Ausmaß. Es ergaben sich keine statistisch signifikanten Unterschiede zugunsten einer der beiden Therapien (LSMD [95 % KI] p-Wert: 2,437 [-0,806; 5,681] 0,1397).

*RAID*

Nach 24-wöchiger Behandlung zeigte sich ein deutlicher Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verbesserung der Lebensqualität gemessen mittels des krankheitsspezifischen Fragebogens RAID.

Die Patienten unter Sarilumab erreichten nach 24 Wochen Behandlung eine statistisch signifikant stärkere Verbesserung ihrer Lebensqualität als die Patienten unter Adalimumab (LSMD [95 % KI] p-Wert: -1,346 [-2,022; -0,671] 0,0001). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges´g zeigt, dass es sich dabei um einen klinisch relevanten Effekt von kleiner Effektstärke handelt (Hedges´g [95 % KI]: -0,651 [-0,978; -0,325]).

Des Weiteren wurde eine Responderanalyse hinsichtlich der minimalen klinisch relevanten Verbesserung des RAID um mindestens 3 Punkte durchgeführt. Dieses Kriterium erfüllten 49,4 % der Patienten unter Sarilumab, während es unter Adalimumab 30,5 % der Patienten waren. Dieser Gruppenunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 1,62 [1,10; 2,37] 0,0142).

***Sicherheit/Verträglichkeit***

Zusammenfassend zeigte sich ein vergleichbares Sicherheitsprofil von Sarilumab und Adalimumab innerhalb der 24-wöchigen randomisierten Behandlungsphase.

Die Gesamtrate der unerwünschten Ereignisse war in beiden Behandlungsgruppen nach 24 Wochen vergleichbar. Bei 66,7 % der mit Sarilumab behandelten Patienten sowie bei 68,3 % der mit Adalimumab behandelten Patienten wurden unerwünschte Ereignisse dokumentiert. Somit ergab sich bei einem RR von 0,98 kein statistisch signifikanter Gruppenunterschied ([95 % KI] p-Wert: [0,79; 1,20] 0,8216). Am häufigsten wurden in beiden Behandlungsarmen Infektionen verzeichnet. In der Mehrzahl der Fälle handelte es sich dabei um Atemwegsinfekte. Weitere häufige unerwünschte Ereignisse bei den mit Sarilumab behandelten Patienten waren Erkrankungen des Blutes und des Lymphsystems. Dabei handelte es sich fast ausschließlich um Neutropenien.

Schwerwiegende unerwünschte Ereignisse traten innerhalb der 24-wöchigen randomisierten Behandlungsphase selten auf. Es ergab sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (RR [95 % KI] p-Wert: 0,75 [0,21; 2,72] 0,6657).

Therapieabbrüche aufgrund von unerwünschten Ereignissen traten innerhalb der 24-wöchigen randomisierten Behandlungsphase bei 8,0 % der mit Sarilumab behandelten und 3,7 % der mit Adalimumab behandelten Patienten auf. Der Gruppenunterschied war bei einem RR von 2,2 statistisch nicht signifikant ([95 % KI] p-Wert: [0,59; 8,26] 0,2421). Die häufigste Ursache für Therapieabbrüche unter Sarilumab waren Neutropenien.

Es gab einen Todesfall im Sarilumab-Arm innerhalb der 24-wöchigen Behandlungsphase. Dabei handelte es sich um ein akutes Herzversagen nach Aortendissektion.

Als unerwünschte Ereignisse von besonderem Interesse wurden Infektionen, Leukopenien, Thrombozytopenien, Leberwerterhöhungen, Divertikulitis, gastrointestinale Ulzerationen, Erhöhung der Lipidwerte, Überempfindlichkeit, Anaphylaxie, Reaktionen an der Injektionsstelle, Krebserkrankungen, Lupus-ähnliches Syndrom und demyelinisierende

Erkrankung präspezifiziert. Für die meisten dieser Ereignisse ist bekannt, dass die Behandlung mit einem TNF- oder IL-6R-Antagonisten das Risiko für diese Ereignisse erhöht.

Infektionen traten in der doppelblinden Behandlungsphase bei 28,7 % der mit Sarilumab behandelten bzw. 28,0 % der mit Adalimumab behandelten Patienten auf. Das Risiko für das Auftreten einer Infektion war somit in den Behandlungsgruppen vergleichbar (RR [95 % KI] p-Wert: 1,02 [0,63; 1,66] 0,9212). Bei zwei Patienten im Sarilumab-Arm wurden die Infektionen als schwerwiegende Infektionen eingestuft. In beiden Fällen war die Infektion nicht mit einer Neutropenie assoziiert.

Leukopenien traten innerhalb von 24 Wochen unter Sarilumab bei 16,1 % der Patienten auf, unter Adalimumab waren im Gegensatz nur 2,4 % der Patienten betroffen. Bis auf einen Fall handelte es sich dabei um Neutropenien. Der Gruppenunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 6,60 [1,54; 28,3] 0,0112). Ein Fall im Sarilumab-Arm wurde als schwerwiegend eingestuft. Bei diesem Patienten lag eine Überdosierung von Sarilumab aufgrund der Nichteinhaltung des Behandlungsintervalls von zwei Wochen vor.

Eine Erhöhung der Lebertransaminasen trat bei 6,9 % der Patienten im Sarilumab-Arm und bei 3,7 % der Patienten im Adalimumab-Arm auf. Der Gruppenunterschied ist bei einem RR von 1,89 statistisch nicht signifikant ([95 % KI] p-Wert: [0,49; 7,32] 0,3589).

Eine Lipidwerterhöhung wurde lediglich im Adalimumab-Arm bei 4,9 % der Patienten festgestellt, unter Sarilumab traten keine Lipidwerterhöhungen auf. Der Gruppenunterschied zwischen Sarilumab und Adalimumab ist statistisch signifikant (Peto-OR [95 % KI] p-Wert: 0,12 [0,02; 0,89] 0,0376).

Überempfindlichkeit trat bei 1,1 % der Sarilumab-Patienten und 4,9 % der Adalimumab-Patienten auf. Der Gruppenunterschied ist mit einem RR von 0,24 statistisch nicht signifikant ([95 % KI] p-Wert: [0,03; 2,08] 0,1926). In der Studie traten keine anaphylaktischen Reaktionen auf.

Reaktionen an der Injektionsstelle traten bei 8,0 % der mit Sarilumab behandelten Patienten und 4,9 % der mit Adalimumab behandelten Patienten auf. Der Gruppenunterschied ist bei einem RR von 1,65 statistisch nicht signifikant ([95 % KI] p-Wert: [0,50; 5,45] 0,4107).

Thrombozytopenien, Divertikulitis und gastrointestinale Ulzerationen, Krebserkrankungen, demyelinisierende Erkrankungen sowie das Lupus-ähnliche Syndrom traten in der 24-wöchigen Behandlungsphase nicht auf.

***Auswertungen nach 48 Wochen (offene Extensionsphase)***

Die Ergebnisse der offenen Extensionsphase zeigten eine langanhaltende Wirksamkeit von Sarilumab in Bezug auf die Verbesserung der Morbidität und Lebensqualität sowie ein im Vergleich zur 24-wöchigen doppelblinden Behandlungsphase gleichbleibendes Sicherheitsprofil. Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bestätigen

somit die oben zusammengefassten Ergebnisse der 24-wöchigen doppelblinden Behandlungsphase.

*Morbiditätsendpunkte*

In der offenen Extensionsphase kam es unter Sarilumab zu einer weiteren Verringerung der Krankheitsaktivität gemessen mittels DAS28-ESR. Auch die Anzahl der Patienten, die eine Remission erreichten, stieg nach Woche 24 weiter an, unabhängig von den angewendeten Remissionskriterien (DAS28-ESR, CDAI, SDAI sowie booleschen ACR/EULAR-Kriterien). Die Zahl der Patienten, die nach 48 Wochen Behandlung eine niedrige Krankheitsaktivität, gemessen mittels DAS28-ESR, CDAI oder SDAI, erreichten, nahm ebenfalls in der offenen Extensionsphase weiter zu. Der Anteil der Patienten mit einem guten Therapieansprechen gemäß EULAR-Kriterien stieg ebenso wie der Anteil der Patienten mit einem 50%igen oder 70%igen Therapieansprechen gemäß ACR-Kriterien weiter an.

Die Auswertungen nach 48 Wochen hinsichtlich einer klinisch relevanten Verbesserung der Schmerzen und der globalen Einschätzung der Patienten (PtGA) zeigten gleichbleibende Ergebnisse im Vergleich zur Auswertung nach Woche 24. Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bezüglich einer klinisch relevanten Verbesserung des körperlichen Funktionsstatus gemäß HAQ-DI sowie der Fatigue gemäß FACIT-Fatigue zeigten lediglich einen leichten Rückgang der jeweiligen Responderzahlen im Vergleich zu den Auswertungen nach 24 Wochen.

*Gesundheitsbezogene Lebensqualität*

Die Auswertungen hinsichtlich einer klinisch relevanten Verbesserung der Lebensqualität unter Sarilumab zeigten nach 48 Wochen für den RAID-Fragebogen einen leichten Rückgang der Responderzahlen im Vergleich zu den Auswertungen nach 24 Wochen und gleichbleibende Ergebnisse für die körperliche Summenskala des SF-36.

*Sicherheit/Verträglichkeit*

Die Auswertung der unerwünschten Ereignisse bis Woche 48 bestätigte das Sicherheitsprofil von Sarilumab, das sich bereits in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat.

Die Gesamtraten der UE und der SUE sowie der Therapieabbrüche aufgrund von UE in der offenen Extensionsphase (Woche 25 bis 48) blieben im Vergleich zu den ersten 24 Wochen der Behandlung konstant. Es kam nicht zu einer Zunahme schwerwiegender unerwünschter Ereignisse. Todesfälle traten bei den Patienten der relevanten Studienpopulation in der offenen Extensionsphase nicht auf.

Die Zahl der Infektionen blieb bis Woche 48 konstant und es war keine Zunahme schwerwiegender Infektionen zu beobachten. Die Rate der Leukopenien war in der offenen Extensionsphase etwas niedriger als die Rate der Leukopenien innerhalb der 24-wöchigen doppelblinden Behandlungsphase. Keine der aufgetretenen Leukopenien nach Woche 24 wurde als schwerwiegend eingestuft. Auch die EMA hat für die Gesamtschau der Zulassungsstudien von Sarilumab bestätigt, dass sowohl die Raten der Neutropenien und

schwerwiegenden Neutropenien als auch die Raten der Infektionen und schwerwiegenden Infektionen im Zeitverlauf nicht ansteigen.

Bezüglich der weiteren UE von besonderem Interesse (Leberwerterhöhungen, Lipidwerterhöhungen, Überempfindlichkeitsreaktionen sowie Reaktionen an der Injektionsstelle) zeigten sich keine relevanten Änderungen nach Woche 24.

***Subgruppenanalysen***

Die post hoc durchgeführten Subgruppenanalysen ergaben keine klinisch relevante Effektmodifikation. Somit kann die Bewertung des Zusatznutzens auf Basis der gesamten relevanten Studienpopulation erfolgen.

**Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen**

***Endpunktkategorie Morbidität***

Die Behandlung mit Sarilumab führte im Vergleich mit der zweckmäßigen Vergleichstherapie Adalimumab zu einer deutlichen Verminderung der Krankheitslast und Abschwächung schwerwiegender Symptome.

Bei den Patienten der Studie MONARCH handelte es sich um Patienten mit einer schweren rheumatoiden Arthritis. Für diese Patienten stellt das Erreichen einer niedrigen Krankheitsaktivität das vorrangige Therapieziel dar, da ein Großteil dieser Patienten den Zustand der Remission nicht mehr erreichen kann. Alle Patienten hatten zu Beginn der Studie eine hohe Krankheitsaktivität, starke Schmerzen sowie eine deutlich eingeschränkte körperliche Funktionsfähigkeit. Aus diesem Grund sind die untersuchten Morbiditätsendpunkte als schwerwiegende Symptome der Erkrankung einzuschätzen.

Das Erreichen der Remission stellt prinzipiell das primäre Therapieziel für Patienten mit rheumatoider Arthritis dar. Die Wahrscheinlichkeit, eine Remission gemäß DAS28-ESR zu erreichen, war unter Sarilumab gegenüber Adalimumab statistisch signifikant erhöht. Das Ausmaß dieses Effekts entspricht einem erheblichen Zusatznutzen. Bei Anwendung der strengeren CDAI- und SDAI- sowie der booleschen ACR/EULAR-Kriterien für eine Remission zeigte sich ein numerischer Vorteil für Sarilumab gegenüber Adalimumab, die Behandlungsunterschiede erreichten jedoch keine Signifikanz.

Für diejenigen Patienten, die den Zustand der Remission nicht erreichen können, stellt auch gemäß aktuellster Empfehlungen das Erreichen einer niedrigen Krankheitsaktivität das primäre Therapieziel dar, insbesondere für Patienten mit einer lang andauernden, schweren RA, welche Bestandteil der hier dargestellten Studie MONARCH waren. Die Wahrscheinlichkeit, eine niedrige Krankheitsaktivität zu erreichen, war unter Sarilumab im Vergleich zu Adalimumab ebenfalls signifikant erhöht, und zwar sowohl bei Betrachtung des DAS28-ESR als auch des CDAI und SDAI. Das Ausmaß dieses Effektes für den Endpunkt „Niedrige Krankheitsaktivität“ entspricht in allen drei Fällen einem erheblichen Zusatznutzen.

Die Behandlung mit Sarilumab führte zu einem signifikant besseren Therapieansprechen als die Behandlung mit Adalimumab. Die Wahrscheinlichkeit für ein gutes bis moderates Ansprechen gemäß EULAR-Kriterien war unter Sarilumab im Vergleich zu Adalimumab signifikant erhöht. Das Ausmaß des Effektes entspricht einem geringen Zusatznutzen. Dabei zeigte sich vor allem ein deutlicher Vorteil von Sarilumab hinsichtlich des Anteils der Patienten mit gutem Ansprechen. Die Wahrscheinlichkeit, eine 50%ige oder 70%ige Verringerung der Krankheitsaktivität gemäß ACR-Kriterien zu erreichen, war unter Sarilumab im Vergleich zu Adalimumab ebenfalls signifikant erhöht. Das Ausmaß der Effekte entspricht jeweils einem beträchtlichen Zusatznutzen.

Die Behandlung mit Sarilumab führte des Weiteren zu einer signifikanten und klinisch relevanten Verbesserung des körperlichen Funktionsstatus gemessen mittels HAQ-DI. Das Ausmaß des Effektes für den Endpunkt „Körperlicher Funktionsstatus“ entspricht einem geringen Zusatznutzen.

Für alle weiteren Morbiditätsendpunkte – Gelenksymptome, Schmerzen, Morgensteifigkeit, Fatigue und den allgemeinen Gesundheitszustand – konnte kein Zusatznutzen belegt werden. Allerdings zeigten sich über alle Endpunkte hinweg konsistent numerische Vorteile zugunsten von Sarilumab oder gar statistisch signifikante Vorteile, für die jedoch die klinische Relevanz nicht belegt werden konnte.

Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bestätigen die Ergebnisse der 24-wöchigen doppelblinden Behandlungsphase hinsichtlich der Verbesserung der Morbidität.

In der Gesamtschau ergibt sich gemäß § 5 Abs. 7 AM-NutzenV für die Endpunktkategorie „Morbidität“ aufgrund der erreichten langfristigen Freiheit von schwerwiegenden Symptomen eine große Verbesserung des therapierelevanten Nutzens und damit ein **erheblicher Zusatznutzen** von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab.

***Endpunktkategorie Gesundheitsbezogene Lebensqualität***

Die Behandlung mit Sarilumab führte im Vergleich mit der zweckmäßigen Vergleichstherapie Adalimumab zu einer deutlichen Verbesserung der gesundheitsbezogenen Lebensqualität.

Die Auswertung des generischen Messinstruments SF-36 ergab tendenziell einen Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der körperlichen Summenskala (PCS). Die Wahrscheinlichkeit, eine Verbesserung der SF-36-PCS um mindestens 2,5, 5,1 und 7,2 Punkte zu erreichen, war unter Sarilumab numerisch erhöht. Die Behandlungsunterschiede waren jedoch statistisch nicht signifikant. Hinsichtlich der psychischen Summenskala (MCS) des SF-36 zeigte sich kein relevanter Unterschied zwischen den Behandlungsarmen. Insgesamt kann für die Verbesserung der gesundheitsbezogenen Lebensqualität – gemessen mit dem generischen Fragebogen SF-36 – kein Zusatznutzen abgeleitet werden.

Die Auswertung des spezifischen Fragebogens RAID, der speziell für Patienten mit rheumatoider Arthritis entwickelt und validiert wurde, ergab jedoch eine statistisch signifikant

stärkere Verbesserung der Lebensqualität der Patienten unter Sarilumab. Das Ausmaß dieses Effektes entspricht einem geringen Zusatznutzen.

Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bestätigen die Ergebnisse der 24-wöchigen doppelblinden Behandlungsphase hinsichtlich der Verbesserung der Lebensqualität.

In der Gesamtschau ergibt sich für die Endpunktkategorie „Gesundheitsbezogene Lebensqualität“ aufgrund der festgestellten relevanten Verbesserung der Lebensqualität ein **geringer Zusatznutzen** von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab.

***Endpunktkategorie Sicherheit/Verträglichkeit***

Zusammenfassend zeigte sich in der Studie MONARCH überwiegend ein vergleichbares Sicherheitsprofil von Sarilumab und der zweckmäßigen Vergleichstherapie Adalimumab.

In Bezug auf die Gesamtraten der unerwünschten Ereignisse, der schwerwiegenden unerwünschten Ereignisse sowie der unerwünschten Ereignisse, die zum dauerhaften Abbruch der Behandlung führen, ergaben sich keine statistisch signifikanten Vor- oder Nachteile.

Unter der Behandlung mit Sarilumab traten signifikant häufiger Leukopenien auf als unter Adalimumab, wobei es sich fast ausschließlich um Neutropenien handelte. Diese führten aber nicht zu einer erhöhten Zahl von Infektionen. Dieser Effekt würde formal einem geringeren Nutzen vom Ausmaß beträchtlich entsprechen. Bei Neutropenien handelt es sich jedoch um Veränderungen eines Laborparameters, die nicht per se symptomatisch und für den Patienten spürbar sind. Mit einer Ausnahme handelte es sich um nicht schwerwiegende Neutropenien.

In der Studie MONARCH zeigte sich kein Unterschied zwischen den Behandlungsarmen bezüglich der Häufigkeit von Infektionen. Schwerwiegende Infektionen traten sehr selten auf und waren nicht mit einer Neutropenie assoziiert. Somit stellt das erhöhte Neutropenie-Risiko unter Sarilumab keinen patientenrelevanten Nachteil dar und das Auftreten von Neutropenien ist in diesem Fall als Surrogatparameter ohne klinische Relevanz zu betrachten.

In der Studie MONARCH zeigte sich des Weiteren ein signifikanter Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich des Auftretens von Lipidwerterhöhungen. Dabei handelte es sich um reine Laborwertveränderungen, die nicht als schwerwiegend eingestuft wurden. Das Ausmaß des Behandlungseffekts würde formal einem geringen Zusatznutzen entsprechen.

Für alle übrigen unerwünschten Ereignisse von besonderem Interesse wurden keine signifikanten Vor- oder Nachteile für Sarilumab beobachtet.

Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bestätigen die Ergebnisse der 24-wöchigen doppelblinden Behandlungsphase hinsichtlich der Sicherheit und Verträglichkeit.

In der Gesamtschau ergibt sich für die Endpunktkategorie „Sicherheit/Verträglichkeit“ aufgrund der fehlenden klinischen Relevanz der beobachteten Unterschiede **kein Zusatznutzen und kein geringerer Nutzen** für Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab.

***Zusammenfassung der Aussagen zum Zusatznutzen***

Bezüglich der Endpunktkategorie Morbidität konnte ein erheblicher Zusatznutzen von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab abgeleitet werden. Bezüglich der Lebensqualität zeigte sich ein geringer Zusatznutzen von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab. Die Daten zu Sicherheit und Verträglichkeit belegen in der Gesamtschau keinen Zusatznutzen und keinen geringeren Nutzen von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab.

Die qualitative Ergebnissicherheit der vorliegenden Studie ist insgesamt als hoch einzuschätzen. Über alle Endpunkte hinweg zeigten sich konsistente Ergebnisse mit gleicher Effektrichtung. Somit kann in der Gesamtschau von einer hohen Aussagesicherheit und damit einem Hinweis auf einen Zusatznutzen ausgegangen werden.

In der Gesamtbetrachtung ergibt sich unter Berücksichtigung der Schwere und der chronischen Natur der Erkrankung für Sarilumab ein **Hinweis auf einen erheblichen Zusatznutzen**, da eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte große Verbesserung des therapierelevanten Nutzens erreicht wird, insbesondere eine langfristige Freiheit von schwerwiegenden Symptomen sowie eine relevante Verbesserung der gesundheitsbezogenen Lebensqualität.

## 4.2 Methodik

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

### 4.2.1 Fragestellung

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

– Patientenpopulation
– Intervention
– Vergleichstherapie
– Endpunkte
– Studientypen

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

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

Das vorliegende Modul 4A bezieht sich auf die Wirksamkeit und Sicherheit von Sarilumab (Kevzara®) zur Behandlung der mittelschweren bis schweren aktiven rheumatoiden Arthritis bei erwachsenen Patienten, die auf ein oder mehrere krankheitsmodifizierende antirheumatische Arzneimittel (DMARDs) unzureichend angesprochen oder diese nicht vertragen haben. Sarilumab wird gemäß Fachinformation in Kombination mit MTX verabreicht und kann als Monotherapie gegeben werden, wenn MTX nicht vertragen wird oder wenn eine Behandlung mit MTX ungeeignet ist [1]. Aus diesem Anwendungsgebiet ergeben sich gemäß Beratungsgespräch mit dem Gemeinsamen Bundesausschuss (G-BA) vier Teilanwendungsgebiete (siehe Tabelle 4-2) [2]. Die **Teilanwendungsgebiete 1 bis 4** werden gemeinsam in Modul 4A dargestellt[1].

[1] Anmerkung bezüglich Teilanwendungsgebiet 1: Es handelt sich dabei um Patienten, welche zwar formal von der Zulassung für Sarilumab umfasst sind, die aber im klinischen Alltag noch nicht mit einem Biologikum behandelt werden würden. Die EULAR-Leitlinie sieht für diese Patienten eine Therapie mit einem weiteren csDMARD vor [3]. Aus diesem Grund ist davon auszugehen, dass die Patienten in Teilanwendungsgebiet 1 im Regelfall nicht mit Sarilumab behandelt werden. Das Teilanwendungsgebiet 1 wird somit für die Nutzenbewertung von Sarilumab als nicht relevant erachtet.

Die **zweckmäßige Vergleichstherapie** (zVT) gemäß G-BA-Beratung ist ebenfalls in Tabelle 4-2 aufgeführt [2]. Für die Teilanwendungsgebiete 2 und 3 zieht Sanofi als zweckmäßige Vergleichstherapie **Adalimumab** zur Ableitung des Zusatznutzens heran. Für die Teilanwendungsgebiete 1 und 4 steht keine direkt vergleichende Evidenz zur Verfügung.

Tabelle 4-2: Teilanwendungsgebiete und zweckmäßige Vergleichstherapie

| | **Benennung der Teilanwendungs-gebiete im Dossier** | **Formulierung der Teilanwendungsgebiete gemäß G-BA-Beratung (2016-B-101)** | **Formulierung der zweckmäßigen Vergleichstherapie gemäß G-BA-Beratung (2016-B-101)** |
|---|---|---|---|
| 1) [1] | Patienten, die mit **einem csDMARD** vorbehandelt sind und **keine ungünstigen Prognosefaktoren[2]** aufweisen | Patienten, bei denen keine ungünstigen Prognosefaktoren[2] vorliegen und die unzureichend auf eine vorangegangene Behandlung mit einem krankheitsmodifizierenden Antirheumatikum (klassische DMARDs, inklusive Methotrexat) ansprachen | – alternative klassische DMARDs, sofern geeignet (z.B. MTX, Leflunomid) als Mono- oder Kombinationstherapie |
| 2) | Patienten, die mit **einem csDMARD** vorbehandelt sind und **ungünstige Prognosefaktoren**[2] aufweisen | Patienten mit ungünstigen Prognosefaktoren[2], die unzureichend auf eine vorangegangene Behandlung mit einem krankheitsmodifizierenden Antirheumatikum (klassische DMARDs, inklusive Methotrexat) ansprachen | – biotechnologisch hergestellte DMARDs (bDMARD) in Kombination mit MTX (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab)<br>– ggf. als Monotherapie unter Berücksichtigung des jeweiligen Zulassungsstatus bei MTX-Unverträglichkeit |
| 3) | Patienten, die mit **mehreren csDMARDs** vorbehandelt sind | Patienten, die unzureichend auf eine vorangegangene Behandlung mit mehreren krankheitsmodifizierenden Antirheumatika (klassische DMARDs, inklusive Methotrexat) ansprachen | – bDMARD in Kombination mit MTX (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab);<br>– ggf. als Monotherapie unter Berücksichtigung des jeweiligen Zulassungsstatus bei MTX-Unverträglichkeit |
| 4) | Patienten, die mit einem oder mehreren **bDMARDs** vorbehandelt sind | Patienten, die unzureichend auf eine vorangegangene Behandlung mit einem oder mehreren bDMARDs ansprachen | – Wechsel der bDMARD-Therapie (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab) in Kombination mit MTX;<br>– ggf. als Monotherapie unter Berücksichtigung des jeweiligen Zulassungsstatus bei MTX-Unverträglichkeit;<br>– oder bei Patienten mit schwerer rheumatoider Arthritis Rituximab unter Berücksichtigung des Zulassungsstatus;<br>– in Abhängigkeit von der Vortherapie |

[1] Anmerkung bezüglich Teilanwendungsgebiet 1: Es handelt sich dabei um Patienten, welche zwar formal von der Zulassung für Sarilumab umfasst sind, die aber im klinischen Alltag noch nicht mit einem Biologikum behandelt werden würden. Die EULAR-Leitlinie sieht für diese Patienten eine Therapie mit einem weiteren csDMARD vor [3]. Aus diesem Grund ist davon auszugehen, dass die Patienten in Teilanwendungsgebiet 1 im Regelfall nicht mit Sarilumab behandelt werden. Das Teilanwendungsgebiet 1 wird somit für die Nutzenbewertung von Sarilumab als nicht relevant erachtet.

[2] Ungünstige Prognosefaktoren wie z.B.: Nachweis von Autoantikörpern (z.B. Rheumafaktoren, hohe Spiegel von Antikörpern gegen citrullinierte Peptid-Antigene), hohe Krankheitsaktivität (nachgewiesen durch DAS bzw. DAS28-Bewertungssystem, geschwollene Gelenke, Parameter der Akute-Phase-Reaktion wie z.B. CRP, ESR), frühes Auftreten von Gelenkerosionen

Für die Nutzenbewertung von Sarilumab für die Teilanwendungsgebiete 2 und 3 steht die Studie MONARCH zur Verfügung (siehe Abschnitt 4.3.1). Da in der Studie jedoch ausschließlich die Monotherapie mit Sarilumab und Adalimumab untersucht wurde, liegt für diejenigen Patienten der Teilanwendungsgebiete 2 und 3, die für eine Kombinationstherapie mit MTX geeignet wären, keine Evidenz vor. Dementsprechend wird für die Ableitung des Zusatznutzens eine weitere Unterscheidung zwischen Patienten, die für MTX geeignet sind (Teilanwendungsgebiete 2a und 3a), und Patienten mit MTX-Intoleranz (Teilanwendungsgebiete 2b und 3b) vorgenommen (siehe Tabelle 4-3).

Tabelle 4-3: Untergliederung der Teilanwendungsgebiete 2 und 3 anhand der MTX-Eignung

| Benennung der Teilanwendungsgebiete | Zweckmäßige Vergleichstherapie |
|---|---|
| 2 a) Für **MTX geeignete** Patienten, die mit **einem csDMARD** vorbehandelt sind und **ungünstige Prognosefaktoren** aufweisen | – bDMARD in Kombination mit MTX (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab) |
| b) Patienten mit **MTX-Intoleranz**, die mit **einem csDMARD** vorbehandelt sind und **ungünstige Prognosefaktoren** aufweisen | – bDMARD als Monotherapie unter Berücksichtigung des jeweiligen Zulassungsstatus bei MTX-Unverträglichkeit (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Tocilizumab) |
| 3 a) Für **MTX geeignete** Patienten, die mit **mehreren csDMARDs** vorbehandelt sind | – bDMARD in Kombination mit MTX (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab) |
| b) Patienten mit **MTX-Intoleranz**, die mit **mehreren csDMARDs** vorbehandelt sind | – bDMARD als Monotherapie unter Berücksichtigung des jeweiligen Zulassungsstatus bei MTX-Unverträglichkeit (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Tocilizumab) |

Die **Intervention** Sarilumab ist ein monoklonaler Antikörper, der als Injektionslösung (Fertigpen oder Fertigspritze) alle zwei Wochen in einer Wirkstärke von 200 mg subkutan verabreicht wird [1]. Bei Patienten, bei denen unter der Behandlung mit Sarilumab eine Neutropenie, Thrombozytopenie oder Erhöhung der Leberenzymwerte auftritt, ist die Dosis auf 150 mg zu reduzieren. Gemäß Fachinformation wird Sarilumab in Kombination mit MTX verabreicht. Bei Patienten, bei denen eine MTX-Unverträglichkeit vorliegt oder eine

Fortsetzung der Therapie mit MTX ungeeignet erscheint, kann Sarilumab als Monotherapie gegeben werden.

Zur Beurteilung des Zusatznutzens werden **patientenrelevante Endpunkte** der Kategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität (HrQoL) und Sicherheit im Rahmen von randomisierten, kontrollierten Studien (**RCT**) untersucht.

Zusammenfassend leitet sich daraus folgende **Fragestellung** für Modul 4A ab:

Welchen Zusatznutzen hat Sarilumab als Monotherapie oder in Kombination mit MTX im Vergleich zur gewählten zweckmäßigen Vergleichstherapie bei der Behandlung der mittelschweren bis schweren aktiven rheumatoiden Arthritis bei erwachsenen Patienten, die auf ein oder mehrere DMARDs unzureichend angesprochen oder diese nicht vertragen haben, hinsichtlich der patientenrelevanten Endpunkte Mortalität, Morbidität, Lebensqualität und Sicherheit im Rahmen von RCT?

### 4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister erfolgen, während ein Kongressabstract allein in der Regel nicht für eine Studienbewertung ausreicht.

*Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.*

Tabelle 4-4: Ein- und Ausschlusskriterien

| | | **Einschlusskriterium (E)** | **Ausschlusskriterium (A)** |
|---|---|---|---|
| 1 | Patienten-population | Erwachsene Patienten mit mittelschwerer bis schwerer aktiver rheumatoider Arthritis, die auf ein oder mehrere DMARDs unzureichend angesprochen oder diese nicht vertragen haben | Die Studienpopulation ist nicht wie definiert. |
| 2 | Intervention | Sarilumab 200 mg s.c. q2w in Kombination mit MTX oder als Monotherapie | Die Intervention ist nicht wie definiert. |
| 3 | Vergleichs-therapie | csDMARDs[a] und/oder bDMARDs[b] in Kombination oder als Monotherapie in den jeweils für Deutschland zugelassenen Dosierungen und Anwendungsarten | Die Vergleichstherapie ist nicht wie definiert. |
| 4 | Endpunkte | Mindestens ein patientenrelevanter Endpunkt (Mortalität, Morbidität, Lebensqualität, Sicherheit) wird berichtet. | Kein patientenrelevanter Endpunkt wird berichtet. |
| 5 | Studientyp | Die Studie ist eine randomisierte, kontrollierte, doppelblinde Studie. | Die Studie ist keine randomisierte, kontrollierte, doppelblinde Studie. |
| 6 | Studiendauer | Die Studiendauer der Studie beträgt ≥ 24 Wochen. | Die Studiendauer ist nicht wie definiert. |
| 7 | Publikations-typ | Eine Vollpublikation oder der Studienbericht ist verfügbar. | Nur ein Review, Abstract oder Kongress-Poster ist verfügbar. |

[a] Methotrexat, Sulfasalazin, Leflunomid, Chloroquin, Hydroxychloroquin, Azathioprin, Ciclosporin, parenterales Gold
[b] Adalimumab, Etanercept, Certolizumab-Pegol, Golimumab, Abatacept, Tocilizumab, Rituximab

**Erläuterungen und Begründungen zu den Ein- und Ausschlusskriterien**

***Patientenpopulation (Ein-/Ausschlusskriterium 1):***

Es wurden Studien eingeschlossen, deren Patienten zur Patientenpopulation im zugelassenen Anwendungsgebiet von Sarilumab gehören [1]. Die identifizierten Studien wurden anhand der Vortherapie und Basischarakteristika der Patienten, so wie sie in den Einschlusskriterien der jeweiligen Studien definiert sind, den Teilanwendungsgebieten 1 bis 4 zugeordnet.

Teilanwendungsgebiet 1: Patienten, die mit einem csDMARD vorbehandelt sind und keine ungünstigen Prognosefaktoren aufweisen

Teilanwendungsgebiet 2: Patienten, die mit einem csDMARD vorbehandelt sind und ungünstige Prognosefaktoren aufweisen

Teilanwendungsgebiet 3: Patienten, die mit mehreren csDMARDs vorbehandelt sind

Teilanwendungsgebiet 4: Patienten, die mit einem oder mehreren bDMARDs vorbehandelt sind

***Intervention (Ein-/Ausschlusskriterium 2):***

Es wurden nur Studien eingeschlossen, in denen Sarilumab in der empfohlenen Dosierung sowie Art der Anwendung laut Fachinformation eingesetzt wurde [1]. Die empfohlene Dosierung beträgt 200 mg subkutan alle zwei Wochen. Bei Patienten, bei denen unter der Behandlung mit Sarilumab eine Neutropenie, Thrombozytopenie oder Erhöhung der Leberenzymwerte auftritt, ist die Dosis auf 150 mg zu reduzieren. Gemäß Fachinformation wird Sarilumab in Kombination mit MTX verabreicht. Bei Patienten, bei denen eine MTX-Unverträglichkeit vorliegt oder eine Fortsetzung der Therapie mit MTX ungeeignet erscheint, kann Sarilumab als Monotherapie gegeben werden.

***Vergleichstherapie (Ein-/Ausschlusskriterium 3):***

Die in den identifizierten Studien eingesetzte Vergleichstherapie musste je nach Patientenpopulation der vom G-BA festgelegten zweckmäßigen Vergleichstherapie der Teilanwendungsgebiete 1 bis 4 entsprechen (G-BA-Beratung 2016-B-101 vom 15.09.2016 [2]).

Teilanwendungsgebiet 1: alternative csDMARDs, sofern geeignet (z.B. MTX, Leflunomid) als Mono- oder Kombinationstherapie

Teilanwendungsgebiet 2: bDMARD in Kombination mit MTX (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab), ggf. als Monotherapie unter Berücksichtigung des jeweiligen Zulassungsstatus bei MTX-Unverträglichkeit

Teilanwendungsgebiet 3: bDMARD in Kombination mit MTX (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab); ggf. als Monotherapie unter Berücksichtigung des jeweiligen Zulassungsstatus bei MTX-Unverträglichkeit

Teilanwendungsgebiet 4: Wechsel der bDMARD-Therapie (Adalimumab oder Etanercept oder Certolizumab-Pegol oder Golimumab oder Abatacept oder Tocilizumab); in Kombination mit MTX und ggf. als Monotherapie unter Berücksichtigung des jeweiligen Zulassungsstatus bei MTX-Unverträglichkeit; oder bei Patienten mit schwerer rheumatoider Arthritis Rituximab unter Berücksichtigung des Zulassungsstatus

***Endpunkte (Ein-/Ausschlusskriterium 4):***

Es wurden Studien eingeschlossen, die mindestens einen patientenrelevanten Endpunkt berichteten, da die Nutzenbewertung auf Basis von patientenrelevanten Endpunkten erfolgt. Nähere Angaben zu patientenrelevanten Endpunkten in der Indikation rheumatoide Arthritis finden sich in Abschnitt 4.2.5.2.

***Studientyp (Ein-/Ausschlusskriterium 5):***

Es wurden ausschließlich randomisierte, kontrollierte, doppelblinde Studien eingeschlossen, da die Nutzenbewertung auf Basis der höchsten Evidenz erfolgt.

***Studiendauer (Ein-/Ausschlusskriterium 6):***

Die erforderliche Studiendauer sollte immer im Hinblick auf die Indikation, den zu untersuchenden Wirkstoff und die jeweilige Fragestellung festgelegt werden. Bei der rheumatoiden Arthritis handelt es sich um eine chronische Erkrankung, deren Therapie in der Regel eine Dauermedikation erfordert. Aus diesem Grund ist in der vorliegenden Indikation eine Studiendauer von weniger als 24 Wochen als unzureichend anzusehen.

*Aussagen der Bewertungsinstitutionen*

Auch das IQWiG hat für die aktuell laufende Nutzenbewertung der Biologika in der rheumatoiden Arthritis eine Mindeststudiendauer von sechs Monaten (24 Wochen) als adäquat erachtet und deshalb als Einschlusskriterium für die Evidenzrecherche festgelegt [4]. Das IQWiG beruft sich dabei auf die europäische Leitlinie zur Behandlung der rheumatoiden Arthritis von 2014 sowie auf die ICH[2]-Leitlinie für Studien zur Bewertung der Sicherheit von neuen Wirkstoffen in einer chronischen Indikation [5, 6].

In der Vergangenheit hat auch der G-BA für eine andere chronische immunvermittelte Erkrankung, die Plaque Psoriasis (Nutzenbewertungsverfahren Secukinumab/Cosentyx®) eine Mindeststudiendauer von sechs Monaten als adäquat festgelegt [7]. In den Beratungsgesprächen zur Nutzenbewertung von Secukinumab/Cosentyx® in den Indikationen Psoriasis-Arthritis und ankylosierende Spondylitis – zwei Erkrankungen des rheumatischen Formenkreises – hat der G-BA gemäß Angabe im Dossier des Herstellers ebenfalls eine minimale Studiendauer von 24 Wochen festgelegt (G-BA-Beratungen 2013-B-080 und 2013-B-081) [8, 9].

*Behandlungsleitlinien zur rheumatoiden Arthritis*

Die aktuelle europäische Leitlinie der EULAR (*European League Against Rheumatism*) macht Angaben dazu, wie schnell der Wirkungseintritt der DMARD-Therapien zu erwarten ist und wann eine unzureichende Therapie angepasst werden muss:

"*3. Monitoring should be frequent in active disease (every 1–3 months); if there is no improvement by at most 3 months after the start of treatment or the target has not been reached by 6 months, therapy should be adjusted*." [3].

Das bedeutet, Patienten, die nach sechsmonatiger Behandlung ihr Therapieziel nicht erreicht haben, profitieren nicht mehr davon, wenn die Behandlung unverändert fortgesetzt wird. Dies

[2] ICH: *International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use*

gilt gleichermaßen für csDMARDs und bDMARDs [3]. Es gibt somit keine medizinische Rationale dafür, Patienten, die in einer Studie nach sechs Monaten Behandlung kein ausreichendes Therapieansprechen zeigen, weiter auf dieser unzureichenden Therapie zu belassen. Eine Beurteilung des Therapieerfolgs nach sechs Monaten ist somit sachgerecht.

*Therapieansprechen bei den im Anwendungsgebiet zugelassenen Arzneimitteln*

Die Fachinformationen der meisten zVT-Arzneimittel machen Angaben dazu, in welchem Zeitraum ein Ansprechen auf die Therapie zu erwarten ist und wann bei ausbleibender Wirksamkeit die Therapie angepasst werden sollte.

Ein erstes Ansprechen auf die Therapie sollte in der Regel innerhalb der ersten drei Monate der Therapie sichtbar werden [10-18]. Die Fachinformationen von Adalimumab, Certolizumab pegol, Etanercept sowie Golimumab geben konkret vor, dass die Fortführung der Therapie sorgfältig zu überdenken ist, wenn ein Patient innerhalb von 12 Wochen keinen klinischen Nutzen der Therapie zeigt [14-16, 18]. Die Weiterbehandlung mit Rituximab sollte überdacht werden, wenn ein Patient innerhalb von 24 Wochen keinen klinischen Nutzen der Therapie zeigt [19]. Für Abatacept sieht der Hersteller bei ausbleibendem Ansprechen einen maximalen Behandlungszeitraum von 6 Monaten vor [20].

Ein primäres Therapieversagen würde sich also innerhalb der ersten Monate der Therapie zeigen, und zur Beurteilung der Wirksamkeit im Sinne eines Ansprechens auf die Therapie ist somit eine Studiendauer von sechs Monaten adäquat.

*Fazit zur Studiendauer*

- Das IQWiG hat für die aktuell laufende Nutzenbewertung der Biologika in der rheumatoiden Arthritis eine Mindeststudiendauer von sechs Monaten (24 Wochen) festgelegt [4].

- Die EULAR-Leitlinie gibt an, dass Patienten, die nach sechsmonatiger Behandlung mit einem DMARD ihr Therapieziel nicht erreicht haben, nicht mehr davon profitieren, wenn die Behandlung unverändert fortgesetzt wird. Eine Therapieanpassung sollte also immer dann erfolgen, wenn nach sechs Monaten Behandlung das Therapieziel nicht erreicht wurde [3].

- Die Fachinformationen diverser DMARDs weisen darauf hin, dass in der Indikation rheumatoide Arthritis ein klinisches Ansprechen üblicherweise innerhalb von 12 Behandlungswochen erreicht wird. Die Fortsetzung der Therapie sollte bei einem Patienten, der innerhalb dieser Zeitspanne nicht anspricht, sorgfältig überdacht werden.

Somit ist für den Vergleich der Wirksamkeit und Sicherheit von Sarilumab mit der zweckmäßigen Vergleichstherapie eine Studiendauer von 24 Wochen adäquat.

***Publikationstyp (Ein-/Ausschlusskriterium 7):***

Es wurden lediglich Studien berücksichtigt, für die eine Vollpublikation oder ein Studienbericht vorlag, um die Kriterien des CONSORT Statements zu erfüllen und eine Einschätzung der Studienergebnisse zu ermöglichen [21, 22].

### 4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

#### 4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

#### 4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

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

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

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

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

Die bibliographische Literaturrecherche wurde am 30.06.2017 über die Oberfläche Ovid in den Literaturdatenbanken MEDLINE und EMBASE durchgeführt. Weiterhin wurde eine Recherche in den Cochrane-Datenbanken durchgeführt.

Für jede Datenbank wurde eine adaptierte Suchstrategie formuliert, die in Anhang 4-A detailliert beschrieben ist. Für die Suche wurden Synonyme des zu bewertenden Arzneimittel Sarilumab als Freitext- oder ggf. als MeSH-Begriffe verwendet. Eine Einschränkung auf die Indikation wurde nicht vorgenommen, um eine möglichst breite Suche zu gewährleisten.

Eine Suche für RCT für einen indirekten Vergleich, nach nicht randomisierten vergleichenden Studien oder nach weiteren Untersuchungen wurde nicht durchgeführt.

#### 4.2.3.3 Suche in Studienregistern

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

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

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

Die Suche soll mindestens in den Studienregistern clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), Klinische Prüfungen PharmNet.Bund (http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search

Portal, Suchportal der WHO: http://apps.who.int/trialsearch/) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

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

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

Neben der bibliographischen Literaturrecherche wurde vom 26.06. bis zum 30.06.2017 eine Suche nach Studien in den Studienregistern www.clinicaltrials.gov, im International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, http://apps.who.int/trialsearch/AdvSearch.aspx), dem EU Clinical Trials Register (http://www.clinicaltrialsregister.eu/ctr-search) und im Register des PharmNet.Bund (www.pharmnet-bund.de) durchgeführt. Die verwendeten Suchbegriffe bezogen sich ausschließlich auf Synonyme des zu bewertenden Arzneimittels Sarilumab. Die detaillierte Suchstrategie ist für jedes Studienregister getrennt in Anhang 4-B dargestellt.

Eine Suche in Studienregistern nach Studien für einen indirekten Vergleich, nach nicht randomisierten vergleichenden Studien oder nach weiteren Untersuchungen wurde nicht durchgeführt.

#### 4.2.3.4 Selektion relevanter Studien

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

Die Auswahl einzuschließender Studien aus den Treffern in den medizinischen Datenbanken und den Studienregistern wurde nach Entfernung von Dubletten und unter Anwendung der festgelegten Ein- und Ausschlusskriterien (vgl. Abschnitt 4.2.2) durchgeführt.

Die einzelnen Schritte zur Selektion der Studien wurden unabhängig von zwei Personen durchgeführt. Sofern sich bezüglich des Ein- und Ausschlusses von Treffern abweichende Beurteilungen ergaben, wurden diese im abschließenden Konsensverfahren durch die beiden am Verfahren beteiligten Personen geklärt.

### 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.*

Von der oben angegebenen Methodik wurde nicht abgewichen. Die Verzerrungsaspekte der Ergebnisse wurden für jede eingeschlossene Studie beschrieben. Die Bewertung der einzelnen Aspekte wurde für jeden Endpunkt im Bewertungsbogen in Anhang 4-F dokumentiert.

### 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)[3]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[4] bzw. STROBE-Statements[5] folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

*Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.*

Das Design und die Methodik der eingeschlossenen Studien wurden ausführlich in Abschnitt 4.3.1.2.1 und Anhang 4-E dokumentiert. Zur Bewertung des Zusatznutzens von Sarilumab wurden ausschließlich randomisierte kontrollierte Studien herangezogen. Somit erfolgte die Darstellung der Studien nach den Anforderungen des CONSORT-Statements (Items 2b bis 14 und CONSORT-Flow-Chart) [21, 22].

[3] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.
[4] 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.
[5] 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.

#### 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.*

**Patientencharakteristika**

In der für die Nutzenbewertung von Sarilumab relevanten Studie MONARCH wurden demographische und andere Patientencharakteristika (Basischarakteristika der Erkrankung) erhoben. Im vorliegenden Dossier werden folgende Patientencharakteristika dargestellt:

***Demographische Daten:***

- Alter
- Geschlecht
- Rasse
- BMI
- Region

***Basischarakteristika der Erkrankung:***

- Krankheitsdauer (Zeit seit Diagnose)
- Funktionelle Klasse der rheumatoiden Arthritis gemäß ACR-Kriterien

- Anteil der Patienten mit Rheumafaktoren
- Anteil der Patienten mit anti-CCP-Antikörpern
- Anzahl der druckschmerzhaften Gelenke
- Anzahl der geschwollenen Gelenke
- Funktionsstatus (HAQ-DI)
- CRP
- ESR
- DAS28-ESR
- CDAI

Die Darstellung erfolgt jeweils für die randomisierte Population (für die Gesamtpopulation der Studie sowie die drei im Dossier dargestellten Teilpopulationen).

**Endpunkte**

Für die Nutzenbewertung von Sarilumab wurden Endpunkte der Kategorien Morbidität, gesundheitsbezogene Lebensqualität und Sicherheit untersucht. Ergebnisse zum Endpunkt Mortalität liegen nicht vor. Die folgende Tabelle 4-5 zeigt die patientenrelevanten Endpunkte sowie weitere, im Dossier ergänzend dargestellte Endpunkte, die in der Studie MONARCH erhoben wurden, sowie ihre Zuordnung zu den Endpunktkategorien. Im Anschluss wird die Patientenrelevanz der einzelnen Endpunkte beschrieben und die Validität der verwendeten Messinstrumente diskutiert.

Tabelle 4-5: Patientenrelevante und ergänzend dargestellte Endpunkte der Studie MONARCH

| Kategorie | Endpunkt | Operationalisierung |
|---|---|---|
| **Morbidität** | Änderung der Krankheitsaktivität | Änderung des DAS28-ESR |
| | | Änderung des CDAI |
| | | Änderung des SDAI |
| | Remission | DAS28-ESR < 2,6 |
| | | CDAI ≤ 2,8 |
| | | SDAI ≤ 3,3 |
| | | Boolesche ACR/EULAR-Remission |
| | Niedrige Krankheitsaktivität | DAS28-ESR < 3,2 |
| | | CDAI ≤ 10 |
| | | SDAI ≤ 11 |
| | Therapieansprechen | Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR |
| | | Ansprechen gemäß ACR-Kriterien (ACR20/50/70) |
| | Einzelkomponenten der Scores zur Krankheitsaktivität | Gelenksymptome: Druckschmerzhafte Gelenke (68TJC + 28TJC) |
| | | Gelenksymptome: Geschwollene Gelenke (66SJC + 28SJC) |
| | | Schmerzen |
| | | Globale Einschätzung des Arztes (PhGA) |
| | | Globale Einschätzung des Patienten (PtGA) |
| | | Körperlicher Funktionsstatus – HAQ-DI |
| | | Laborwerte: CRP + ESR |
| | Morgensteifigkeit | Morgensteifigkeit-VAS |
| | Fatigue | FACIT-Fatigue |
| | Änderung des allgemeinen Gesundheitszustands | EQ-5D-VAS |
| **Lebensqualität** | Änderung der Lebensqualität - SF-36 | |
| | Änderung der Lebensqualität - RAID | |
| **Sicherheit/ Verträglichkeit** | Unerwünschte Ereignisse (UE)<br>Schwerwiegende unerwünschte Ereignisse (SUE)<br>Therapieabbrüche aufgrund unerwünschter Ereignisse<br>Unerwünschte Ereignisse mit Todesfolge<br>Unerwünschte Ereignisse von besonderem Interesse (UEBI) | |

***Endpunkte zur Morbidität***

*Messinstrumente zur Bestimmung der Krankheitsaktivität*

Zur Bestimmung der Krankheitsaktivität bei Patienten mit rheumatoider Arthritis stehen verschiedene Messinstrumente zur Verfügung, die validiert und sowohl im klinischen Alltag als auch in klinischen Studien gebräuchlich sind. Aufgrund der heterogenen klinischen Manifestation der Erkrankung ist es nicht möglich, die Krankheitsaktivität durch einen einzelnen Parameter valide abzubilden. Aus diesem Grund wurde eine Reihe von Messinstrumenten entwickelt, die verschiedene Einzelkomponenten jeweils zu einem Gesamtscore kombinieren. Alle Scores beruhen auf der Quantifizierung der druckschmerzhaften (TJC, *Tender Joint Count*) und geschwollenen (SJC, *Swollen Joint Count*) Gelenke sowie einer Einschätzung des Patienten zur Krankheitsaktivität (PtGA, *Patient Global Assessment*). Zusätzlich berücksichtigen manche Messinstrumente Laborwerte, wie die CRP (C-reaktives Protein)-Konzentration im Blut oder die Erythrozytensedimentationsrate (ESR), sowie die Einschätzung des Arztes zur Krankheitsaktivität (PhGA, *Physicians Global Assessment*).

Die folgende Tabelle 4-6 fasst zusammen, wie die Einzelkomponenten zu den Messinstrumenten DAS28-ESR (*Disease Activity Score with 28 Joint Counts*), CDAI (*Clinical Disease Activity Index*), SDAI (*Simplified Disease Acitvity Index*) und ACR (*American College of Rheumatology*)-Kriterien kombiniert werden.

Tabelle 4-6: Einzelkomponenten der Scores zur Bestimmung der Krankheitsaktivität

| Einzelkomponenten | DAS28-ESR | CDAI | SDAI | ACR |
|---|---|---|---|---|
| Gelenksymptome: | | | | |
| - Druckschmerzhafte Gelenke (68TJC) | | | | X |
| - Druckschmerzhafte Gelenke (28TJC) | X | X | X | |
| - Geschwollene Gelenke (66SJC) | | | | X |
| - Geschwollene Gelenke (28SJC) | X | X | X | |
| Schmerzen | | | | X |
| Globale Einschätzung des Arztes (PhGA) | | X | X | X |
| Globale Einschätzung des Patienten (PtGA) | X | X | X | X |
| Körperlicher Funktionsstatus - HAQ-DI | | | | X |
| Laborparameter: | | | | |
| - CRP | | | X | X |
| - ESR | X | | | |

Im Folgenden werden die Patientenrelevanz und Validität der Einzelkomponenten und der verwendeten zusammengesetzten Messinstrumente dargestellt.

Einzelkomponenten der Messinstrumente: Gelenksymptome

Bezüglich der Gelenksymptomatik werden zum einen druckschmerzhafte (TJC) und zum anderen geschwollene (SJC) Gelenke berücksichtigt. Für die Auswertung der TJC werden nach ACR 68[6] Gelenke auf Druckschmerzhaftigkeit untersucht. Für die Auswertung der SJC werden entsprechend 66 Gelenke auf Schwellungen untersucht. Dabei handelt es sich um dieselben Gelenke, die zur Bestimmung der TJC begutachtet werden, mit Ausnahme der Hüftgelenke, für die eine Schwellung schwieriger zu beurteilen ist. Als vereinfachte Variante können TJC und SJC auch basierend auf 28 Gelenken[7] bestimmt werden (siehe DAS28-ESR, CDAI, SDAI). Dabei handelt es sich um die Gelenke, die bei Patienten mit rheumatoider Arthritis am häufigsten betroffen sind. Die Gelenkzählung basierend auf 28 Gelenken wurde gegenüber der Zählung anhand von 68 Gelenken validiert [23]. Druckschmerzhaftigkeit und Schwellungen der Gelenke sind als direkt patientenrelevante Parameter einzuschätzen.

Einzelkomponenten der Messinstrumente: Schmerzen

Schmerzen sind unabhängig vom Krankheitsbild und der Schmerzursache generell patientenrelevant, da sie direkt das Wohlbefinden und die Lebensqualität der Patienten beeinträchtigen. Die Patientenrelevanz für Patienten mit rheumatoider Arthritis wurde vom G-BA im Rahmen des Beratungsgesprächs 2015-B-178 bestätigt [24]. Das Ausmaß der Schmerzen wurde in der Studie MONARCH mittels einer 100 mm-VAS erhoben.

Als klinisch relevante Schmerzreduktion wurde für Patienten mit rheumatoider Arthritis eine Verbesserung um 20,4 auf einer 100 mm-VAS definiert [25]. Somit wurde im vorliegenden Dossier ein Wert von 20,4 als MCID (*Minimal Clinically Important Difference*) für eine Responderanalyse gewählt.

Einzelkomponenten der Messinstrumente: Globale Einschätzung des Arztes (PhGA)

Die Einschätzung des Arztes zur Krankheitsaktivität der Patienten wird auf einer visuellen Analogskala (VAS) vorgenommen, auf der der Arzt einen Wert zwischen 0 (keine Krankheitsaktivität) und 100 (höchstmögliche Krankheitsaktivität) ankreuzt. Die Einschätzung des Arztes ist als nicht direkt patientenrelevant zu bewerten. Die PhGA wird im vorliegenden Dossier lediglich aus Transparenzgründen dargestellt, da der Parameter in die Berechnung von CDAI und SDAI eingeht, sowie eines der ACR-Kriterien darstellt.

[6] Kiefergelenke (n=2), Sternoklavikulargelenke (Brustbein-Schlüsselbein-Gelenke, n=2), Akromioklavikulargelenke (Schultereckgelenke, n=2), Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke der Hände (Fingermittelgelenke, n=8), distale Interphalangealgelenke der Hände (Fingerendgelenke, n=8), Hüftgelenke (n=2), Kniegelenke (n=2), obere Sprungelenke (n=2), untere Sprungelenke (n=2), Metatarsophalangealgelenke (Zehengrundgelenke, n=10), Interphalangealgelenke der großen Zehen (n=2), proximale/distale Interphalangealgelenke der Zehen (Zehenmittel- und -endgelenke, n=8)

[7] Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke der Hände (Fingermittelgelenke, n=8), Kniegelenke (n=2)

Einzelkomponenten der Messinstrumente: Globale Einschätzung des Patienten (PtGA)

Für die Selbsteinschätzung der Patienten zu ihrer Krankheitsaktivität wird ebenfalls eine VAS verwendet, auf der die Patienten einen Wert zwischen 0 (keine Krankheitsaktivität) und 100 (höchstmögliche Krankheitsaktivität) ankreuzen. Die PtGA gibt direkt wieder, wie die Patienten die Schwere ihrer Erkrankung empfinden, und ist somit patientenrelevant. Für die PtGA wurde in einer Studie an Patienten mit rheumatoider Arthritis eine Verbesserung um 18,4 Punkte als minimale klinisch relevante Verbesserung definiert [25]. Entsprechend wurde im vorliegenden Dossier ein Wert von 18,4 als MCID für eine Responderanalyse gewählt.

Einzelkomponenten der Messinstrumente: Körperlicher Funktionsstatus - HAQ-DI

Funktionelle Einschränkungen haben einen großen Anteil an der Beeinträchtigung der Lebensqualität von Patienten mit rheumatoider Arthritis. Somit hat die Verbesserung der körperlichen Funktionen eine große Patientenrelevanz.

Das am häufigsten verwendete und am besten validierte Messinstrument zur Beurteilung der funktionellen Beeinträchtigung bei Patienten mit rheumatoider Arthritis ist der HAQ-DI (*Health Assessment Questionnaire Disability Index*). Der HAQ-DI ist ein generisches Selbstbeurteilungsinstrument zur Erfassung von Bewegungs- und Funktionseinschränkungen, welches ursprünglich für Patienten mit rheumatoider Arthritis entwickelt und validiert wurde, aber mittlerweile auch in vielen anderen Indikationen Anwendung findet [26]. Im Beratungsgespräch 2015-B-178 hat der G-BA bestätigt, dass der HAQ-DI als valides Messinstrument akzeptiert wird [24].

Der Fragebogen umfasst acht Kategorien, die die Schwierigkeiten der Patienten bei Aktivitäten des täglichen Lebens erfassen: Anziehen, Aufstehen, Essen, Gehen, Körperpflege, Gegenstände reichen, Greifen und allgemeine tägliche Aktivitäten. Zusätzlich wird abgefragt, ob Hilfsmittel oder Hilfe von anderen Personen benötigt werden. Als Gesamtscore ergibt sich für den HAQ-DI ein Wert zwischen 0 und 3, wobei höhere Werte eine stärkere körperliche Beeinträchtigung anzeigen.

Es wurde gezeigt, dass der HAQ-DI ein wichtiger Prädiktor für die Mortalität von Patienten mit rheumatoider Arthritis sowie für den Langzeitverlauf der Erkrankung ist [27]. Der HAQ-DI zeigt eine gute Korrelation mit klinischen Parametern und Labormarkern der Erkrankung [26].

Eine Verbesserung um mindestens 0,22 Punkte wird standardmäßig als klinisch relevante Verbesserung interpretiert [28, 29]. Auch das IQWiG hat in der Vergangenheit eine Verbesserung des HAQ-DI um mindestens 0,22 Punkte als adäquate individuelle klinisch relevante Verbesserung bewertet [30]. Eine neuere Studie definiert eine Verbesserung um 0,375 Punkte als klinisch relevant [25]. Im Rahmen des vorliegenden Dossiers wurden Responderanalysen mit beiden MCIDs durchgeführt.

Einzelkomponenten der Messinstrumente: Laborparameter

Die Laborparameter CRP und ESR, die die Entzündungsaktivität widerspiegeln, werden im vorliegenden Dossier lediglich aus Transparenzgründen dargestellt, da die Parameter in die

Berechnung des SDAI und des DAS28-ESR eingehen bzw. eines der ACR-Kriterien darstellen. Die Veränderung dieser Laborparameter stellt zwar im klinischen Alltag ein relevantes Kriterium für die Beurteilung der Krankheitsschwere dar, jedoch haben die Laborwerte nach gängiger Interpretation des G-BA als Surrogatparameter keine direkte Patientenrelevanz.

DAS28-ESR

Der DAS28 (*Disease Activity Score with 28 Joint Counts*) gilt als Goldstandard zur Bestimmung der Krankheitsaktivität und wird von den Fachgesellschaften ACR (*American College of Rheumatology*) und EULAR (*European League Against Rheumatism*) zum Einsatz in klinischen Studien empfohlen [31]. Im Beratungsgespräch 2015-B-178 hat der G-BA bestätigt, dass es sich beim DAS28 um das Standardmessinstrument zur Bestimmung der Krankheitsaktivität handelt [24].

Der DAS28 beruht auf der Anzahl der geschwollenen und druckschmerzhaften Gelenke aus einem vordefinierten Satz von 28 Gelenken (28SJC und 28TJC) sowie der Erythrozytensedimentationsrate und der Patientenselbsteinschätzung [32, 33]. Mit Hilfe einer Formel wird ein Wert zwischen 0 und 9,4 berechnet, wobei höhere Werte eine höhere Krankheitsaktivität widerspiegeln. Der DAS28 stellt ein valides Messinstrument dar, um die Krankheitsaktivität und das Ansprechen auf eine Therapie zu ermitteln (siehe unten) [32, 34-36]. Eine alternative Variante des DAS28 verwendet die Serumkonzentration des CRP anstelle der ESR. Der DAS28-CRP zeigt eine starke lineare Korrelation mit dem DAS28-ESR, allerdings ist diese Korrelation bei niedrigen DAS28-Werten schwächer ausgeprägt [34]. Generell führt die Verwendung des CRP-Wertes zu etwas niedrigeren Werten des DAS28 als bei Verwendung der ESR, so dass der DAS28-CRP in Verbindung mit den gültigen Grenzwerten generell zu einer leichten Unterschätzung der Krankheitsaktivität führt [37]. Aus diesem Grund wurde im vorliegenden Dossier der DAS28-ESR zur Bestimmung von Remission, niedriger Krankheitsaktivität und Ansprechen gemäß EULAR-Kriterien herangezogen (siehe unten). Auf die zusätzliche Darstellung des DAS28-CRP wurde im vorliegenden Dossier entsprechend verzichtet.

CDAI

Ein weiteres etabliertes und validiertes Messinstrument zur Bestimmung der Krankheitsaktivität ist der CDAI (*Clinical Disease Activity Index*) [34]. Der CDAI beruht wie der DAS28 auf den Werten für 28TJC, 28SJC und der Patientenselbsteinschätzung [38]. Zusätzlich bezieht der CDAI noch die Arzteinschätzung (PhGA) mit ein. Daraus ergibt sich ein Score, der durch einfache Summierung der Einzelwerte Werte zwischen 0 und 76 annehmen kann. Der CDAI zeigt in Validierungsstudien eine starke Korrelation mit dem DAS28 und dem SDAI (siehe unten) [39]. Auch der Funktionsstatus, erhoben mittels HAQ-DI, sowie das Fortschreiten der Gelenkzerstörungen korrelieren mit dem CDAI [39]. Der CDAI kann ebenso wie der DAS28 zur Bestimmung der Krankheitsaktivität und des Ansprechens auf eine Therapie genutzt werden und wird ebenfalls von EULAR und ACR empfohlen, vor allem wenn kein CRP-Wert verfügbar ist [31, 34, 38, 40, 41]. Darüberhinaus ist die Darstellung des CDAI zusätzlich zu den anderen Messinstrumenten, die

Laborparameter berücksichtigen, medizinisch und aus Patientensicht hoch relevant, da die starke Erniedrigung der Entzündungsparameter durch die IL-6R-Blockade, die Beurteilung der Wirksamkeit mittel DAS28 oder SDAI erschwert. Die Validität eines gemessenen Behandlungseffekts auf die Krankheitsaktivität kann somit durch konsistente Ergebnisse über alle Messinstrumente hinweg belegt werden.

SDAI

Ein weiteres gebräuchliches und validiertes Messinstrument zur Bestimmung der Krankheitsaktivität ist der SDAI (*Simplified Disease Acitvity Index*) [34]. Der SDAI berücksichtigt die gleichen Parameter wie der CDAI, bezieht aber zusätzlich noch den CRP-Wert mit ein [38]. Es ergibt sich ein Score, der durch einfache Summierung der Einzelwerte Werte zwischen 0 und 86 annehmen kann. Validierungsstudien haben gezeigt, dass der SDAI signifikant mit dem DAS28 und dem Funktionsstatus von Patienten mit rheumatoider Arthritis, erhoben mittels HAQ-DI, assoziiert ist [40]. Eine Verringerung des SDAI korreliert ebenfalls mit einer Verringerung des Fortschreitens der Gelenkzerstörung [40]. Der SDAI kann ebenso wie der DAS28 zur Bestimmung der Krankheitsaktivität und des Ansprechens auf eine Therapie genutzt werden und wird ebenfalls von EULAR und ACR empfohlen [31, 34, 38, 40, 41].

*Remission und niedrige Krankheitsaktivität*

Das Ziel der Therapie der rheumatoiden Arthritis ist eine möglichst weitgehende Verringerung der Krankheitsaktivität. Den Zustand der weitestgehenden Symptomfreiheit bezeichnet man als Remission. Bei Patienten, die den Zustand der Remission nicht erreichen können, gilt eine niedrige Krankheitsaktivität als Therapieziel [3]. Sowohl das Erreichen der Remission als auch das Erreichen einer niedrigen Krankheitsaktivität stellen patientenrelevante Endpunkte dar. Es handelt sich um die Hauptkriterien zur Beurteilung des Therapieerfolgs. Wenn Patienten mit ihrer bestehenden Therapie weder eine Remission noch eine niedrige Krankheitsaktivität erreichen, ist eine Anpassung bzw. Intensivierung der Therapie erforderlich.

Zur Beurteilung des Erreichens einer Remission bzw. niedrigen Krankheitsaktivität existieren verschiedene spezifische Grenzwerte für die oben genannten Messinstrumente, die im klinischen Alltag und im Rahmen klinischer Studien Anwendung finden (siehe Tabelle 4-7). Auch im Rahmen der Studie MONARCH wurden unterschiedliche Operationalisierungen für beide Endpunkte verwendet. Diese sind im Folgenden erläutert.

Tabelle 4-7: Bestimmung der Krankheitsaktivität mittels DAS28, CDAI und SDAI [34]

| | | DAS28-ESR | CDAI | SDAI |
|---|---|---|---|---|
| **Krankheitsaktivität:** | Remission | < 2,6 | ≤ 2,8 | ≤ 3,3 |
| | niedrig | < 3,2 | ≤ 10 | ≤ 11 |
| | moderat | ≥ 3,2 und ≤ 5,1 | > 10 und ≤ 22 | > 11 und ≤ 26 |
| | schwer | > 5,1 | > 22 | > 26 |

DAS28: *Disease Activity Score with 28 Joint Counts*; ESR: Erythrozytensedimentationsrate; SDAI: *Simplified Disease Acitvity Index*; CDAI: *Clinical Disease Activity Index*

DAS28-Remission und niedrige Krankheitsaktivität

Ein DAS28-Wert < 2,6 gilt als Remission (siehe Tabelle 4-7). Dieser Schwellenwert wurde anhand der Remissionskriterien der *American Rheumatism Association* (ARA) validiert [42, 43]. Ein weiterer klinisch relevanter Schwellenwert ist der DAS28 < 3,2. Dieser wird als Kriterium für das Erreichen einer niedrigen Krankheitsaktivität verwendet. Auch die EULAR-Leitlinie nennt den DAS28 < 3,2 als eine von drei möglichen Definitionen einer niedrigen Krankheitsaktivität [3].

CDAI-Remission und niedrige Krankheitsaktivität

Ein CDAI ≤ 2,8 definiert den Zustand der Remission, während ein CDAI ≤ 10 als Grenzwert für eine niedrige Krankheitsaktivität gilt (siehe Tabelle 4-7) [38]. Die Schwellenwerte für die verschiedenen Krankheitslevel wurden vergleichend für CDAI, SDAI (s. unten) und DAS28 validiert. Dabei zeigte sich eine gute Übereinstimmung zwischen den drei Messinstrumenten und ihren jeweiligen Schwellenwerten [38].

SDAI-Remission und niedrige Krankheitsaktivität

Ein SDAI ≤ 3,3 wurde als Grenzwert für Remission festgelegt (siehe Tabelle 4-7) [44]. Für eine niedrige Krankheitsaktivität wurde ein Erreichen eines SDAI ≤ 11 definiert [44]. Die Schwellenwerte für die verschiedenen Krankheitslevel wurden vergleichend für CDAI, SDAI und DAS28 validiert. Dabei zeigte sich eine gute Übereinstimmung zwischen den drei Messinstrumenten und ihren jeweiligen Schwellenwerten [38]. Die Übereinstimmung zwischen CDAI und SDAI war dabei besonders hoch.

ACR/EULAR-Remission

2011 wurde von den Fachgesellschaften ACR und EULAR eine neue Definition für die Remission der rheumatoiden Arthritis veröffentlicht [41]. Dabei handelt es sich im Gegensatz zu den oben aufgeführten Scores um eine boolesche Definition („*Boolean Remission*“), die vier Kriterien betrachtet: die Anzahl der druckschmerzhaften und die Anzahl der geschwollenen Gelenke (jeweils basierend auf 28 Gelenken), die Patientenselbsteinschätzung auf einer VAS (0 bis 10 cm) sowie den CRP-Wert (in mg/dL). Patienten in Remission nach den booleschen ACR/EULAR-Kriterien müssen für jedes dieser Kriterien einen Wert von ≤ 1 aufweisen. Damit stellt die ACR/EULAR-Remission eine wesentlich strengere Definition als die oben aufgeführten Kriterien dar. Allerdings konnte in verschiedenen Studien gezeigt

werden, dass mit Ausnahme des DAS28 alle Remissionskriterien vergleichbar gut mit der Langzeitprognose bezüglich irreversibler Gelenkveränderungen korrelieren [41, 45]. Alternativ zur booleschen ACR/EULAR-Remission empfehlen ACR und EULAR die Verwendung des SDAI als Index-basiertes Remissionskriterium (siehe oben).

Kritisiert wurde, dass die boolesche ACR/EULAR-Remission maßgeblich von der Patientenselbsteinschätzung (≤ 1 cm auf der VAS) beeinflusst wird, was bei multimorbiden Patienten zu einer Verzerrung führen kann [46]. So können Patienten fälschlicherweise als nicht in Remission befindlich klassifiziert werden, wenn sie die PtGA aus Gründen, die nicht direkt mit ihrer rheumatoiden Arthritis zusammenhängen, mit > 1 bewerten. Dies könnte bei manchen Patienten zu einer Übertherapierung führen, wenn die boolesche ACR/EULAR-Remission als Entscheidungskriterium für eine Therapieanpassung herangezogen wird. Die oben aufgeführten Scores hingegen sind weniger anfällig für eine derartige Verzerrung. Aus diesem Grund ist es sinnvoll, im Rahmen klinischer Studien verschiedene Definitionen für Remission zu betrachten.

*Therapieansprechen*

Neben dem Erreichen einer möglichst geringen Krankheitsaktivität (Remission oder niedrige Krankheitsaktivität) stellt eine klinisch relevante Verringerung der Krankheitsaktivität ebenfalls einen patientenrelevanten Endpunkt in klinischen Studien dar. In der Studie MONARCH wurden parallel die Kriterien der EULAR und des ACR angewandt, um ein klinisch relevantes Ansprechen darzustellen.

Ansprechen gemäß EULAR-Kriterien

Die ursprünglichen Kriterien der EULAR für ein Therapieansprechen beruhen auf der Bestimmung der Krankheitsaktivität mittels DAS und berücksichtigen sowohl die durch eine Therapie erzielte numerische Verringerung des DAS als auch das erreichte Niveau der Krankheitsaktivität [35]. Die EULAR-Kriterien dienen einer Unterscheidung von Patienten mit gutem, moderatem oder fehlendem Ansprechen. Mit der Einführung des heute standardmäßig verwendeten DAS28 wurden die Kriterien für das Therapieansprechen an den modifizierten Score angepasst (siehe Tabelle 4-8) [36]. Die Unterscheidung anhand der erreichten Krankheitsaktivität beruht auf der üblichen Definition einer niedrigen, moderaten oder hohen Krankheitsaktivität gemäß DAS28 [42, 43]. Die Schwellenwerte für relevante Veränderungen des DAS28 von 0,6 und 1,2 beruhen auf Analysen zum Standardmessfehler des DAS [35].

Tabelle 4-8: Kriterien für ein Therapieansprechen gemäß EULAR [35]

<table>
<tr><th>DAS28</th><th colspan="3">DAS28-Senkung (Verbesserung)</th></tr>
<tr><th></th><th>> 1,2</th><th>> 0,6 und ≤ 1,2</th><th>≤ 0,6</th></tr>
<tr><td>≤ 3,2</td><td>gutes Ansprechen</td><td>moderates</td><td rowspan="3">kein Ansprechen</td></tr>
<tr><td>> 3,2 und ≤ 5,1</td><td rowspan="2"></td><td>Ansprechen</td></tr>
<tr><td>> 5,1</td><td></td></tr>
<tr><td colspan="4">DAS28: Disease Activity Score with 28 Joint Counts</td></tr>
</table>

Die EULAR-Kriterien, basierend auf dem DAS und dem DAS28, wurden im Vergleich zu den ACR-Kriterien (siehe unten) und zum Funktionsstatus, erhoben mittels HAQ-DI (siehe unten), validiert [35, 36].

Ansprechen gemäß ACR-Kriterien (ACR20/50/70)

Die ACR-Kriterien für ein Therapieansprechen berücksichtigen sieben zentrale Parameter der rheumatoiden Arthritis: die Anzahl der druckschmerzhaften und die Anzahl der geschwollenen Gelenke, die Einschätzung des Arztes und die Einschätzung des Patienten zur Krankheitsaktivität, Schmerzen, den Funktionsstatus (z.B. erhoben mittels HAQ-DI) und einen Marker für die Akute-Phase-Reaktion (CRP oder ESR) [47].

Ein 20%iges Ansprechen (ACR20) ist als eine mindestens 20%ige Verbesserung der Anzahl der druckschmerzhaften und geschwollenen Gelenke sowie eine mindestens 20%ige Verbesserung für mindestens drei der übrigen fünf Kriterien definiert [47]. Da mittlerweile die Schwelle von 20% für viele Patienten als zu niedrig angesehen wird und lediglich mit einer Linderung der Symptome einhergeht, wurden weitere Schwellenwerte wie ACR50 (entsprechend einer 50%igen Verbesserung) und ACR70 (entsprechend einer 70%igen Verbesserung) definiert und validiert [48]. Eine 50%ige sowie eine 70%ige Verringerung der Krankheitsaktivität sind im Gegensatz zur 20%igen Verbesserung als klinisch relevantes Therapieansprechen und somit als patientenrelevant einzuschätzen, insbesondere bei Patienten mit einer langjährigen hohen Krankheitsaktivität.

Kritisiert wird, dass das Ansprechen gemäß ACR-Kriterien – im Gegensatz zum Ansprechen gemäß EULAR-Kriterien – lediglich eine relative Aussage erlaubt und keine Information über die absolute Höhe der Krankheitsaktivität enthält [49]. So benötigen Patienten mit einer hohen Ausgangsaktivität eine stärkere Verringerung ihrer Krankheitsaktivität für eine klinisch relevante Verbesserung als Patienten mit moderater Krankheitsaktivität. Der Wert für das Ansprechen gemäß ACR-Kriterien erlaubt zudem keinen Rückschluss darauf, ob ein Patient den angestrebten Zustand der Remission oder niedrigen Krankheitsaktivität erreicht hat. Demzufolge wird die Erhebung von ACR20, ACR50 und ACR70 nicht für den klinischen Alltag empfohlen, wohingegen diese Parameter im Rahmen klinischer Studien eine Abschätzung des Anteils der Patienten mit relevantem Ansprechen in den verschiedenen Behandlungsarmen erlauben. Dementsprechend fordern die Zulassungsbehörden die Erhebung dieser Parameter im Rahmen der Zulassungsstudien für DMARDs [50].

*Morgensteifigkeit*

Morgensteifigkeit ist ein häufiges Symptom von Patienten mit rheumatoider Arthritis und bezeichnet die eingeschränkte Beweglichkeit der betroffenen Gelenke nach einer längeren Ruhephase, insbesondere am Morgen. Da ein Großteil der Patienten Morgensteifigkeit von über einer Stunde Dauer erlebt, bedeutet diese eine erhebliche Einschränkung der Lebensqualität der Patienten [51]. Eine Verringerung der Morgensteifigkeit ist somit als patientenrelevant einzuschätzen.

Das Ausmaß der Morgensteifigkeit wurde in der Studie MONARCH mittels einer 100 mm-VAS erhoben, auf der die Patienten angaben, wie stark sie durch Morgensteifigkeit in der zurückliegenden Woche beeinträchtigt waren.

*Fatigue*

Ein Großteil der Patienten mit rheumatoider Arthritis ist von einer klinisch relevanten Fatigue betroffen [52-54]. Chronische Fatigue beeinträchtigt die Fähigkeit der Patienten, alltägliche Aufgaben zu verrichten oder am sozialen Leben teilzuhaben, und hat einen negativen Einfluss auf die Lebensqualität der Patienten [52]. Somit stellt eine Verbesserung der Fatigue einen patientenrelevanten Endpunkt dar. Dies wurde auch vom G-BA im Rahmen des Beratungsgesprächs 2015-B-178 bestätigt [24].

Im Rahmen der Studie MONARCH wurde das Ausmaß der Fatigue mittels des Fragebogens FACIT-Fatigue (*Functional Assessment of Chronic Illness Therapy-Fatigue*) erhoben. Der FACIT-Fatigue wurde ursprünglich für Patienten mit Krebserkrankungen entwickelt und ist eine unabhängige Skala des FACIT-Messinstruments. Der FACIT-Fatigue wurde für verschiedene weitere Erkrankungen, unter anderem für Patienten mit rheumatoider Arthritis validiert [55]. So wurde gezeigt, dass die Ergebnisse des FACIT-Fatigue mit anderen, für rheumatoide Arthritis validierten Fatigue-Messinstrumenten korrelieren.

Der Fragebogen beinhaltet 13 Fragen, die das Ausmaß der Fatigue in den vergangen sieben Tagen erfassen, und die jeweils auf einer 5-Punkte Likert-Skala (von 0 = „gar nicht“ bis 4 = „sehr“) bewertet werden. Daraus ergibt sich nach Reversion der Rohwerte für die Fragen Nr. 1 bis 6 sowie Nr. 9 bis 13 (Reversion bedeutet, die Ausprägung 0 erhält 4 Punkte, die Ausprägung 4 demgegenüber 0 Punkte) ein Gesamtscore von 0 bis 52, wobei höhere Werte eine geringere Fatigue bedeuten. Cella et al. definierten in einer Validierungsstudie für Patienten mit rheumatoider Arthritis eine MCID zwischen 3 und 4, je nach angewandter Methodik [55]. Im Rahmen des vorliegenden Dossiers wurde für die Responderanalyse als MCID ein Wert von 4 gewählt, da dies den konservativeren Schwellenwert darstellt.

*Änderung des allgemeinen Gesundheitszustands*

Die Änderung des allgemeinen Gesundheitszustands wurde mit Hilfe des EQ-5D erhoben. Der EQ-5D ist ein standardisiertes, generisches Erhebungsinstrument zur Bestimmung des Gesundheitszustandes und der gesundheitsbezogenen Lebensqualität [56]. Der Fragebogen dient der schriftlichen Selbsteinschätzung in folgenden fünf Bereichen: Mobilität/Beweglichkeit, Fähigkeit, für sich selbst zu sorgen, allgemeine Tätigkeiten, Schmerzen/körperliche Beschwerden, Angst/Niedergeschlagenheit. Zusätzlich nehmen die Patienten auf einer VAS (0-100) eine Selbsteinschätzung zu ihrem Gesundheitszustand vor.

Der EQ-5D ist ein international anerkanntes und validiertes Messinstrument, der auch für Patienten mit rheumatoider Arthritis eingesetzt wird [57, 58]. Da der G-BA jedoch standardmäßig nur die Auswertung der EQ-5D-VAS zur Bewertung des Zusatznutzens in der Kategorie Morbidität heranzieht und die übrigen Ergebnisse des EQ-5D nicht berücksichtigt,

wurde im vorliegenden Dossier lediglich die Änderung des allgemeinen Gesundheitszustandes anhand der EQ-5D-VAS dargestellt.

***Endpunkte zur gesundheitsbezogenen Lebensqualität***

Rheumatoide Arthritis ist eine chronische Erkrankung, die mit einer erheblichen Einschränkung der Lebensqualität der Patienten einhergeht. Eine Verbesserung der Lebensqualität beispielsweise infolge einer Verringerung der funktionellen Einschränkung oder einer Verringerung von Schmerzen ist von großer Relevanz für die Patienten.

*SF-36*

Der *Medical Outcomes Study Short Form*-36 (SF-36) ist ein generisches Messinstrument zur Erhebung der gesundheitsbezogenen Lebensqualität. Beim SF-36 handelt es sich um einen häufig benutzten Fragebogen, der für den Einsatz in verschiedenen Indikationen validiert wurde [59, 60]. Der SF-36 enthält die Dimensionen „Körperliche Funktionsfähigkeit", „Körperliche Rollenfunktion", „Körperlicher Schmerz", „Allgemeine Gesundheitswahrnehmung", „Vitalität", „Soziale Funktionsfähigkeit", „Emotionale Rollenfunktion" und „Psychisches Wohlbefinden", die zu einer körperlichen und einer psychischen Summenskala zusammengefasst werden.

Für Patienten mit rheumatoider Arthritis wurde gezeigt, dass eine Verringerung der Krankheitsaktivität anhand der Parameter TJC, SJC, Schmerzen sowie PtGA und PhGA mit einer Verbesserung der Lebensqualität gemäß SF-36 einhergeht [61]. Jedoch zeigte sich in verschiedenen Studien, dass diese Korrelation in erster Linie auf die Domänen der körperlichen Gesundheit zurückzuführen ist, während die Domänen der psychischen Gesundheit weniger stark mit einer Veränderung der Krankheitsaktivität assoziiert sind [61, 62]. Der SF-36 scheint also weniger geeignet, die psychischen Aspekte der Lebensqualität von Patienten mit rheumatoider Arthritis zu erfassen.

In Studien mit Patienten mit rheumatoider Arthritis wurden Werte zwischen 2,5 und 5 Punkten als Schwellenwerte für eine klinisch relevante Verbesserung herangezogen [63-68]. Die Validität der meisten verwendeten Schwellenwerte ist jedoch unklar. In einer Studie von Ward et al. wurden für die körperliche Summenskala SF-36-PCS (SF-36 *Physical Component Summary*) Schwellenwerte für eine minimale klinisch relevante Verbesserung für Patienten mit rheumatoider Arthritis definiert [62]. Je nach verwendeter Methode wurden Werte von 2,5, 5,1 und 7,2 Punkten als MCID für die SF-36-PCS abgeleitet. Aufgrund der geringen Änderungssensitivität der psychischen Summenskala SF-36-MCS (SF-36 *Mental Component Summary*) konnte in der Studie von Ward et al. jedoch keine MCID für die SF-36-MCS abgeleitet werden [62].

Im vorliegenden Dossier wurden für die Auswertung der Lebensqualität gemäß SF-36 die MCIDs 2,5, 5,1 und 7,2 für die SF-36-PCS herangezogen. Für die Auswertung der SF-36-MCS wurde keine Responderanalyse durchgeführt, da gut validierte Responsekriterien fehlen.

*RAID*

Der *Rheumatoid Arthritis Impact of Disease* (RAID) ist ein krankheitsspezifischer Fragebogen zur Erhebung der Lebensqualität von Patienten mit rheumatoider Arthritis, der auf Initiative der EULAR entwickelt und in einer Studie mit europäischen Patienten mit rheumatoider Arthritis validiert wurde [69]. Der Fragebogen umfasst sieben numerische Ratingskalen (von 0 bis 10) für die Domänen „Schmerz“, „Funktionelle Einschränkungen“, „Fatigue“, „Schlaf“, „Körperliches Wohlbefinden“, „Emotionales Wohlbefinden“ und „Coping/Situationsbewältigung“. Die Werte der Skalen werden gewichtet zu einem Score zwischen 0 und 10 aufaddiert. Höhere Werte bedeuteten einen schlechteren Gesundheitszustand. Als minimale klinisch relevante Verbesserung wurde bei Patienten mit rheumatoider Arthritis eine Verbesserung um 3 Punkte oder 50 % definiert [70]. Im vorliegenden Dossier wurde für die Auswertung der Lebensqualität gemäß RAID eine Verbesserung um 3 Punkte als MCID für eine Responderanalyse verwendet.

***Endpunkte zur Sicherheit***

Die Erhebung von unerwünschten Ereignissen, die im zeitlichen Zusammenhang mit der Therapie auftreten, ist ein wesentlicher Bestandteil der Bewertung der Sicherheit eines Arzneimittels. Die Verringerung von therapieassoziierten Nebenwirkungen ist grundsätzlich als patientenrelevanter Aspekt eines zu bewertenden Arzneimittels anzusehen. Dennoch sind nicht alle untersuchten unerwünschten Ereignisse direkt patientenrelevant, da beispielsweise Änderungen von Laborparametern nicht per se für den Patienten spürbar und nicht zwingend symptomatisch sind.

Im Rahmen der Studie MONARCH wurde das Auftreten folgender unerwünschter Ereignisse (UE) erhoben:

- Gesamtrate UE
- UE nach Schweregrad (mild, moderat, schwer)
- Gesamtrate schwerwiegender UE (SUE)
- Gesamtrate der UE mit Todesfolge
- Gesamtrate der UE, die zum Studienabbruch führten
- UE von besonderem Interesse (UEBI)

Als UE von besonderem Interesse waren folgende Ereignisse präspezifiziert: Leukopenie, Thrombozytopenie, Infektionen, Leberwerterhöhungen, Divertikulitis, gastrointestinale Ulzerationen, Erhöhung der Lipidwerte, Überempfindlichkeit und Anaphylaxie, Reaktionen an der Injektionsstelle, Krebserkrankungen, Lupus-ähnliches Syndrom, demyelinisierende Erkrankung, Schwangerschaft, Überdosierung.

#### 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[8] erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär 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[9] werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes $I^2$ und des statistischen Tests auf Vorliegen von Heterogenität[10] erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

[8] *DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-188.*
[9] *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.*
[10] *Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.*

Es wird lediglich eine Studie zum Nachweis des Zusatznutzens von Sarilumab vorgelegt. Die Notwendigkeit einer Meta-Analyse entfällt daher.

#### 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.*

In der eingeschlossenen Studie MONARCH wurden zwei Sensitivitätsanalysen für den primären Endpunkt „Änderung des DAS28-ESR nach 24 Wochen“ a priori geplant und durchgeführt:

- In der ersten Analyse wurden die Daten der Patienten, die die Studie abgebrochen haben, eingeschlossen. Dafür wurden Patienten in Woche 24 erneut untersucht, auch wenn sie die Studienteilnahme beendet hatten.

- In der zweiten Analyse wurden alle Daten nach Studienabbruch oder der Dosiserhöhung in der Adalimumab-Gruppe als fehlend gesetzt und mittels Multipler Imputation ersetzt.

#### 4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

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

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

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

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

*Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

In der Studie MONARCH wurden für den primären Endpunkt „Änderung des DAS28-ESR nach 24 Wochen“ sowie für den sekundären Endpunkt „Anteil der Patienten mit DAS28-ESR-Remission in Woche 24“ Subgruppenanalysen für folgende Faktoren a priori geplant:

- Geschlecht (Männer / Frauen)
- Rasse (kaukasisch bzw. weiß / andere)
- Region (Region 1 / Region 2 / Region 3)[11]
- Alter (< 65 Jahre / ≥ 65 Jahre)
- Gewicht (< 60 kg / ≥ 60 kg und < 100 kg / ≥ 100 kg)
- BMI (< 25 / ≥ 25 und < 30 / ≥ 30 kg/m$^2$)
- Rheumafaktor (positiv / negativ)
- Anti-CCP-Antikörper (positiv/ negativ)
- Baseline CRP (≤ 15 mg/L / > 15 mg/L)
- Baseline ESR (≤ Median / > Median)

[11] Region 1 (Westliche Länder): Tschechien, Deutschland, Ungarn, Israel, Spanien, USA; Region 2 (Südamerika): Chile, Peru; Region 3 (andere): Südkorea, Polen, Südafrika, Rumänien, Russland, Ukraine

- Dauer der RA (≤ Median / > Median in Jahren; ≤ 3 Jahre / > 3 Jahre)
- Anzahl der vorangegangenen DMARDs (≤ 1 / 2 / ≥ 3)
- MTX-Status (intolerant / ungeeignet / Non-Responder)
- Raucherstatus (Nicht-Raucher / Raucher / Ex-Raucher)

Für alle weiteren Endpunkte wurden keine Subgruppenanalysen im Studienprotokoll oder SAP a priori festgelegt.

Die Ergebnisse der a priori geplanten Subgruppenanalysen sind für den Nachweis des Zusatznutzens von Sarilumab nicht relevant, da die a priori geplanten Subgruppenanalysen die gesamte Studienpopulation der Studie MONARCH betreffen. Diese umfasst auch Patienten, die für die Nutzenbewertung nicht relevant sind (siehe Abschnitt 4.3.1.2.1). Somit ermöglichen die a priori geplanten Subgruppenanalysen keine Aussage zur relevanten Studienpopulation und werden im vorliegenden Dossier nicht dargestellt. Im Übrigen ergaben die im Studienbericht gezeigten Subgruppenanalysen der gesamten Studienpopulation von MONARCH keine klinisch relevante Effektmodifikation [71].

Um das Vorliegen einer Effektmodifikation in der relevanten Studienpopulation ausschließen zu können, wurden die a priori geplanten Subgruppenanalysen (siehe oben) post hoc für die relevante Studienpopulation berechnet [72].

Weiterhin wurden für die Nutzenbewertung von Sarilumab post hoc Subgruppenanalysen für die Endpunkte, für die keine Subgruppenanalysen a priori geplant waren, nach Geschlecht (Männer / Frauen), Alter (< 65 Jahre / ≥ 65 Jahre) und Region (Region 1: Westliche Staaten / Region 2: Süd-Amerika/ Region 3: andere) durchgeführt [72, 73]. Für diese post-hoc-Analysen wurde ausschließlich die relevante Studienpopulation herangezogen. Eine Subgruppenanalyse nach Krankheitsschwere wurde als nicht sinnvoll erachtet, da alle Patienten der Studie MONARCH gemäß der im Studienprotokoll definierten Einschlusskriterien eine hohe Krankheitsaktivität aufwiesen.

Für die Subgruppenanalysen der stetigen Endpunkte wurde das MMRM-Modell um die Interaktionsterme Behandlung*Visite, Subgruppe*Behandlung und Subgruppe*Behandlung* Visite erweitert. Für die Subgruppenanalysen binärer Endpunkte wurden Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte mittels der SAS-Prozedur PROC GLIMMIX berechnet. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Für die Bewertung möglicher Effektmodifikatoren wurde entsprechend der IQWiG Methodik ein Interaktions-p-Wert < 0,05 als Beleg für eine Effektmodifikation durch ein Merkmal gewertet [74]. Wurden nach dieser Methodik zwischen den Subgruppen signifikant unterschiedliche Effekte ermittelt (p (Interaktion) < 0,05), wurden die Ergebnisse auf Konsistenz über verschiedene Endpunkte hinweg sowie auf klinische Plausibilität geprüft.

### 4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich[12]. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“[13], „Multiple-Treatment-Meta-Analysen“[14] oder auch „Netzwerk-Meta-Analysen“[15], sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist[16].

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in Frage kommen, vorliegen. Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen[17].

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

- *Benennung des Brückenkomparators bzw. der Brückenkomparatoren und Begründung für die Auswahl.*

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

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*

- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*

- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*

- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*

- *Art und Umfang von Sensitivitätsanalysen.*

Aussagen zur Nutzenbewertung von neuen Interventionen sollten vorzugsweise aus Ergebnissen direkt vergleichender Studien abgeleitet werden. Stehen jedoch keine Head-to-Head-Studien zur Verfügung, die das zu bewertende Arzneimittel mit der zweckmäßigen Vergleichstherapie direkt vergleichen, ist eine Durchführung von adjustierten indirekten Vergleichen möglich. Die Aussagen, die auf indirekten Vergleichen beruhen, sind jedoch grundsätzlich mit einer geringeren Ergebnissicherheit verbunden als Nutzenbewertungen basierend auf direkten Vergleichen.

Für die Nutzenbewertung von Sarilumab liegt direkt vergleichende Evidenz für Sarilumab als Monotherapie bei Patienten mit MTX-Intoleranz in den Teilanwendungsgebieten 2b und 3b vor. Für die übrigen Teilanwendungsgebiete steht keine direkt vergleichende Evidenz zur Verfügung. Auf die Durchführung von indirekten Vergleichen wurde im vorliegenden Dossier jedoch aus methodischen Gründen verzichtet.

Um valide Aussagen basierend auf indirekten Vergleichen treffen zu können, müssen grundlegende Annahmen hinsichtlich Ähnlichkeit, Homogenität und Konsistenz erfüllt sein. Der Begriff der Ähnlichkeit bezieht sich auf die Untersuchung der Vergleichbarkeit der Studien hinsichtlich möglicher Effektmodifikatoren. Das bedeutet konkret, dass die eingeschlossenen Studien eine möglichst hohe Übereinstimmung hinsichtlich der Patientencharakteristika und des Studiendesigns aufweisen müssen, um vergleichbar zu sein. Der Begriff der Homogenität bezeichnet die Vergleichbarkeit der geschätzten Effekte innerhalb aller paarweisen Vergleiche. Der Begriff der Konsistenz hingegen bezieht sich auf die Vergleichbarkeit der geschätzten Effekte aus indirekter und direkter Evidenz und kann bei Fehlen von direkter Evidenz dementsprechend nicht untersucht werden. Sofern eine oder mehrere der grundlegenden Annahmen nicht in ausreichendem Maße erfüllt sind, sollte gemäß IQWiG-Methodenpapier von der Anwendung indirekter Vergleiche abgesehen werden [74].

Die Therapie der rheumatoiden Arthritis wurde in den letzten zwanzig Jahren grundlegend weiterentwickelt. Mit den bDMARDs stehen Patienten heute deutlich mehr und wirksamere Therapieoptionen zur Verfügung [3]. Zudem wurden die Kriterien zur Diagnose der rheumatoiden Arthritis dahingehend weiterentwickelt, dass heute eine frühere Diagnosestellung und damit ein früherer Beginn der Behandlung mit DMARDs möglich ist [75, 76]. Diese Faktoren haben dazu geführt, dass die durchschnittliche Krankheitsschwere der Patienten im Laufe der Jahre gesunken ist [77]. Zum einen zeigen Studien einen gestiegenen Anteil der Patienten, die eine Remission oder niedrige Krankheitsaktivität erreichen, zum anderen hat die Schwere der irreversiblen Gelenkschäden abgenommen [78-81]. Heutige Patienten mit rheumatoider Arthritis sind somit nicht mit den Patienten vor 20 Jahren vergleichbar. Zudem beeinträchtigt die regionale Variabilität hinsichtlich Patientencharakteristika und Therapiestandards die Vergleichbarkeit der Patientenpopulationen in klinischen Studien zur rheumatoiden Arthritis.

Zusammenfassend ist davon auszugehen, dass eine hinreichend valide Bewertung des Zusatznutzens von Sarilumab auf Basis von indirekten Vergleichen klinischer Studien aufgrund der oben dargelegten hohen methodischen Anforderungen in Verbindung mit den beschriebenen Besonderheiten der Indikation nicht möglich ist. Hinzu kommt, dass bei Nicht-Erfüllung der genannten Voraussetzungen indirekte Vergleiche regelhaft im Rahmen der Nutzenbewertung keine Berücksichtigung finden. Entsprechend wurde im vorliegenden Dossier auf die Durchführung von indirekten Vergleichen verzichtet.

## 4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

### 4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

#### 4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

##### 4.3.1.1.1 Studien des pharmazeutischen Unternehmers

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

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

- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

*Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.*

Tabelle 4-9: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

| Studie | Zulassungsstudie[a] (ja/nein) | Sponsor (ja/nein) | Status (abgeschlossen / abgebrochen / laufend) | Studiendauer[b] | Therapiearme[c] |
|---|---|---|---|---|---|
| **Phase III-/ Phase II-Studien** | | | | | |
| ASCERTAIN / SFY13370 / NCT01768572 | ja | ja | abgeschlossen | 24 Wochen (6 Wochen) | Sarilumab 150 mg q2w<br>Sarilumab 200 mg q2w<br>Tocilizumab 4-8 mg/kg q4w |
| COMPARE / EFC11574 / NCT01764997 | ja | ja | abgebrochen[d] | 24 Wochen (6 Wochen) | Sarilumab 150 mg q2w<br>Sarilumab 200 mg q2w<br>Etanercept 50 mg qw |
| MONARCH / EFC14092 / NCT02332590 | ja | ja | laufend | 24 Wochen (max. 276 Wochen/bis 2020) | Sarilumab 200 mg q2w<br>Adalimumab 40 mg q2w |
| MOBILITY / EFC11072 / NCT01061736 | ja | ja | abgeschlossen | A: 12 Wochen<br>B: 52 Wochen<br>(6 Wochen) | Teil A:<br>Sarilumab 100 mg q2w<br>Sarilumab 150 mg q2w<br>Sarilumab 200 mg q2w<br>Sarilumab 100 mg q2w<br>Sarilumab 150 mg q2w<br>Placebo<br>Teil B:<br>Sarilumab 150 mg q2w<br>Sarilumab 200 mg q2w<br>Placebo |
| TARGET / EFC10832 / NCT01709578 | ja | ja | abgeschlossen | 24 Wochen (6 Wochen) | Sarilumab 150 mg q2w<br>Sarilumab 200 mg q2w<br>Placebo |
| RA-ONE / EFC13752 / NCT02121210 | ja | ja | abgeschlossen | 24 Wochen (6 Wochen) | Sarilumab 200 mg q2w<br>Sarilumab 150 mg q2w |

| Studie | Zulassungsstudie[a] (ja/nein) | Sponsor (ja/nein) | Status (abgeschlossen / abgebrochen / laufend) | Studiendauer[b] | Therapiearme[c] |
|---|---|---|---|---|---|
| RA-EASY / MSC12665 / NCT02057250 | ja | ja | abgeschlossen | 12 Wochen (52 Wochen) | Sarilumab 150 mg q2w (Spritze)<br>Sarilumab 200 mg q2w (Spritze)<br>Sarilumab 150 mg q2w (Pen)<br>Sarilumab 200 mg q2w (Pen) |
| HARUKA / LTS13618 / NCT02373202 | nein | ja | abgeschlossen | 52 Wochen (6 Wochen) | Sarilumab 200 mg q2w<br>Sarilumab 150 mg q2w |
| KAKEHASI / EFC14059 / NCT02293902 | nein | ja | abgeschlossen | 52 Wochen (6 Wochen) | Sarilumab 200 mg q2w<br>Sarilumab 150 mg q2w<br>Placebo |
| ACT11575 / NCT01217814 | ja | ja | abgebrochen[e] | 12 Wochen | Sarilumab 150 mg qw<br>Golimumab<br>Placebo |
| **Phase-I-Studien** | | | | | |
| TDU10808 / 6R88-RA-0703 | ja | ja | abgeschlossen | 1 Tag (16 Wochen) | Sarilumab 0.6mg/kg i.v.<br>Sarilumab 2 mg/kg i.v.<br>Placebo |
| TDU10809 / 6R88-RA-0801 / NCT01055899 | ja | ja[e] | abgeschlossen | 1 Tag (16 Wochen) | Sarilumab 50 mg<br>Sarilumab 100 mg<br>Sarilumab 200 mg<br>Placebo |
| ACT10804 / 6R88-RA-0803 / NCT01026519 | ja | ja[e] | abgeschlossen | 1 Tag (6 Wochen) | Sarilumab 50 mg<br>Sarilumab 100 mg<br>Sarilumab 200 mg<br>Placebo |
| PKM12058 / NCT01328522 | ja | ja | abgeschlossen | 1 Tag (5 Wochen) | Sarilumab 200 mg<br>Sarilumab 200 mg |
| TDU13402 / NCT01850680 | ja | ja | abgeschlossen | 1 Tag (8 Wochen) | Sarilumab 50 mg<br>Sarilumab 100 mg<br>Sarilumab 200 mg<br>Placebo |
| TDR10805 / 6R88-RA-0802 / NCT01011959 | ja | ja[e] | abgeschlossen | 5 Wochen (5 Wochen) | Sarilumab 50 mg qw<br>Sarilumab 100 mg qw |

| Studie | Zulassungsstudie[a] (ja/nein) | Sponsor (ja/nein) | Status (abgeschlossen / abgebrochen / laufend) | Studiendauer[b] | Therapiearme[c] |
|---|---|---|---|---|---|
| | | | | | Sarilumab 150 mg qw<br>Sarilumab 100 mg q2w<br>Sarilumab 150 mg q2w<br>Sarilumab 200 mg q2w<br>Placebo |
| PDY14191 / NCT02404558 | nein | ja | abgeschlossen | 1 Tag<br>(6 Wochen) | Sarilumab 150 mg<br>Tocilizumab 162 mg |
| 6R88-RA-1309 / NCT02097524 | ja | ja[e] | abgeschlossen | 1 Tag<br>(6 Wochen) | Sarilumab 150 mg<br>Sarilumab 200 mg<br>Tocilizumab 4 mg/kg<br>Tocilizumab 8 mg/kg |

[a] Als Zulassungsstudien gelten alle Studien, die im Rahmen des Zulassungsantrags für Sarilumab bei der EMA eingereicht wurden und somit Bestandteil des Modul 5 des CTD sind.

[b] Studiendauer: Angegeben ist die Dauer der randomisierten Behandlungsphase sowie die Dauer des Follow-ups bzw. ggf. die Dauer der offen Extensionsphase (in Klammern).

[c] Sofern nicht anders angegeben ist die Applikationsroute für Sarilumab immer subkutan.

[d] Die Studie COMPARE wurde vorzeitig abgebrochen, da der Einschluss von Patienten in die randomisierte doppelblinde Behandlungsphase nicht im vorgesehenen Zeitplan realisiert werden konnte. Die Beendigung der Studie geschah nicht aus Sicherheitsgründen.

[e] Die Studie ACT11575 wurde vorzeitig aufgrund von Verzögerungen in der Studiendurchführung abgebrochen. Die Beendigung der Studie geschah nicht aus Sicherheitsgründen.

[f] Hauptsponsor der Studien war Regeneron Pharmaceuticals unter Beteiligung von Sanofi in Form einer Kollaboration.

*Geben Sie an, welchen Stand die Information in Tabelle 4-9 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.*

Der hier abgebildete Studienstatus entspricht dem Stand vom 26.06.2017.

*Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-9 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-10: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

| **Studienbezeichnung** | **Begründung für die Nichtberücksichtigung der Studie** |
|---|---|
| ASCERTAIN / SFY13370 / NCT01768572 | A3: Die Dosierung des Komparators entspricht nicht der in Deutschland zugelassenen Dosierung.[a] |
| COMPARE / EFC11574 / NCT01764997 | A7: Die Studie wurde vorzeitig beendet, es liegt lediglich ein verkürzter Studienbericht vor.[b] |
| MOBILITY / EFC11072 / NCT01061736 | A3: Der Komparator entspricht nicht der zVT. |
| TARGET / NCT01709578 | A3: Der Komparator entspricht nicht der zVT. |
| RA-ONE / EFC13752 / NCT02121210 | A3: Der Komparator entspricht nicht der zVT. |
| RA-EASY / MSC12665 / NCT02057250 | A6: Studiendauer < 24Wochen |
| HARUKA / LTS13618 / NCT02373202 | A3: Der Komparator entspricht nicht der zVT. |
| KAKEHASI / EFC14059 / NCT02293902 | A3: Der Komparator entspricht nicht der zVT. |
| ACT11575 / NCT01217814 | A6: Studiendauer < 24Wochen |
| TDU10808 / 6R88-RA-0703 | A6: Studiendauer < 24Wochen |
| TDU10809 / 6R88-RA-0801 / NCT01055899 | A6: Studiendauer < 24Wochen |
| ACT10804 / 6R88-RA-0803 / NCT01026519 | A6: Studiendauer < 24Wochen |
| PKM12058 / NCT01328522 | A6: Studiendauer < 24Wochen |
| TDU13402 / NCT01850680 | A6: Studiendauer < 24Wochen |
| TDR10805 / 6R88-RA-0802 / NCT01011959 | A6: Studiendauer < 24Wochen |
| PDY14191 / NCT02404558 | A6: Studiendauer < 24Wochen |
| 6R88-RA-1309 / NCT02097524 | A6: Studiendauer < 24Wochen |

[a] In der Studie ASCERTAIN wurde die Dosierung des Studienkomparators Tocilizumab der US-amerikanischen Zulassung entsprechend gewählt. Diese sieht eine initiale Dosis von 4 mg/kg Tocilizumab i.v. alle vier Wochen vor, welche bei unzureichendem Ansprechen auf 8 mg/kg alle vier Wochen erhöht werden kann. In Deutschland ist lediglich die Dosierung 8 mg/kg Tocilizumab i.v. alle vier Wochen zugelassen.

[b] Die Studie COMPARE wurde vorzeitig abgebrochen, da der Einschluss von Patienten in die randomisierte doppelblinde Behandlungsphase nicht im vorgesehenen Zeitplan realisiert werden konnte. Die Beendigung der Studie geschah nicht aus Sicherheitsgründen. Insgesamt wurden 43 Patienten in die doppelblinde Behandlungsphase eingeschlossen und auf drei Behandlungsarme randomisiert. Von diesen 43 Patienten erhielten lediglich 11 Patienten die vorgesehene 24-wöchige Behandlung mit Sarilumab oder Etanercept. Aus diesem Grund wurde keine Auswertung zur Wirksamkeit der Behandlung durchgeführt. Die Auswertungen zu den geplanten Sicherheitsendpunkten sind im verkürzten Studienbericht dargestellt. Aufgrund der kleinen Patientenzahl und überwiegend zu kurzen Behandlungsdauer sind diese Auswertungen für die Nutzenbewertung von Sarilumab nicht relevant.

#### 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.*

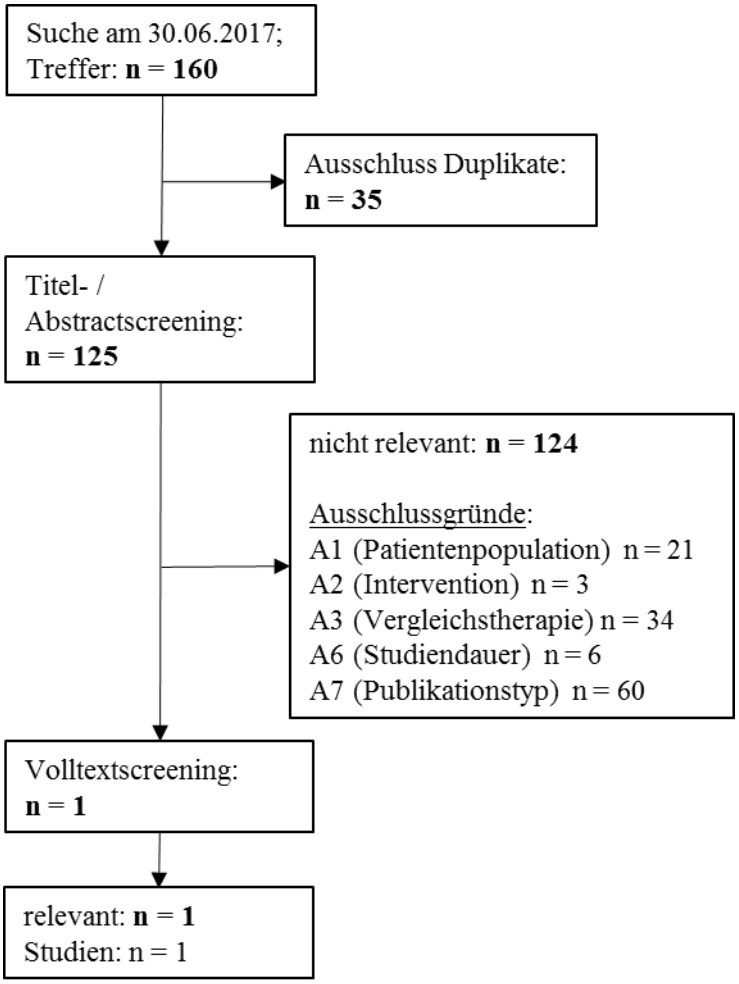



Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die bibliographische Literaturrecherche erfolgte am 30.06.2017. Die Suche in MEDLINE und EMBASE über Ovid sowie in der Cochrane Library ergab insgesamt 160 Treffer, davon waren 35 Dubletten. Die verbleibenden 125 Treffer wurden anhand von Titel und Abstract gescreent. 124 Treffer wurden als nicht relevant ausgeschlossen (Ausschlussgründe siehe Flussdiagramm). Die verbleibende Publikation zur Studie MONARCH wurde nach Volltextsichtung als relevant eingestuft.

#### 4.3.1.1.3 Studien aus der Suche in Studienregistern

*Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe*

*Tabelle 4-9) 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 4-11: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

| **Studie** | **Identifikationsorte (Name des Studienregisters und Angabe der Zitate[a])** | **Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)** | **Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)** | **Status (abgeschlossen/ abgebrochen/ laufend)** |
|---|---|---|---|---|
| MONARCH | clinicaltrials.gov [82]<br>EU-CTR [83]<br>ICTRP [84, 85]<br>PharmNet.Bund [86] | ja | ja | laufend[b] |
| a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.<br>b: Die 24-wöchige randomisierte Behandlungsphase ist abgeschlossen, die offene Extensionsphase der Studie läuft voraussichtlich bis 2020. | | | | |

*Geben Sie an, welchen Stand die Information in Tabelle 4-11 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 Suche nach relevanten Studien in den Studienregistern wurde vom 26.06. bis zum 30.06.2017 durchgeführt. Die Suche ergab insgesamt 79 Treffer zu 24 Studien. Davon wurde die Studie MONARCH als relevant eingeschlossen, alle übrigen Studien wurden ausgeschlossen (Ausschlussgründe siehe Anhang 4-D).

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

*Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.*

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

Tabelle 4-12: 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)** | **Studienbericht (ja/nein [Zitat])** | **Registereintrag**[c] **(ja/nein [Zitat])** | **Publikation (ja/nein [Zitat])** |
| **aktivkontrolliert, zweckmäßige Vergleichstherapie(n)** | | | | | | |
| **MONARCH** | **ja** | **ja** | **nein** | **ja [71]** | **ja [82-86]** | **ja [87]** |

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

b: Studie, für die der Unternehmer Sponsor war.

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

### 4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

#### 4.3.1.2.1 Studiendesign und Studienpopulationen

*Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.*

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

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

| **Studie** | **Studiendesign** <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.> | **Population** <relevante Charakteristika, z. B. Schweregrad> | **Interventionen (Zahl der randomisierten Patienten)** | **Studiendauer** <ggf. Run-in, Behandlung, Nachbeobachtung> | **Ort und Zeitraum der Durchführung** | **Primärer Endpunkt; patientenrelevante sekundäre Endpunkte** |
|---|---|---|---|---|---|---|
| MONARCH | RCT, doppelblind, parallel | Erwachsene mit aktiver rheumatoider Arthritis (RA) die entweder intolerant gegenüber Methotrexat (MTX) oder ungeeignet für eine Weiterbehandlung mit MTX sind, oder die nach mindestens 12 Wochen Behandlung mit MTX unzureichend angesprochen haben. | <u>Sarilumab 200 mg</u><br>(n = 184)<br>relevante Studienpopulation: Patienten mit MTX-Intoleranz (n = 87)<br><br><u>Adalimumab 40 mg</u><br>(n = 185)<br>relevante Studienpopulation: Patienten mit MTX-Intoleranz (n = 82) | Screening:<br>4 Wochen<br><br>doppelblinde Behandlungsphase:<br>24 Wochen<br><br>offene Extensionsphase:<br>max. 276 Wochen | 86 Zentren in 14 Ländern aufgeteilt in 3 Regionen:<br>Region 1 (Westliche Länder): Tschechien, Deutschland, Ungarn, Israel, Spanien, USA<br>Region 2 (Südamerika): Chile, Peru<br>Region 3 (andere): Südkorea, Polen, Südafrika, Rumänien, Russland, Ukraine<br><br>02/2015 – 12/2020[a] | <u>primärer Endpunkt</u>:<br>Änderung des DAS28-ESR<br><u>sekundäre Endpunkte</u>:<br>DAS28-ESR-Remission, DAS28-CRP-Remission<br>DAS28-ESR-Niedrige Krankheitsaktivität , DAS28-CRP-Niedrige Krankheitsaktivität<br>CDAI-Remission<br>Ansprechen gemäß EULAR-Kriterien<br>ACR20/50/70<br>68TJC, 66SJC<br>Schmerz<br>HAQ-DI<br>PtGA<br>FACIT-Fatigue<br>Morgensteifigkeit<br><u>Lebensqualität</u>:<br>SF-36<br>RAID<br>EQ-5D-3L<br><u>Sicherheitsendpunkte</u> |

[a] Spätestes geplantes Ende der offenen Extensionsphase von MONARCH.

Tabelle 4-14: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

| Studie | Intervention | Komparator |
|---|---|---|
| MONARCH | Sarilumab 200 mg s.c. + Placebo für Adalimumab s.c. q2w | Adalimumab 40 mg s.c. + Placebo für Sarilumab q2w[a] |

s.c.: subkutan; q2w: alle 2 Wochen

[a] Eine Dosiserhöhung auf wöchentlich 40 mg Adalimumab ab Woche 16 bis einschließlich Woche 23 war erlaubt, falls ein Patient in zwei aufeinanderfolgenden Visiten eine Verbesserung um < 20% gegenüber dem Ausgangswert bei der Anzahl der geschwollenen (SJC) und druckschmerzhaften Gelenke (TJC) hatte. Alternativ konnten solche Patienten die Studie abbrechen und erhielten dann eine individuelle Therapie.

Tabelle 4-15: Charakterisierung der Studienpopulation (demographische Charakteristika) – RCT mit dem zu bewertenden Arzneimittel

| **Studie** | **Alter** | **Geschlecht** | | **Rasse[a]** | | **BMI** | **Region[b]** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Teilpopulation** | **Jahre** | **w** | **m** | **kauk.** | **andere** | **kg/m²** | **1** | **2** | **3** |
| Behandlung (RP) | MW (SD) | n (%) | n (%) | n (%) | n (%) | MW (SD) | n (%) | n (%) | n (%) |
| **MONARCH** | | | | | | | | | |
| Gesamte Studienpopulation | | | | | | | | | |
| Sarilumab (N = 184) | 50,9 (12,6) | 157 (85,3) | 27 (14,7) | 171 (92,9) | 13 (7,0) | 27,09 (5,64) | 61 (33,2) | 36 (19,6) | 87 (47,3) |
| Adalimumab (N = 185) | 53,6 (11,9) | 150 (81,1) | 35 (18,9) | 164 (88,6) | 21 (11,4) | 27,26 (6,45) | 62 (33,5) | 35 (18,9) | 88 (47,6) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz)[c] | | | | | | | | | |
| Sarilumab (N = 87) | 53,0 (12,7) | 76 (87,4) | 11 (12,6) | 86 (98,9) | 1 (1,1) | 27,73 (5,82) | 40 (46,0) | 17 (19,5) | 30 (34,5) |
| Adalimumab (N = 82) | 54,0 (12,0) | 71 (86,6) | 11 (13,4) | 79 (96,3) | 3 (3,6) | 27,84 (5,70) | 31 (37,8) | 24 (29,3) | 27 (32,9) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD)[c] | | | | | | | | | |
| Sarilumab (N = 35) | 53,4 (15,0) | 29 (82,9) | 6 (17,1) | 35 (100) | 0 | 27,23 (6,29) | 14 (40,0) | 5 (14,3) | 16 (45,7) |
| Adalimumab (N = 35) | 53,1 (11,3) | 29 (82,9) | 6 (17,1) | 34 (97,1) | 1 (2,9) | 28,98 (6,62) | 14 (40,0) | 5 (14,3) | 16 (45,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs)[c] | | | | | | | | | |
| Sarilumab (N = 52) | 52,8 (11,0) | 47 (90,4) | 5 (9,6) | 51 (98,1) | 1 (1,9) | 26,89 (5,51) | 26 (50,0) | 12 (23,1) | 14 (26,9) |
| Adalimumab (N = 47) | 54,7 (12,5) | 42 (89,4) | 5 (10,6) | 45 (95,7) | 2 (4,3) | 27,00 (4,81) | 17 (36,2) | 19 (40,4) | 11 (23,4) |

BMI: Body Mass Index; MW: Mittelwert; RP: randomisierte Population; SD: Standardabweichung

[a] kaukasisch: weiß; andere: schwarz, asiatisch, orientalisch oder andere

[b] Region 1 (Westliche Länder): Tschechien, Deutschland, Ungarn, Israel, Spanien, USA; Region 2 (Südamerika): Chile, Peru; Region 3 (andere): Südkorea, Polen, Südafrika, Rumänien, Russland, Ukraine

[c] Zur Definition der im Dossier dargestellten Teilpopulationen der Studie MONARCH siehe Abschnitt „Für die Nutzenbewertung relevante Studienpopulation“ auf Seite 83.

Quellen: [71, 88]

Tabelle 4-16: Charakterisierung der Studienpopulation (Basischarakteristika der Erkrankung I) – RCT mit dem zu bewertenden Arzneimittel

| **Studie** | **Zeit seit Diagnose** | **Funktionelle Klasse[a]** | | | **Rheuma-faktor** | **Anti-CCP-Antikörper** | **HAQ-DI** |
|---|---|---|---|---|---|---|---|
| **Teilpopulation** | **Jahre** | **I** | **II** | **III** | **positiv** | **positiv** | |
| Behandlung (RP) | MW (SD) | n (%) | n (%) | n (%) | n (%) | n (%) | MW (SD) |
| **MONARCH** | | | | | | | |
| Gesamte Studienpopulation | | | | | | | |
| Sarilumab (N = 184) | 8,1 (8,1) | 29 (15,8) | 125 (67,9) | 30 (16,3) | 119 (66,9) | 134 (75,3) | 1,64 (0,55) |
| Adalimumab (N = 185) | 6,6 (7,8) | 37 (20,0) | 115 (62,2) | 33 (17,8) | 116 (64,8) | 138 (76,7) | 1,63 (0,64) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz)[b] | | | | | | | |
| Sarilumab (N = 87) | 9,4 (9,2) | 18 (20,7) | 51 (58,6) | 18 (20,7) | 58 (69,0) | 66 (78,6) | 1,70 (0,58) |
| Adalimumab (N = 82) | 7,6 (8,6) | 21 (25,6) | 48 (58,5) | 13 (15,9) | 49 (62,0) | 60 (75,9) | 1,71 (0,72) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD)[b] | | | | | | | |
| Sarilumab (N = 35) | 6,3 (8,3) | 7 (20,0) | 23 (65,7) | 5 (14,3) | 23 (67,6) | 25 (71,4) | 1,68 (0,62) |
| Adalimumab (N = 35) | 3,9 (4,1) | 7 (20,0) | 22 (62,9) | 6 (17,1) | 16 (45,7) | 21 (61,8) | 1,54 (0,66) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs)[b] | | | | | | | |
| Sarilumab (N = 52) | 11,4 (9,3) | 11 (21,2) | 28 (53,8) | 13 (25,0) | 35 (70,0) | 41 (83,7) | 1,71 (0,56) |
| Adalimumab (N = 47) | 10,5 (9,9) | 14 (29,8) | 26 (55,3) | 7 (14,9) | 33 (75,0) | 39 (86,7) | 1,83 (0,73) |

CCP: cyclisches Citrullin; HAQ-DI: *Health Assessment Questionnaire Disability Index;* MW: Mittelwert; RP: randomisierte Population; SD: Standardabweichung

[a] Funktionelle Klasse gemäß ACR-Kriterien von 1991

[b] Zur Definition der im Dossier dargestellten Teilpopulationen der Studie MONARCH siehe Abschnitt „Für die Nutzenbewertung relevante Studienpopulation“ auf Seite 83.

Quellen: [71, 88]

Tabelle 4-17: Charakterisierung der Studienpopulation (Basischarakteristika der Erkrankung II) – RCT mit dem zu bewertenden Arzneimittel

| **Studie**<br>**Gruppe**<br>(RP) | **68TJC**<br>MW (SD) | **66SJC**<br>MW (SD) | **CRP**<br>**mg/L**<br>MW (SD) | **ESR**<br>**mm/h**<br>MW (SD) | **DAS28-ESR**<br>MW (SD) | **CDAI**<br>MW (SD) |
|---|---|---|---|---|---|---|
| **MONARCH** | | | | | | |
| Gesamte Studienpopulation | | | | | | |
| Sarilumab<br>(N = 184) | 28,0<br>(13,2) | 18,6<br>(10,7) | 17,4<br>(21,3) | 46,5<br>(21,8) | 6,83<br>(0,76) | 43,62<br>(12,10) |
| Adalimumab<br>(N = 185) | 26,7<br>(13,6) | 17,5<br>(10,3) | 24,1<br>(31,0) | 47,5<br>(23,2) | 6,76<br>(0,83) | 42,40<br>(11,97) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz)[a] | | | | | | |
| Sarilumab<br>(N = 87) | 27,72<br>(12,58) | 19,07<br>(10,80) | 16,81<br>(21,63) | 50,08<br>(23,89) | 6,89<br>(0,64) | 43,52<br>(11,03) |
| Adalimumab<br>(N = 82) | 27,78<br>(13,69) | 19,30<br>(11,00) | 23,70<br>(30,14) | 48,56<br>(23,26) | 6,85<br>(0,82) | 43,68<br>(12,26) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD)[a] | | | | | | |
| Sarilumab<br>(N = 35) | 27,46<br>(12,92) | 18,86<br>(11,29) | 17,49<br>(26,88) | 48,09<br>(22,27) | 6,90<br>(0,53) | 43,57<br>(9,86) |
| Adalimumab<br>(N = 35) | 27,89<br>(11,83) | 18,63<br>(9,92) | 21,68<br>(31,18) | 45,37<br>(19,04) | 6,75<br>(0,72) | 42,91<br>(11,82) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs)[a] | | | | | | |
| Sarilumab<br>(N = 52) | 27,89<br>(12,48) | 19,22<br>(10,57) | 16,35<br>(17,51) | 51,42<br>(25,04) | 6,89<br>(0,71) | 43,49<br>(11,85) |
| Adalimumab<br>(N = 47) | 27,70<br>(15,05) | 19,81<br>(11,82) | 25,20<br>(29,58) | 50,94<br>(25,91) | 6,92<br>(0,89) | 44,26<br>(12,68) |

TJC: *Tender Joint Count* von 68 Gelenken; SJC: *Swollen Joint Count* von 66 Gelenken; CRP: C-reaktives Protein; ESR: Erythrozytensedimentationsrate; DAS28: *Disease Activity Score with 28 Joint Counts*; CDAI: *Clinical Disease Activity Index*; RP: randomisierte Population; MW: Mittelwert; SD: Standardabweichung

[a] Zur Definition der im Dossier dargestellten Teilpopulationen der Studie MONARCH siehe Abschnitt „Für die Nutzenbewertung relevante Studienpopulation“ auf Seite 83.

Quellen: [71, 88]

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

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

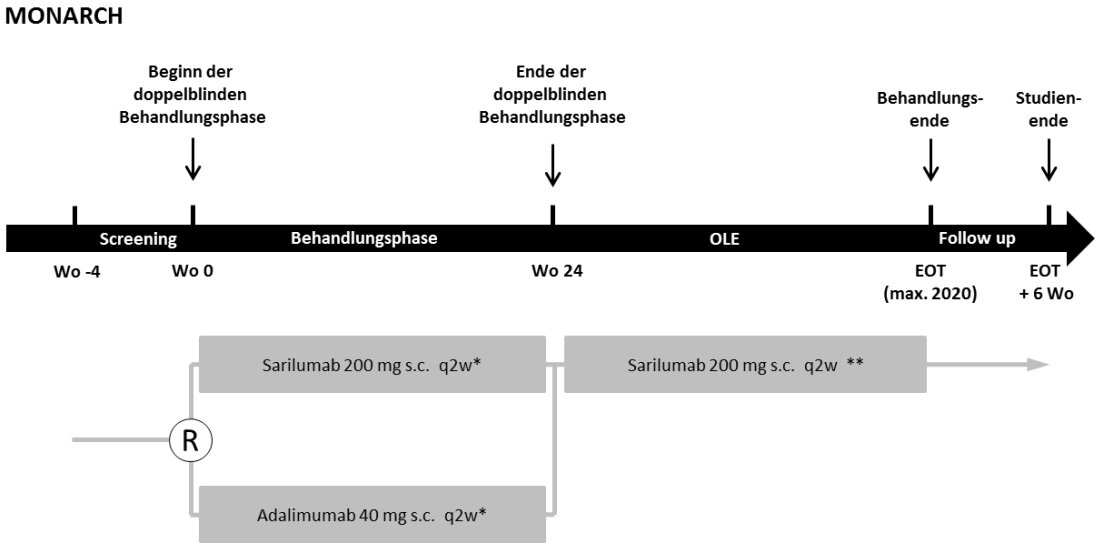
**Studie MONARCH**

Die Studie SARIL-RA-MONARCH (EFC14092, NCT02332590) – im Folgenden mit MONARCH bezeichnet – ist eine Phase-III-Zulassungsstudie, in der die Wirksamkeit und Sicherheit einer Sarilumab-Monotherapie im Vergleich zu einer Adalimumab-Monotherapie bei Patienten mit rheumatoider Arthritis untersucht wurde.

In die Studie wurden Patienten mit schwerer aktiver rheumatoider Arthritis[18] eingeschlossen, wobei ein DAS28-ESR-Wert > 5,1 als Schwellenwert für eine hohe Krankheitsaktivität galt. Die eingeschlossenen Patienten hatten im Mittel einen DAS28-ESR von 6,8 sowie 27 druckschmerzhafte und 18 geschwollene Gelenke (siehe Tabelle 4-17). Die überwiegende Mehrheit der Patienten wies Autoantikörper auf: etwa 66 % der Patienten war positiv für Rheumafaktoren und etwa 76 % waren positiv für anti-CCP-Antikörper (Tabelle 4-16).

Das durchschnittliche Alter der Patienten betrug 52,2 Jahre. Etwa 83 % der Patienten waren weiblich. Die Patienten wiesen entweder eine ärztlich festgestellte Intoleranz gegenüber MTX auf oder waren nach Einschätzung des Arztes für eine Fortführung der MTX-Therapie ungeeignet oder hatten auf eine mindestens 12-wöchige Behandlung mit MTX ein unzureichendes Ansprechen gezeigt.

Das Studiendesign ist schematisch in Abbildung 4-2 dargestellt.




* Ab Woche 16 konnten Patienten mit < 20% Verbesserung bezüglich TJC und SJC bei zwei aufeinanderfolgenden Visiten die Adalimumab-Dosis (bzw. Placebo in der Sarilumab-Gruppe) auf 40 mg qw erhöhen.
** In der OLE-Phase konnte bei Auftreten prädefinierter unerwünschter Ereignisse (Neutropenie, Thrombozytopenie, Erhöhung der Leberwerte) die Sarilumab-Dosis auf 150 mg s.c. q2w reduziert werden.

OLE: *open label extension*; EOT: *end of treatment*; s.c.: subkutan; q2w: alle zwei Wochen; qw: wöchentlich; TJC: *tender joint count*; SJC: *swollen joint count*

Abbildung 4-2: Schematische Darstellung des Studiendesigns der Studie MONARCH

Die Patienten wurden nach Region stratifiziert randomisiert und im Verhältnis 1:1 der Behandlung mit 200 mg Sarilumab oder 40 mg Adalimumab zugeteilt. Beide Wirkstoffe wurden beginnend in Woche 0 alle zwei Wochen mittels subkutaner Injektion verabreicht.

[18] Aktive rheumatoide Arthritis wurde definiert als: 68TJC ≥ 8 und 66SJC ≥ 6 sowie CRP ≥ 8 mg/L oder ESR ≥ 28 mm/h.

Die Patienten erhielten keine csDMARDs in Kombination mit Adalimumab oder Sarilumab, jedoch konnte eine bestehende orale Kortikosteroid-Therapie (≤ 10 mg Prednison-Äquivalent) oder eine Behandlung mit einem nichtsteroidalen Antirheumatikum (NSAID) beibehalten werden. Während der Studie erhielten etwa 55 % der Patienten Kortikosteroide sowie 65 % der Patienten NSAIDs als Hintergrundtherapie. Ab Woche 16 konnten die Patienten, bei denen in zwei aufeinanderfolgenden Visiten keine 20%ige Verbesserung bezüglich der Anzahl der geschwollenen und druckschmerzhaften Gelenke erreicht wurde, die Adalimumab-Dosis auf wöchentlich 40 mg erhöhen. Die Dosisanpassung geschah verblindet, das heißt, die Patienten des Sarilumab-Arms konnten entsprechend die Placebo-Dosis auf eine wöchentliche Gabe erhöhen.

Die doppelblinde, randomisierte Behandlungsphase wurde bis Woche 24 durchgeführt. Im Anschluss konnten alle Patienten in eine offene Extensionsphase eintreten, in der eine Behandlung mit 200 mg Sarilumab s.c. alle zwei Wochen fortgeführt wurde. In der offenen Extensionsphase konnte bei Auftreten von prädefinierten unerwünschten Ereignissen (Neutropenie, Thrombozytopenie oder Erhöhung der Leberwerte) die Sarilumab-Dosis auf 150 mg gesenkt werden.

***Studiendauer***

Die doppelblinde, randomisierte Behandlungsphase war nach 24 Wochen beendet. Dies stellt einen ausreichend langen Zeitraum dar, um die Wirksamkeit und Sicherheit von Sarilumab und Adalimumab im Vergleich zu beurteilen.

Die europäische Leitlinie der EULAR empfiehlt, wenn ein Patienten mit rheumatoider Arthritis nach 6 Monaten Behandlung mit einem DMARD das Therapieziel nicht erreicht, die Therapie anzupassen [3]. Patienten, die nach sechsmonatiger Behandlung ihr Therapieziel nicht erreicht haben, profitieren nicht davon, wenn die Behandlung unverändert fortgesetzt wird. Es gibt somit keine medizinische Rationale dafür, Patienten, die in einer Studie nach sechs Monaten Behandlung das Therapieziel Remission oder zumindest eine niedrige Krankheitsaktivität nicht erreicht haben, weiter auf dieser unzureichenden Therapie zu belassen. Auch die Fachinformation von Adalimumab gibt konkret vor, dass die Fortführung der Therapie sorgfältig zu überdenken ist, wenn ein Patient unter Adalimumab innerhalb von 12 Wochen kein Ansprechen auf die Therapie zeigt [18].

Ergänzend liegen für Sarilumab Daten aus der offenen Extensionsphase vor, die eine Beurteilung der Wirksamkeit und Sicherheit von Sarilumab nach insgesamt 48-wöchiger Behandlung ermöglichen. Von den 87 Patienten unter Sarilumab in der relevanten Studienpopulation (siehe Abschnitt „Für die Nutzenbewertung relevante Studienpopulation“ auf Seite 83) setzten 75 Patienten die Behandlung mit Sarilumab in der offenen Extensionsphase fort.

***Patientenpopulationen der Studie MONARCH***

Im Rahmen der Studie MONARCH wurden verschiedene Populationen für unterschiedliche Fragestellungen definiert. Diese Populationen sind im Folgenden näher beschrieben.

*Population zur Darstellung der Patientencharakteristika – Randomisierte Population (RP)*

Die RP umfasst alle randomisierten Patienten.

*Population zur Auswertung der Wirksamkeitsendpunkte – ITT(Intention-to-treat)-Population*

Die ITT-Population in der randomisierten doppelblinden Behandlungsphase umfasst alle randomisierten Patienten und ist somit mit der RP identisch. Patienten wurden ausgewertet wie randomisiert.

Die ITT-Population in der offenen Extensionsphase umfasst alle Patienten des Sarilumab-Arms der ursprünglichen ITT-Population, die nach Woche 24 in die Extensionsphase der Studie eintraten.

*Population zur Auswertung der Sicherheitsendpunkte – Safety-Population*

Die Safety-Population umfasst alle randomisierten Patienten, die mindestens eine Dosis der Studienintervention (Sarilumab oder Adalimumab) erhalten haben. Patienten wurden ausgewertet wie behandelt.

*Für die Nutzenbewertung relevante Studienpopulation*

Für die Nutzenbewertung von Sarilumab sind lediglich diejenigen Patienten relevant, bei denen eine Intoleranz gegenüber MTX vorlag oder die für eine Weiterbehandlung mit MTX nicht geeignet waren. Diese Patienten werden im Folgenden als **relevante Studienpopulation** bezeichnet. Es handelt sich dabei fast ausschließlich um Patienten mit einer ärztlich bestätigten MTX-Intoleranz (siehe Tabelle 4-18). Lediglich ein Patient der relevanten Studienpopulation gehört zur Gruppe der Patienten, die für eine weitere MTX-Behandlung nicht geeignet sind. Aus diesem Grund wird im Folgenden nicht zwischen diesen beiden Patientengruppen unterschieden.

Tabelle 4-18: Charakterisierung der Studienpopulation (MTX-Historie) – RCT mit dem zu bewertenden Arzneimittel

| **Studie**<br>**Studienpopulation** | **Adalimumab**<br>**(N = 185)** | **Sarilumab**<br>**(N = 184)** |
|---|---|---|
| MONARCH | | |
| Relevante Studienpopulation: | | |
| – Patienten mit MTX-Intoleranz | 81 (43,8%) | 87 (47,3%)[a] |
| – Patienten, die für eine MTX-Behandlung nicht geeignet sind | 1 (0,5%) | 0 |
| Für die Zusatznutzendarstellung nicht berücksichtigte Studienpopulation: | | |
| – Patienten mit unzureichendem Ansprechen auf MTX | 103 (55,7%) | 97 (52,7%) |

MTX: Methotrexat

[a] Von den 87 Patienten im Sarilumab-Arm brachen 12 Patienten innerhalb der 24-wöchigen doppelblinden, randomisierten Behandlungsphase die Behandlung ab (davon 7 Patienten aufgrund von UE). Dementsprechend setzten 75 Patienten die Behandlung mit Sarilumab in der offenen Extensionsphase fort.

Die Patienten der Studie MONARCH, die bei Einschluss in die Studie ein unzureichendes Ansprechen auf MTX zeigten, werden für die Nutzenbewertung von Sarilumab nicht herangezogen. Die Fachinformationen von Adalimumab und Sarilumab sehen eine generelle Kombination mit MTX vor. Nur im Falle einer MTX-Unverträglichkeit, oder wenn eine Behandlung mit MTX nicht angemessen ist, können Adalimumab und Sarilumab als Monotherapie eingesetzt werden [1, 18]. Dies entspricht auch den Leitlinien, die eine generelle Kombination von bDMARDs mit MTX empfehlen [3].

Zur Auswertung der relevanten Studienpopulation wurde eine Subgruppenanalyse nach MTX-Historie durchgeführt. Die Patientencharakteristika der relevanten Studienpopulation weisen keine relevanten Abweichungen von den Patientencharakteristika der Gesamtpopulation der Studie MONARCH auf (vgl. Tabelle 4-15 bis Tabelle 4-17). Die Patienten im Adalimumab-Arm weisen lediglich – wie schon in der Gesamtpopulation der Studie – eine etwas kürzere Erkrankungsdauer (7,6 vs. 9,4 Jahre) und im Mittel etwas höhere CRP-Werte (23,7 vs. 16,9 mg/L) auf. Alle anderen demografischen Charakteristika und Parameter der Krankheitsaktivität sind zwischen den Gruppen sehr ähnlich.

Die Patienten der relevanten Studienpopulation lassen sich den vom G-BA definierten Teilanwendungsgebieten 2 und 3 zuordnen (vgl. Tabelle 4-2).

Etwa 41 % der Patienten der relevanten Studienpopulation wurden vor Beginn der Studie lediglich mit MTX behandelt. Bei Patienten, die mit nur einem csDMARD vorbehandelt sind, trifft der G-BA eine Unterscheidung bezüglich des Vorliegens ungünstiger Prognosefaktoren. Alle Patienten der Studie MONARCH wiesen eine hohe Krankheitsaktivität gemäß DAS28 und damit mindestens einen ungünstigen Prognosefaktor auf. Zudem war die Mehrheit der Patienten Rheumafaktor- und/oder anti-CCP-Antikörper-positiv. Demzufolge erfüllen die Patienten das Kriterium „Vorliegen ungünstiger Prognosefaktoren“. Somit sind die Patienten, die lediglich mit MTX vorbehandelt waren, dem Teilanwendungsgebiet 2 zuzuordnen. Diese Patienten werden im Folgenden mit **Teilpopulation 2** bezeichnet.

Die restlichen 59 % der Patienten der relevanten Studienpopulation hatten vor Beginn der Studie neben MTX bereits andere csDMARDs erhalten und gehören somit zum Teilanwendungsgebiet 3. Diese Patienten werden im Folgenden mit **Teilpopulation 3** bezeichnet.

Die Patientencharakteristika der Teilpopulationen 2 und 3 weisen ebenfalls keine relevanten Abweichungen von den Patientencharakteristika der Gesamtpopulation der Studie MONARCH auf (vgl. Tabelle 4-15 bis Tabelle 4-17). Es zeigen sich leichte Verschiebungen bei der regionalen Zuordnung der Patienten im Vergleich zur Gesamtpopulation. Einzig hinsichtlich der Krankheitsdauer ergibt sich ein deutlicher Unterschied: die Patienten der Teilpopulation 2 weisen eine wesentlich kürzere Krankheitsdauer auf (6,3 Jahre im Sarilumab-Arm bzw. 3,9 Jahre im Adalimumab-Arm) als die Patienten der Teilpopulation 3 (11,4 bzw. 10,5 Jahre). Dies erklärt sich durch die post hoc vorgenommene Aufteilung der Patienten nach Anzahl der Vortherapien. Hinsichtlich der Parameter der Krankheitsaktivität

ergeben sich jedoch keine Unterschiede zur Gesamtpopulation oder der relevanten Studienpopulation.

***Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext***

Die Studie MONARCH wurde weltweit durchgeführt. Die Studienpopulation bestand zu etwa 90 % aus weißen (kaukasischen) Studienteilnehmern. Je nach ausgewerteter Teilpopulation stammt ein Drittel bis die Hälfte der Patienten aus westlichen Ländern (siehe Tabelle 4-15), bei denen von einem ähnlichen Standard der Gesundheitsversorgung wie in Deutschland ausgegangen werden kann. Zudem ergaben die durchgeführten Subgruppenanalysen in der Studie MONARCH keine Effektmodifikation durch den Faktor „Region“ (siehe Abschnitt 4.3.1.3.11). Auch die EMA bestätigte im Rahmen der Zulassung die externe Validität der Zulassungsstudien von Sarilumab im Hinblick auf die Vergleichbarkeit der europäischen und nicht-europäischen Patienten [89].

Die für die Nutzenbewertung relevante Studienpopulation umfasst Patienten mit einer schweren aktiven rheumatoiden Arthritis, die bereits mit mindestens einem csDMARD vorbehandelt wurden, prognostisch ungünstige Faktoren aufweisen und die eine Behandlung mit MTX nicht vertragen. Diese Patienten entsprechen dem Anwendungsgebiet für eine Sarilumab-Monotherapie [1]. Die Behandlung der relevanten Studienpopulation in MONARCH entspricht zudem den Empfehlungen der für Deutschland relevanten europäischen Therapie-Leitlinie [3].

Zusammenfassend ist somit die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext gegeben.

#### 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 4-19: 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 | | | |
| MONARCH | ja | ja | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Bei der Studie MONARCH handelt es sich um eine randomisierte, aktiv kontrollierte, doppelblinde Phase-III-Studie. Sowohl die Erzeugung der Randomisierungssequenz als auch die Verdeckung der Gruppenzuteilung wurde mittels IVRS realisiert. Patienten und behandelnde Personen waren verblindet. Die Erhebung und Auswertung der Endpunkte erfolgte verblindet. Eine ergebnisunabhängige Berichterstattung liegt vor. Änderungen des Studienprotokolls, die nach Studienbeginn vorgenommen wurden, beinhalten keine relevanten Parameter für das Verzerrungspotential. Abschließend wird das Verzerrungspotential der Studie MONARCH als niedrig eingestuft.

#### 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.*

Die in Tabelle 4-20 aufgelisteten Endpunkte wurden nach 24-wöchiger Behandlung mit Sarilumab oder Adalimumab erhoben. Zusätzlich wurden ausgewählte Endpunkte, auf denen die abschließende Bewertung des Zusatznutzens beruht, nach 48 Wochen Behandlung mit Sarilumab deskriptiv ausgewertet und im folgenden Abschnitt ergänzend dargestellt.

Tabelle 4-20: Patientenrelevante und ergänzend dargestellte Endpunkte der Studie MONARCH

| Kategorie | Endpunkt | Operationalisierung |
|---|---|---|
| **Morbidität** | Änderung der Krankheitsaktivität | Änderung des DAS28-ESR |
| | | Änderung des CDAI |
| | | Änderung des SDAI |
| | Remission | DAS28-ESR < 2,6 |
| | | CDAI ≤ 2,8 |
| | | SDAI ≤ 3,3 |
| | | Boolesche ACR/EULAR-Remission |
| | Niedrige Krankheitsaktivität | DAS28-ESR < 3,2 |
| | | CDAI ≤ 10 |
| | | SDAI ≤ 11 |
| | Therapieansprechen | Ansprechen gemäß EULAR-Kriterien (basierend auf DAS28-ESR) |
| | | Ansprechen gemäß ACR-Kriterien (ACR20/50/70) |
| | Einzelkomponenten der Scores zur Krankheitsaktivität | Gelenksymptome: Druckschmerzhafte Gelenke (68TJC + 28TJC) |
| | | Gelenksymptome: Geschwollene Gelenke (66SJC + 28SJC) |
| | | Schmerzen |
| | | Globale Einschätzung des Arztes (PhGA) |
| | | Globale Einschätzung des Patienten (PtGA) |
| | | Körperlicher Funktionsstatus – HAQ-DI |
| | | Laborwerte: CRP +ESR |
| | Morgensteifigkeit | Morgensteifigkeit-VAS |
| | Fatigue | FACIT-Fatigue |
| | Änderung des allgemeinen Gesundheitszustands | EQ-5D-VAS |

| | |
|---|---|
| **Lebensqualität** | Änderung der Lebensqualität - SF-36 |
| | Änderung der Lebensqualität - RAID |
| **Sicherheit/ Verträglichkeit** | Unerwünschte Ereignisse (UE)<br>Schwerwiegende unerwünschte Ereignisse (SUE)<br>Therapieabbrüche aufgrund unerwünschter Ereignisse<br>Unerwünschte Ereignisse mit Todesfolge<br>Unerwünschte Ereignisse von besonderem Interesse (UEBI) |

Lagen bei stetigen Effektmaßen statistisch signifikante Mittelwertdifferenzen vor, wurden standardisierte Mittelwertdifferenz in Form von Hedges' g zur Bewertung der klinischen Relevanz des statistisch signifikanten Gruppenunterschiedes berechnet.

Für binäre Endpunkte wurden als Effektschätzer das relative Risiko (RR, *Risk Ratio*), das *Odds Ratio* (OR) sowie die Risikodifferenz (RD) berechnet. Zur Bewertung des Behandlungseffekts im Rahmen der Zusatznutzenableitung wurden sofern möglich die RR herangezogen.

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

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

– Ergebnisse der ITT-Analyse
– Zahl der Patienten, die in die Analyse eingegangen sind
– dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
– bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
– bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler
– Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

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

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

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

#### 4.3.1.3.1 Morbidität: Änderung der Krankheitsaktivität – RCT

##### 4.3.1.3.1.1 Änderung des DAS28-ESR

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

Tabelle 4-21: Operationalisierung von „Änderung des DAS28-ESR“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Der DAS28 (*Disease Activity Score for 28 Joints*) ist ein Score zur Bestimmung der Krankheitsaktivität bei Patienten mit rheumatoider Arthritis.<br>Der Score berücksichtigt die folgenden vier Parameter:<br>– Anzahl der druckschmerzhaften Gelenke von 28 Gelenken (28TJC, *Tender Joint Count*): Die Anzahl der druckschmerzhaften Gelenke von insgesamt 68 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt. Für die Berechnung des DAS28 wurden 28 Gelenke[19] berücksichtigt.<br>– Anzahl der geschwollenen Gelenke von 28 Gelenken (28SJC, *Swollen Joint Count*): Die Anzahl der geschwollenen Gelenke von insgesamt 66 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt. Für die Berechnung des DAS28 wurden 28 Gelenke berücksichtigt (siehe Fußnote).<br>– Globale Einschätzung des Patienten (PtGA, *Patient Global Assessment*): Die Patienten schätzten ihre Krankheitsaktivität auf einer 100 mm-VAS ein, wobei 0 „keine Krankheitsaktivität“ und 100 „höchstmögliche Krankheitsaktivität“ bedeutet.<br>– Erythrozytensedimentationsrate (ESR): Die Bestimmung des Entzündungsparameters ESR (in mm/h) wurde lokal im jeweiligen Studienzentrum anhand einer Nüchternblutprobe durchgeführt.<br>Der DAS28-ESR wurde mittels folgender Formel berechnet:<br>$DAS28\text{-}ESR = 0{,}56 \times \sqrt{28TJC} + 0{,}28 \times \sqrt{28SJC} + 0{,}70 \times LN[\text{ESR (mm/h)}] + 0{,}014 \times PtGA(VAS)$<br>Der Score kann Werte zwischen 0 und 9,4 annehmen, wobei höhere Werte eine höhere Krankheitsaktivität bedeuten.<br>Der DAS28 wurde zu Beginn der Studie (Woche 0) sowie in der randomisierten Behandlungsphase in Woche 12 und Woche 24 sowie in der offenen Extensionsphase in Woche 36 und 48 erhoben. Die dargestellte Änderung des DAS28-ESR ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0.<br>Die Änderung des DAS28-ESR nach 24 Wochen im Vergleich zum Ausgangswert war der primäre Endpunkt der Studie MONARCH. |

*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.*

[19] Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke (Fingermittelgelenke, n=8), Kniegelenke (n=2)

Tabelle 4-22: Bewertung des Verzerrungspotenzials für „Änderung des DAS28-ESR“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Änderung des DAS28-ESR“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

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

Tabelle 4-23: Ergebnisse für „Änderung des DAS28-ESR (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **DAS28-ESR-Basiswert** | | **DAS28-ESR in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 6,87 (0,62) | 76 | 3,32 (1,35) | -3,55 (1,27) | -1,253 [-1,660; -0,846] < 0,0001 **-0,990 [-1,311; -0,668]** | |
| Adalimumab (N = 82) | 71 | 6,77 (0,81) | 71 | 4,66 (1,28) | -2,11 (1,35) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 6,87 (0,51) | 31 | 3,39 (1,35) | -3,47 (1,29) | -1,365 [-1,966; -0,763] < 0,0001 **-1,134 [-1,634; -0,634]** | 0,8371 |
| Adalimumab (N = 35) | 31 | 6,71 (0,75) | 31 | 4,74 (1,15) | -1,97 (1,04) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | |
| Sarilumab (N = 52) | 45 | 6,87 (0,69) | 45 | 3,27 (1,37) | -3,60 (1,27) | -1,179 [-1,744; -0,613] <0,0001 **-0,883 [-1,307; -0,459]** | |
| Adalimumab (N = 47) | 40 | 6,82 (0,87) | 40 | 4,60 (1,39) | -2,22 (1,55) | | |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 15 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Der Endpunkt „Änderung des DAS28-ESR (Woche 24)“ stellte den primären Endpunkt der Studie MONARCH dar.

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verminderung der Krankheitsaktivität gemessen mittels DAS28-ESR (siehe Tabelle 4-23).

In der relevanten Studienpopulation zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung des DAS28 (LSMD [95% KI] p-Wert: -1,253 [-1,660; -0,846] < 0,0001). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g belegt, dass es sich dabei um einen klinisch relevanten Effekt mit mittlerer Effektstärke handelt (Hedges´g [95% KI]: -0,990 [-1,311; -0,668]).

Bei den Patienten der Teilpopulationen 2 und 3 zeigte sich ebenfalls unter Sarilumab im Vergleich zu Adalimumab eine statistisch signifikant stärkere Verringerung des DAS28 (LSMD [95% KI] p-Wert in Teilpopulation 2: -1,365 [-1,966; -0,763] < 0,0001; LSMD [95% KI] p-Wert in Teilpopulation 3: -1,179 [-1,744; -0,613] < 0,0001). Dabei handelte es sich wie im Falle der relevanten Studienpopulation um klinisch relevante Effekte mit mittlerer (Teilpopulation 2, Hedges´g [95% KI]: -1,134 [-1,634; -0,634]) bzw. geringer Effektstärke (Teilpopulation 3, Hedges´g [95% KI]: -0,883 [-1,307; -0,459]).

Der Interaktionstest ergab keine relevante Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Änderung des DAS28-ESR (Woche 24)“ das Ergebnis der relevanten Studienpopulation betrachtet werden kann.

Von den 87 Patienten im Sarilumab-Arm brachen 12 Patienten innerhalb der 24-wöchigen doppelblinden, randomisierten Behandlungsphase die Behandlung ab (davon 7 Patienten aufgrund von UE). Dementsprechend setzten 75 Patienten die Behandlung mit Sarilumab in der offenen Extensionsphase fort.

Die Auswertung des DAS28-ESR nach 48 Wochen Behandlung mit Sarilumab zeigte die langanhaltende Verringerung der Krankheitsaktivität durch Sarilumab (siehe Tabelle 4-24). Der mittlere Behandlungseffekt blieb in allen betrachteten Teilpopulationen über 48 Wochen bestehen. Im Vergleich zur mittleren Krankheitsaktivität nach 24 Wochen Behandlung kam es nach 48 Wochen zu einer weiteren leichten Reduktion der Krankheitsaktivität gemessen mittels DAS28-ESR.

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

Tabelle 4-24: Ergebnisse für „Änderung des DAS28-ESR (Woche 48)“

| **Studie MONARCH** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Behandlung** | **DAS28-ESR-Basiswert[a]** | | **DAS28-ESR in Woche 48[b]** | | **absolute Änderung im Vergleich zum Basiswert** | | **absolute Änderung von Woche 24 zu Woche 48** | |
| | N | MW (SD) | N | MW (SD) | N | MD (SD) | N | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | | |
| Sarilumab | 75 | 6,87 (0,62) | 71 | 2,78 (1,31) | 71 | -4,09 (1,31) | 71 | -0,48 (1,15) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | | |
| Sarilumab | 30 | 6,88 (0,51) | 30 | 2,71 (1,32) | 30 | -4,17 (1,33) | 30 | -0,62 (1,22) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | | |
| Sarilumab | 45 | 6,87 (0,69) | 41 | 2,83 (1,32) | 41 | -4,02 (1,31) | 41 | -0,39 (1,10) |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate; MW: Mittelwert; MD: Mittelwertdifferenz; SD: Standardabweichung

[a] Dargestellt ist der DAS28-ESR-Mittelwert zu Studienbeginn für die Patienten, die nach Woche 24 an der offenen Extensionphase teilnahmen.

[b] Dargestellt ist der DAS28-ESR-Mittelwert zu Woche 48 für die Patienten, die die 48-wöchige Behandlung beendet haben.

Quelle: [90]

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Änderung des DAS28-ESR“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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

#### 4.3.1.3.1.2 Änderung des CDAI

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

Tabelle 4-25: Operationalisierung von „Änderung des CDAI“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Der CDAI (*Clinical Disease Activity Index*) ist ein Score zur Bestimmung der Krankheitsaktivität bei Patienten mit rheumatoider Arthritis.<br>Der Score berücksichtigt die folgenden vier Parameter:<br>– Anzahl der druckschmerzhaften Gelenke von 28 Gelenken (28TJC, *Tender Joint Count*): Die Anzahl der druckschmerzhaften Gelenke von insgesamt 68 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt. Für die Beurteilung wurden 28 Gelenke[20] berücksichtigt.<br>– Anzahl der geschwollenen Gelenke von 28 Gelenken (28SJC, *Swollen Joint Count*): Die Anzahl der geschwollenen Gelenke von insgesamt 66 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt. Für die Beurteilung wurden 28 Gelenke berücksichtigt (siehe Fußnote).<br>– Globale Einschätzung des Patienten (PtGA, *Patient Global Assessment*): Die Patienten schätzten ihre Krankheitsaktivität auf einer 100 mm-VAS ein, wobei 0 „keine Krankheitsaktivität“ und 100 „höchstmögliche Krankheitsaktivität“ bedeutet.<br>– Globale Einschätzung des Arztes (PhGA, *Physicians Global Assessment*): Der behandelnde Arzt schätzte die Krankheitsaktivität des Patienten auf einer 100 mm-VAS ein, wobei 0 „keine Krankheitsaktivität“ und 100 „höchstmögliche Krankheitsaktivität“ bedeutet.<br>Der Score ergibt sich durch Addition der vier Parameter 28TJC, 28SJC, PtGA (in cm) und PhGA (in cm) und kann Werte von 0 bis 76 annehmen, wobei höhere Werte eine höhere Krankheitsaktivität bedeuten.<br>Die Parameter zur Berechnung des CDAI wurden zu Beginn der Studie (Woche 0) und anschließend alle vier Wochen in der randomisierten Behandlungsphase sowie in der offenen Extensionsphase erhoben. Die dargestellte Änderung des CDAI ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. |

*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.*

[20] Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke (Fingermittelgelenke, n=8), Kniegelenke (n=2)

Tabelle 4-26: Bewertung des Verzerrungspotenzials für „Änderung des CDAI“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Änderung des CDAI“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

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

Tabelle 4-27: Ergebnisse für „Änderung des CDAI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **CDAI-Basiswert** | | **CDAI in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 43,52 (10,87) | 76 | 12,36 (9,44) | -31,16 (10,65) | -5,005 [-7,998; -2,011] 0,0012 **-0,536 [-0,856; -0,215]** | |
| Adalimumab (N = 82) | 70 | 43,04 (12,19) | 70 | 17,87 (9,83) | -25,17 (13,97) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 44,05 (10,17) | 31 | 12,07 (8,63) | -31,97 (9,55) | -6,136 [-10,517; -1,756] 0,0068 **-0,692 [-1,186; -0,198]** | |
| Adalimumab (N = 35) | 31 | 42,60 (12,16) | 31 | 18,39 (9,70) | -24,21 (12,79) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,5023 |
| Sarilumab (N = 52) | 45 | 43,16 (11,43) | 45 | 12,56 (10,06) | -30,59 (11,42) | -3,944 [-8,120; 0,232] 0,0639 | |
| Adalimumab (N = 47) | 39 | 43,39 (12,36) | 39 | 17,45 (10,04) | -25,94 (14,97) | | |

CDAI: *Clinical Disease Activity Index*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Baseline; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 31 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verminderung der Krankheitsaktivität gemessen mittels CDAI (siehe Tabelle 4-27).

In der relevanten Studienpopulation zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung des CDAI (LSMD [95% KI] p-Wert: -5,005 [-7,998; -2,011] 0,0012). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g belegt, dass es sich dabei um einen klinisch relevanten Effekt mit kleiner Effektstärke handelt (Hedges´g [95% KI]: -0,536 [-0,856; -0,215]).

Bei den Patienten der Teilpopulation 2 zeigte sich unter Sarilumab im Vergleich zu Adalimumab ebenfalls eine statistisch signifikante Verringerung des CDAI (LSMD [95% KI] p-Wert: -6,136 [-10,517; -1,756] 0,0068). Das Konfidenzintervall zu Hedges´g schließt jedoch den Wert -0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt.

Bei den Patienten der Teilpopulation 3 zeigte sich ebenfalls eine stärkere Verringerung des CDAI unter Sarilumab im Vergleich zu Adalimumab, der Unterschied war jedoch statistisch nicht signifikant (LSMD [95% KI] p-Wert: -3,944 [-8,120; 0,232] 0,0639).

Der Interaktionstest ergab keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Änderung des CDAI (Woche 24)“ das Ergebnis der relevanten Studienpopulation betrachtet werden kann.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Änderung des CDAI“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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

#### 4.3.1.3.1.3 Änderung des SDAI

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

Tabelle 4-28: Operationalisierung von „Änderung des SDAI“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Der SDAI (*Simplified Disease Activity Index*) ist ein Score zur Bestimmung der Krankheitsaktivität bei Patienten mit rheumatoider Arthritis.<br>Der Score berücksichtigt neben den vier Parameter des CDAI – 28TJC, 28SJC, PtGA und PhGA (siehe Tabelle 4-25) – zusätzlich folgenden Parameter:<br>– C-reaktives Protein (CRP): Die Bestimmung des CRP-Werts als Entzündungsparameter wurde anhand einer Nüchternblutprobe in einem Zentrallabor durchgeführt.<br>Der Score ergibt sich durch Addition der fünf Parameter 28TJC, 28SJC, PtGA (in cm), PhGA (in cm) und CRP (0 - 10 mg/dL) und kann Werte von 0 bis 86 annehmen, wobei höhere Werte eine höhere Krankheitsaktivität bedeuten.<br>Der Auswertung des SDAI wurde post hoc vorgenommen. Die Einzelkomponenten wurden zu Beginn der Studie (Woche 0) und anschließend alle vier Wochen in der randomisierten Behandlungsphase sowie in der offenen Extensionsphase erhoben.<br>Die dargestellte Änderung des SDAI ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. |

*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 4-29: Bewertung des Verzerrungspotenzials für „Änderung des SDAI“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Änderung des SDAI“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat

umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-30: Ergebnisse für „Änderung des SDAI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **SDAI-Basiswert** | | **SDAI in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 74 | 45,02 (10,56) | 74 | 12,26 (9,42) | -32,76 (10,81) | -5,922 [-9,116; -2,727] 0,0003 **-0,596 [-0,917; -0,274]** | |
| Adalimumab (N = 82) | 68 | 45,28 (13,06) | 68 | 19,14 (10,23) | -26,14 (14,18) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 30 | 45,36 (8,85) | 30 | 11,64 (8,30) | -33,72 (9,83) | -6,762 [-11,682; -1,842] 0,0078 **-0,679 [-1,174;-0,185]** | |
| Adalimumab (N = 35) | 30 | 44,18 (13,34) | 30 | 19,45 (10,67) | -24,73 (12,76) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,4959 |
| Sarilumab (N = 52) | 44 | 44,79 (11,68) | 44 | 12,68 (10,18) | -32,11 (11,49) | -5,088 [-9,424;-0,752] 0,0220 **-0,501 [-0,928; -0,074]** | |
| Adalimumab (N = 47) | 38 | 46,14 (12,94) | 38 | 18,89 (10,00) | -27,25 (15,28) | | |

SDAI: *Simplified Disease Activity Index*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Baseline; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verminderung der Krankheitsaktivität gemessen mittels SDAI (siehe Tabelle 4-30).

In der relevanten Studienpopulation zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung des SDAI (LSMD [95% KI] p-Wert: -5,922 [-9,116; -2,727] 0,0003). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g belegt, dass es sich dabei um einen klinisch relevanten Effekt mit kleiner Effektstärke handelt (Hedges´g [95% KI]: -0,596 [-0,917; -0,274]).

Bei den Patienten der Teilpopulationen 2 und 3 zeigte sich unter Sarilumab im Vergleich zu Adalimumab ebenfalls eine statistisch signifikante Verringerung des SDAI (LSMD [95% KI] p-Wert in Teilpopulation 2: -6,762 [-11,682; -1,842] 0,0078 und LSMD [95% KI] p-Wert in Teilpopulation 3: -5,088 [-9,424;-0,752] 0,0220). Die Konfidenzintervalle zu Hedges´g schließen jedoch in beiden Gruppen den Wert -0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt.

Der Interaktionstest ergab keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Änderung des SDAI (Woche 24)“ das Ergebnis der relevanten Studienpopulation betrachtet werden kann.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Änderung des SDAI“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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 Morbidität: Remission – RCT

##### 4.3.1.3.2.1 DAS28-Remission

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

Tabelle 4-31: Operationalisierung von „DAS28-Remission“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die DAS28-Remission ist als Erreichen eines DAS28-Werts < 2,6 definiert.<br>Für die Auswertung der DAS28-Remission wurde der DAS28-ESR herangezogen. Die Operationalisierung des DAS28-ESR ist in Tabelle 4-21 dargestellt.<br>Der Anteil der Patienten mit DAS28-ESR < 2,6 wurde in Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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 4-32: Bewertung des Verzerrungspotenzials für „DAS28-Remission“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „DAS28-Remission“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-33: Ergebnisse für „DAS28-Remission (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit DAS28-ESR < 2,6 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 26 (29,9) | 6,20 [2,21; 17,41] **0,0006** | 4,39 [1,78; 10,85] **0,0014** | 0,23 [0,09; 0,37] **0,0010** | |
| Adalimumab (N = 82) | 5 (6,1) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,8593 RR: 0,7824 RD: 0,9562 |
| Sarilumab (N = 35) | 10 (28,6) | 6,74 [1,33; 34,26] **0,0217** | 4,98 [1,17; 21,27] **0,0302** | 0,23 [0,04; 0,42] **0,0190** | |
| Adalimumab (N = 35) | 2 (5,7) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 16 (30,8) | 5,58 [1,44; 21,57] **0,0131** | 3,84 [1,20; 12,27] **0,0233** | 0,22 [0,02; 0,42] **0,0292** | |
| Adalimumab (N = 47) | 3 (6,4) | | | | |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate; ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 20 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24-wöchiger Behandlung zeigte sich ein sehr deutlicher Vorteil von Sarilumab gegenüber Adalimumab bei der Anzahl der Patienten, die eine Remission gemäß DAS28-ESR erreichten.

In der relevanten Studienpopulation erzielten 29,9 % der Patienten unter Sarilumab einen DAS28-ESR < 2,6, wohingegen nur 6,1 % der Patienten unter Adalimumab diesen Schwellenwert unterschritten (siehe Tabelle 4-33). Dieser Gruppenunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 4,39 [1,78; 10,85] 0,0014).

In den Teilpopulationen 2 und 3 zeigten sich ebenfalls vergleichbare Ergebnisse (RR [95 % KI] p-Wert: 4,98 [1,17; 21,27] 0,0302 in Teilpopulation 2 bzw. 3,84 [1,20; 12,27] 0,0233 in Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „DAS28-Remission (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-34: Ergebnisse für „DAS28-Remission (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit DAS28-ESR < 2,6 in Woche 24 (ITT)** | | **Patienten mit DAS28-ESR < 2,6 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 26 (29,9) | 75 | 33 (44,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 10 (28,6) | 30 | 16 (53,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 16 (30,8) | 45 | 17 (37,8) |
| DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate<br>Patienten mit fehlenden Werten wurden als Non-Responder gezählt.<br>Quelle: [90] | | | | |

Die Auswertung des Endpunktes „DAS28-Remission“ nach 48 Wochen Behandlung mit Sarilumab zeigte im Vergleich zur Auswertung nach 24 Wochen einen weiteren Anstieg der Anzahl der Patienten, die die Remission gemäß DAS28-ESR erreichten (siehe Tabelle 4-34).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „DAS28-Remission“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.2.2 CDAI-Remission

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

Tabelle 4-35: Operationalisierung von „CDAI-Remission“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die CDAI-Remission ist als Erreichen eines CDAI-Werts ≤ 2,8 definiert.<br>Die Operationalisierung des CDAI ist in Tabelle 4-25 dargestellt.<br>Der Anteil der Patienten mit CDAI ≤ 2,8 wurde in Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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 4-36: Bewertung des Verzerrungspotenzials für „CDAI-Remission“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „CDAI-Remission“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

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

Tabelle 4-37: Ergebnisse für „CDAI-Remission (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit CDAI ≤ 2,8 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR[b] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 8 (9,2) | 3,73 [0,75; 18,44] 0,1062 | 3,37 [0,73; 15,45] 0,1182 | 0,06 [-0,06; 0,19] 0,3274 | |
| Adalimumab (N = 82) | 2 (2,4) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: 0,3739 RR: n.a. RD: 0,8476 |
| Sarilumab (N = 35) | 3 (8,6) | Peto-OR: 7,84 [0,79; 77,98] 0,0788 | n.a. | 0,14 [-0,78; 1,07] 0,7583 | |
| Adalimumab (N = 35) | 0 | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 5 (9,6) | OR: 1,80 [0,31; 10,33] 0,5085 Peto-OR: 2,24 [0,49; 10,37] 0,3013 | 1,65 [0,34; 7,87] 0,5293 | 0,05 [-0,15; 0,26] 0,6140 | |
| Adalimumab (N = 47) | 2 (4,3) | | | | |

CDAI: *Clinical Disease Activity Index*; ITT: Intent-to treat; KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

[b] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 32 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24-wöchiger Behandlung zeigte sich ein numerischer Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Anzahl der Patienten, die eine Remission gemäß CDAI erreichten.

In der relevanten Studienpopulation erzielten 9,2 % der Patienten unter Sarilumab einen CDAI ≤ 2,8, wohingegen nur 2,4 % der Patienten unter Adalimumab diesen Schwellenwert unterschritten (siehe Tabelle 4-37). Dieser numerische Unterschied ist jedoch statistisch nicht signifikant (RR [95 % KI] p-Wert: 3,37 [0,73; 15,45] 0,1182).

In den Teilpopulationen 2 und 3 zeigte sich ebenfalls kein statistisch signifikanter Unterschied (Peto-OR [95 % KI] p-Wert: 7,84 [0,79; 77,98] 0,0788 für Teilpopulation 2 bzw. 2,24 [0,49; 10,37] 0,3013 in Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „CDAI-Remission (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-38: Ergebnisse für „CDAI-Remission (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit CDAI ≤ 2,8 in Woche 24 (ITT)** | | **Patienten mit CDAI ≤ 2,8 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 8 (9,2) | 75 | 19 (25,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 3 (8,6) | 30 | 8 (26,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 5 (9,6) | 45 | 11 (24,4) |

CDAI: *Clinical Disease Activity Index*
Patienten mit fehlenden Werten wurden als Non-Responder gezählt.
Quelle: [90]

Die Auswertung des Endpunktes „CDAI-Remission“ nach 48 Wochen Behandlung mit Sarilumab zeigte im Vergleich zur Auswertung nach 24 Wochen einen weiteren Anstieg der Anzahl der Patienten, die die Remission gemäß CDAI erreichten (siehe Tabelle 4-38).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „CDAI-Remission“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.2.3 SDAI-Remission

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

Tabelle 4-39: Operationalisierung von „SDAI-Remission“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die SDAI-Remission ist als Erreichen eines SDAI-Werts ≤ 3,3 definiert.<br>Die Operationalisierung des SDAI ist in Tabelle 4-28 dargestellt.<br>Der Anteil der Patienten mit SDAI ≤ 3,3 wurde post hoc für Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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 4-40: Bewertung des Verzerrungspotenzials für „SDAI-Remission“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „SDAI-Remission“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-41: Ergebnisse für „SDAI-Remission (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit SDAI ≤ 3,3 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR[b] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 9 (10,3) | 4,25 [0,87; 20,70] 0,0732 | 3,75 [0,83; 16,88] 0,0844 | 0,07 [-0,07; 0,21] 0,3182 | |
| Adalimumab (N = 82) | 2 (2,4) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: 0,3186 RR: n.a. RD: 0,4997 |
| Sarilumab (N = 35) | 4 (11,4) | Peto-OR: 8,09 [1,09; 60,03] **0,0409** | n.a. | 0,15 [-0,05; 0,36] 0,1429 | |
| Adalimumab (N = 35) | 0 | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 5 (9,6) | OR: 1,80 [0,31; 10,33] 0,5085 Peto-OR: 2,24 [0,49; 10,37] 0,3013 | 1,65 [0,34; 7,87] 0,5293 | 0,05 [-0,15; 0,26] 0,6055 | |
| Adalimumab (N = 47) | 2 (4,3) | | | | |

SDAI: *Simplified Disease Activity Index*; ITT: Intent-to treat; KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

[b] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

Quelle: [72]

Nach 24-wöchiger Behandlung zeigte sich ein numerischer Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Anzahl der Patienten, die eine Remission gemäß SDAI erreichten.

In der relevanten Studienpopulation erzielten 10,3 % der Patienten unter Sarilumab einen SDAI ≤ 3,3, wohingegen nur 2,4 % der Patienten unter Adalimumab diesen Schwellenwert unterschritten (siehe Tabelle 4-41). Dieser numerische Unterschied ist jedoch statistisch nicht signifikant (RR [95 % KI] p-Wert: 3,75 [0,83; 16,88] 0,0844).

In der Teilpopulation 2 zeigten sich vergleichbare Ergebnisse. In der Sarilumab-Gruppe erreichten 11,4 % der Patienten eine Remission gemäß SDAI, während kein Patient unter Adalimumab eine SDAI-Remission erreichte. Der Gruppenunterschied ist statistisch signifikant (Peto-OR [95 % KI] p-Wert: 8,09 [1,09; 60,03] 0,0409).

In der Teilpopulation 3 zeigte sich ebenfalls ein numerischer Vorteil für Sarilumab bezüglich des Erreichens der SDAI-Remission, wobei der Gruppenunterschied statistisch nicht signifikant war (Peto-OR [95 % KI] p-Wert: 2,24 [0,49; 10,37] 0,3013).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „SDAI-Remission (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-42: Ergebnisse für „SDAI-Remission (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit SDAI ≤ 3,3 in Woche 24 (ITT)** | | **Patienten mit SDAI ≤ 3,3 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 9 (10,3) | 75 | 19 (25,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 4 (11,4) | 30 | 7 (23,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 5 (9,6) | 45 | 12 (26,7) |
| SDAI: *Simplified Disease Activity Index*<br>Patienten mit fehlenden Werten wurden als Non-Responder gezählt.<br>Quelle: [90] | | | | |

Die Auswertung des Endpunktes „SDAI-Remission“ nach 48 Wochen Behandlung mit Sarilumab zeigte im Vergleich zur Auswertung nach 24 Wochen einen weiteren Anstieg der Anzahl der Patienten, die die Remission gemäß SDAI erreichten (siehe Tabelle 4-42).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „SDAI-Remission“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.2.4 Boolesche ACR/EULAR-Remission

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

Tabelle 4-43: Operationalisierung von „Boolesche ACR/EULAR-Remission“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Zur Auswertung der Patienten mit Remission gemäß ACR/EULAR-Kriterien wurde folgende Parameter herangezogen:<br>– Anzahl der druckschmerzhaften Gelenke von 28 Gelenken (28TJC, *Tender Joint Count*): Die Anzahl der druckschmerzhaften Gelenke von insgesamt 68 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt. Für die Beurteilung wurden 28 Gelenke[21] berücksichtigt.<br>– Anzahl der geschwollenen Gelenke von 28 Gelenken (28SJC, *Swollen Joint Count*): Die Anzahl der geschwollenen Gelenke von insgesamt 66 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt. Für die Beurteilung wurden 28 Gelenke berücksichtigt (siehe Fußnote).<br>– CRP: Die Bestimmung des CRP-Werts als Entzündungsparameter wurde anhand einer Nüchternblutprobe in einem Zentrallabor durchgeführt.<br>– Globale Einschätzung des Patienten (PtGA): Die Patienten schätzten ihre Krankheitsaktivität auf einer 100 mm-VAS ein, wobei 0 „keine Krankheitsaktivität“ und 100 „höchstmögliche Krankheitsaktivität“ bedeutet.<br>Die boolesche ACR/EULAR-Remission ist wie folgt definiert:<br>– Anzahl der druckschmerzhaften Gelenke (28TJC) ≤ 1<br>– Anzahl der geschwollenen Gelenke (28SJC) ≤ 1<br>– CRP ≤ 1 mg/dL<br>– Globale Einschätzung des Patienten (PtGA 0-10 cm) ≤ 1<br>Die Patienten müssen jedes dieser Kriterien erfüllen, um als Responder zu gelten (Boole-basierte Definition).<br>Der Anteil der Patienten mit ACR/EULAR-Remission wurde post hoc für Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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.*

[21] Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke (Fingermittelgelenke, n=8), Kniegelenke (n=2)

Tabelle 4-44: Bewertung des Verzerrungspotenzials für „Boolesche ACR/EULAR-Remission“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Boolesche ACR/EULAR-Remission“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-45: Ergebnisse für „Boolesche ACR/EULAR-Remission (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit ACR/EULAR-Remission in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 5 (5,7) | 2,30 [0,43; 12,38] 0,3315 | 2,20 [0,43; 11,16] 0,3423 | 0,04 [-0,07; 0,15] 0,4259 | |
| Adalimumab (N = 82) | 2 (2,4) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,9856 RR: 0,9950 RD: 0,9944 |
| Sarilumab (N = 35) | 2 (5,7) | 2,08 [0,17; 24,89] 0,5608 | 2,06 [0,20; 21,78] 0,5453 | 0,04 [-0,39; 0,47] 0,8497 | |
| Adalimumab (N = 35) | 1 (2,9) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 3 (5,8) | 2,15 [0,21; 22,25] 0,5198 | 2,04 [0,22; 18,94] 0,5281 | 0,04 [-0,10; 0,18] 0,5834 | |
| Adalimumab (N = 47) | 1 (2,1) | | | | |

ACR: *American College of Rheumatology*; EULAR: *European League Against Rheumatism*; ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Nach 24-wöchiger Behandlung zeigte sich ein leichter numerischer Vorteil von Sarilumab gegenüber Adalimumab bezüglich des Erreichens einer Remission gemäß der Boole-basierten ACR/EULAR-Definition.

In der relevanten Studienpopulation erfüllten nach 24 Wochen Behandlung 5,7 % der Patienten unter Sarilumab und 2,4 % der Patienten unter Adalimumab die ACR/EULAR-Remissionskriterien (siehe Tabelle 4-45). Dieser Gruppenunterschied ist statistisch nicht signifikant (RR [95 % KI] p-Wert: 2,20 [0,43; 11,16] 0,3423).

In den Teilpopulationen 2 und 3 zeigten sich vergleichbare Ergebnisse (RR [95 % KI] p-Wert: 2,06 [0,20; 21,78] 0,5453 für Teilpopulation 2 bzw. 2,04 [0,22; 18,94] 0,5281 für Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „ACR/EULAR-Remission (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-46: Ergebnisse für „Boolesche ACR/EULAR-Remission (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit ACR/EULAR-Remission in Woche 24 (ITT)** | | **Patienten mit ACR/EULAR-Remission in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 5 (5,7) | 75 | 13 (17,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 2 (5,7) | 30 | 6 (20,0) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 3 (5,8) | 45 | 7 (15,6) |
| Patienten mit fehlenden Werten wurden als Non-Responder gezählt.<br>Quelle: [90] | | | | |

Die Auswertung des Endpunktes „Boolesche ACR/EULAR-Remission“ nach 48 Wochen Behandlung mit Sarilumab zeigte im Vergleich zur Auswertung nach 24 Wochen einen weiteren Anstieg der Anzahl der Patienten, die die Remission gemäß ACR/EULAR-Kriterien erreichen (siehe Tabelle 4-46).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Boolesche ACR/EULAR-Remission" sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.3 Morbidität: Niedrige Krankheitsaktivität – RCT

##### 4.3.1.3.3.1 Niedrige Krankheitsaktivität gemäß DAS28

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

Tabelle 4-47: Operationalisierung von „Niedrige Krankheitsaktivität gemäß DAS28"

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die niedrige Krankheitsaktivität gemäß DAS28 ist als Erreichen eines DAS28-Werts < 3,2 definiert.<br>Für die Auswertung wurde der DAS28-ESR herangezogen. Die Operationalisierung des DAS28-ESR ist in Tabelle 4-21 dargestellt.<br>Der Anteil der Patienten mit DAS28-ESR < 3,2 wurde in Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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 4-48: Bewertung des Verzerrungspotenzials für „Niedrige Krankheitsaktivität gemäß DAS28“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Niedrige Krankheitsaktivität gemäß DAS28“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-49: Ergebnisse für „Niedrige Krankheitsaktivität gemäß DAS28 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit DAS28-ESR < 3,2 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 39 (44,8) | 7,40 [3,11; 17,59] **< 0,0001** | 4,22 [2,10; 8,46] **< 0,0001** | 0,32 [0,19; 0,45] **< 0,0001** | |
| Adalimumab (N = 82) | 8 (9,8) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,6501 RR: 0,5824 RD: 0,4491 |
| Sarilumab (N = 35) | 16 (45,7) | 9,23 [2,33; 36,64] **0,0018** | 5,27 [1,67; 16,63] **0,0048** | 0,37 [0,17; 0,58] **0,0004** | |
| Adalimumab (N = 35) | 3 (8,6) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 23 (44,2) | 6,12 [1,98; 18,95] **0,0018** | 3,52 [1,47; 8,47] **0,0052** | 0,27 [0,09; 0,45] **0,0033** | |
| Adalimumab (N = 47) | 5 (10,6) | | | | |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate; ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 22 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Für Patienten mit hoher Krankheitsaktivität, wie sie in die Studie MONARCH eingeschlossen wurden, stellt das Erreichen einer niedrigen Krankheitsaktivität ein relevantes Therapieziel dar. Nach 24-wöchiger Behandlung zeigte sich ein deutlicher Vorteil für Sarilumab gegenüber Adalimumab bei der Zahl der Patienten, die eine niedrige Krankheitsaktivität gemäß DAS28 erreichten.

In der relevanten Studienpopulation erzielten nach 24 Wochen Behandlung 44,8 % der mit Sarilumab behandelten Patienten und 9,8 % der mit Adalimumab behandelten Patienten einen DAS28-ESR < 3,2 (siehe Tabelle 4-49). Dieser Behandlungsunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 4,22 [2,10; 8,46] < 0,0001). Dieses Ergebnis wird durch die nachfolgend dargestellten Auswertungen zur niedrigen Krankheitsaktivität gemäß CDAI und SDAI in der Größenordnung bestätigt.

Die Ergebnisse in den Teilpopulation 2 und 3 waren denen der relevanten Studienpopulation sehr ähnlich, in beiden Teilpopulationen zeigte sich ein signifikanter Vorteil für Sarilumab (RR [95 % KI] p-Wert: 5,27 [1,67; 16,63] 0,0048 in der Teilpopulation 2 bzw. 3,52 [1,47; 8,47] 0,0052 in der Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Niedrige Krankheitsaktivität gemäß DAS28 (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-50: Ergebnisse für „Niedrige Krankheitsaktivität gemäß DAS28 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit DAS28-ESR < 3,2 in Woche 24 (ITT)** | | **Patienten mit DAS28-ESR < 3,2 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 39 (44,8) | 75 | 44 (58,7) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 16 (45,7) | 30 | 17 (56,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 23 (44,2) | 45 | 27 (60,0) |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate
Patienten mit fehlenden Werten wurden als Non-Responder gezählt.
Quelle: [90]

Die Auswertung des Endpunktes „Niedrige Krankheitsaktivität gemäß DAS28“ nach 48 Wochen Behandlung mit Sarilumab zeigte im Vergleich zur Auswertung nach 24 Wochen einen weiteren Anstieg der Anzahl der Patienten, die eine niedrige Krankheitsaktivität gemäß DAS28-ESR erreichen (siehe Tabelle 4-50).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Niedrige Krankheitsaktivität gemäß DAS28“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.3.2 Niedrige Krankheitsaktivität gemäß CDAI

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

Tabelle 4-51: Operationalisierung von „Niedrige Krankheitsaktivität gemäß CDAI“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die niedrige Krankheitsaktivität gemäß CDAI ist als Erreichen eines CDAI-Werts ≤ 10 definiert.<br>Die Operationalisierung des CDAI ist in Tabelle 4-25 dargestellt.<br>Der Anteil der Patienten mit CDAI ≤ 10 wurde post hoc für Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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 4-52: Bewertung des Verzerrungspotenzials für „Niedrige Krankheitsaktivität gemäß CDAI“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des hier berichteten, nicht präspezifizierten Endpunktes „Niedrige Krankheitsaktivität gemäß CDAI“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-53: Ergebnisse für „Niedrige Krankheitsaktivität gemäß CDAI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit CDAI ≤ 10 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 39 (44,8) | 3,67 [1,78; 7,57] **0,0005** | 2,39 [1,44; 3,97] **0,0008** | 0,24 [0,11; 0,38] **0,0004** | |
| Adalimumab (N = 82) | 15 (18,3) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,6852 RR: 0,6921 RD: 0,6752 |
| Sarilumab (N = 35) | 16 (45,7) | 4,32 [1,38; 13,51] **0,0122** | 2,68 [1,20; 5,99] **0,0168** | 0,27 [0,06; 0,48] **0,0107** | |
| Adalimumab (N = 35) | 6 (17,1) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 23 (44,2) | 3,18 [1,23; 8,25] **0,0174** | 2,17 [1,13; 4,18] **0,0204** | 0,21 [0,03; 0,39] **0,0201** | |
| Adalimumab (N = 47) | 9 (19,1) | | | | |

CDAI: *Clinical Disease Activity Index*; ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Nach 24-wöchiger Behandlung zeigte sich auch bei der Zahl der Patienten, die eine niedrige Krankheitsaktivität gemäß CDAI erreichten, ein deutlicher Vorteil für Sarilumab gegenüber Adalimumab.

In der relevanten Studienpopulation erzielten nach 24 Wochen Behandlung 44,8 % der mit Sarilumab behandelten Patienten und 18,3 % der mit Adalimumab behandelten Patienten einen CDAI ≤ 10 (siehe Tabelle 4-53). Dieser Behandlungsunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 2,39 [1,44; 3,97] 0,0008). Dieses Ergebnis wird durch die

Auswertungen zur niedrigen Krankheitsaktivität gemäß DAS28-ESR und SDAI in der Größenordnung bestätigt.

Die Ergebnisse in den Teilpopulation 2 und 3 waren denen der relevanten Studienpopulation sehr ähnlich, in beiden Teilpopulationen zeigte sich ein signifikanter Vorteil für Sarilumab (RR [95 % KI] p-Wert: 2,68 [1,20; 5,99] 0,0168 in der Teilpopulation 2 bzw. 2,17 [1,13; 4,18] 0,0204 in der Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Niedrige Krankheitsaktivität gemäß CDAI (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-54: Ergebnisse für „Niedrige Krankheitsaktivität gemäß CDAI (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit CDAI ≤ 10 in Woche 24 (ITT)** | | **Patienten mit CDAI ≤ 10 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 39 (44,8) | 75 | 45 (60,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 16 (45,7) | 30 | 19 (63,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 23 (44,2) | 45 | 26 (57,8) |
| CDAI: *Clinical Disease Activity Index*<br>Patienten mit fehlenden Werten wurden als Non-Responder gezählt.<br>Quelle: [90] | | | | |

Die Auswertung des Endpunkts „Niedrige Krankheitsaktivität gemäß CDAI“ nach 48 Wochen Behandlung mit Sarilumab ergab im Vergleich zur Auswertung nach 24 Wochen einen weiteren Anstieg der Anzahl der Patienten, die eine niedrige Krankheitsaktivität gemäß CDAI erreichten (siehe Tabelle 4-54).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Niedrige Krankheitsaktivität gemäß CDAI“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.3.3 Niedrige Krankheitsaktivität gemäß SDAI

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

Tabelle 4-55: Operationalisierung von „Niedrige Krankheitsaktivität gemäß SDAI“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die niedrige Krankheitsaktivität gemäß SDAI ist als Erreichen eines SDAI-Werts ≤ 11 definiert.<br>Die Operationalisierung des SDAI ist in Tabelle 4-28 dargestellt.<br>Der Anteil der Patienten mit SDAI ≤ 11 wurde post hoc für Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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 4-56: Bewertung des Verzerrungspotenzials für „Niedrige Krankheitsaktivität gemäß SDAI“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des hier berichteten, nicht präspezifizierten Endpunktes „Niedrige Krankheitsaktivität gemäß SDAI“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-57: Ergebnisse für „Niedrige Krankheitsaktivität gemäß SDAI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit SDAI ≤ 11 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 40 (46,0) | 3,52 [1,73; 7,17] **0,0006** | 2,29 [1,40; 3,74] **0,0010** | 0,24 [0,11; 0,38] **0,0005** | |
| Adalimumab (N = 82) | 16 (19,5) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,6688 RR: 0,6599 RD: 0,7952 |
| Sarilumab (N = 35) | 16 (45,7) | 2,95 [1,02; 8,54] **0,0457** | 2,03 [1,01; 4,08] **0,0477** | 0,22 [0,02; 0,43] **0,0484** | |
| Adalimumab (N = 35) | 8 (22,9) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 24 (46,2) | 4,04 [1,52; 10,74] **0,0054** | 2,52 [1,26; 5,07] **0,0095** | 0,25 [0,08; 0,43] **0,0052** | |
| Adalimumab (N = 47) | 8 (17,0) | | | | |

SDAI: *Simplified Disease Activity Index*; ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Nach 24-wöchiger Behandlung zeigte sich auch bei der Zahl der Patienten, die eine niedrige Krankheitsaktivität gemäß SDAI erreichten, ein deutlicher Vorteil für Sarilumab gegenüber Adalimumab.

In der relevanten Studienpopulation erzielten nach 24 Wochen Behandlung 46,0 % der mit Sarilumab behandelten Patienten und 19,5 % der mit Adalimumab behandelten Patienten einen SDAI ≤ 11 (siehe Tabelle 4-57). Dieser Behandlungsunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 2,29 [1,40; 3,74] 0,0010). Dieses Ergebnis wird durch die

Auswertungen zur niedrigen Krankheitsaktivität gemäß CDAI und DAS28-ESR in der Größenordnung bestätigt.

Die Ergebnisse in den Teilpopulation 2 und 3 waren denen der relevanten Studienpopulation sehr ähnlich, in beiden Teilpopulationen zeigte sich ein signifikanter Vorteil für Sarilumab (RR [95 % KI] p-Wert: 2,03 [1,01; 4,08] 0,0477 in der Teilpopulation 2 bzw. 2,52 [1,26; 5,07] 0,0095 in der Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Niedrige Krankheitsaktivität gemäß SDAI (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-58: Ergebnisse für „Niedrige Krankheitsaktivität gemäß SDAI (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit SDAI ≤ 11 in Woche 24 (ITT)** | | **Patienten mit SDAI ≤ 11 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 40 (46,0) | 75 | 45 (60,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 16 (45,7) | 30 | 18 (60,0) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 24 (46,2) | 45 | 27 (60,0) |
| SDAI: *Simplified Disease Activity Index*<br>Patienten mit fehlenden Werten wurden als Non-Responder gewertet.<br>Quelle: [90] | | | | |

Die Auswertung des Endpunkts „Niedrige Krankheitsaktivität gemäß SDAI“ nach 48 Wochen Behandlung mit Sarilumab ergab im Vergleich zur Auswertung nach 24 Wochen einen weiteren Anstieg der Anzahl der Patienten, die eine niedrige Krankheitsaktivität gemäß SDAI erreichten (siehe Tabelle 4-58).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Niedrige Krankheitsaktivität gemäß SDAI“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.4 Therapieansprechen – RCT

##### 4.3.1.3.4.1 Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR

*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 „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die Kriterien der EULAR für ein Ansprechen auf eine Therapie beruhen auf der Bestimmung der Krankheitsaktivität mittels DAS28. Je nach Höhe des erreichten DAS28-Werts und der jeweiligen Differenz zum Ausgangswert unterscheidet man Patienten mit gutem, moderatem oder keinem Ansprechen: |

| DAS28-ESR | DAS28-ESR-Senkung (Verbesserung) | | |
|---|---|---|---|
| | > 1,2 | > 0,6 und ≤ 1,2 | ≤ 0,6 |
| ≤ 3,2 | gutes Ansprechen | moderates Ansprechen | kein Ansprechen |
| > 3,2 und ≤ 5,1 | | | |
| > 5,1 | | | |

Für die Auswertung des Ansprechens gemäß EULAR-Kriterien wurde der DAS28-ESR herangezogen. Die Operationalisierung des DAS28-ESR ist in Tabelle 4-21 dargestellt.

In der vorliegenden Auswertung wurde eine Responderanalyse für Woche 24 und Woche 48 für Patienten mit Ansprechen (gut oder moderat) im Vergleich zu Patienten ohne Ansprechen gemäß EULAR-Kriterien durchgeführt.

*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 „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Ansprechen gemäß EULAR-Kriterien“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-61: Ergebnisse für „Ansprechen gemäß EULAR-Kriterien (Woche 24) basierend auf DAS28-ESR“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **Ansprechen in Woche 24** | | | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| | **gut** | **moderat** | **kein** | | | | |
| (ITT) | n (%) | n (%) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 39 (44,8) | 32 (36,8) | 16 (18,4) | 2,13 [1,04; 4,38] **0,0395** | 1,22 [1,02; 1,46] **0,0328** | 0,14 [0,01; 0,27] **0,0296** | |
| Adalimumab (N = 82) | 8 (9,8) | 47 (57,3) | 27 (32,9) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | OR: 0,9527 RR: 0,9114 RD: 0,9558 |
| Sarilumab (N = 35) | 16 (45,7) | 13 (37,1) | 6 (17,1) | 2,22 [0,71; 6,98] 0,1696 | 1,21 [0,92; 1,58] 0,1742 | 0,14 [-0,06; 0,34] 0,1600 | |
| Adalimumab (N = 35) | 3 (8,6) | 21 (60,0) | 11 (31,4) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | |
| Sarilumab (N = 52) | 23 (44,2) | 19 (36,5) | 10 (19,2) | 2,13 [0,83; 5,46] 0,1161 | 1,23 [0,97; 1,57] 0,0919 | 0,15 [-0,02; 0,32] 0,0888 | |
| Adalimumab (N = 47) | 5 (10,6) | 26 (55,3) | 16 (34,0) | | | | |

EULAR: *European League Against Rheumatism*; ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Für die Responderanalyse wurden Patienten mit Ansprechen (gut oder moderat) im Vergleich zu Patienten ohne Ansprechen gemäß EULAR-Kriterien ausgewertet. Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen. Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 33 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Anzahl der Patienten mit einem Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR (siehe Tabelle 4-61).

Für die relevante Studienpopulation ergab die Responderanalyse eine signifikant erhöhte Wahrscheinlichkeit für ein gutes bis moderates Ansprechen unter Sarilumab im Vergleich zu Adalimumab (RR [95 % KI] p-Wert: 1,22 [1,02; 1,46] 0,0328). Dabei wird vor allem bei der

getrennten Betrachtung der Patientengruppen ein erheblicher Unterschied deutlich: 44,8 % der mit Sarilumab behandelten Patienten erreichten ein gutes Ansprechen gemäß EULAR-Kriterien, unter Adalimumab waren es lediglich 9,8 %. Demgegenüber zeigten 32,9 % der mit Adalimumab behandelten Patienten kein Ansprechen gemäß EULAR-Kriterien, während es unter Sarilumab nur 18,4 % waren.

Bei Auswertung der Patienten in den Teilpopulationen 2 und 3 zeigten sich numerisch sehr vergleichbare Ergebnisse bezüglich des Anteils der Patienten mit gutem bis moderatem Ansprechen, allerdings waren die Ergebnisse statistisch nicht signifikant (RR [95 % KI] p-Wert für Teilpopulation 2: 1,21 [0,92; 1,58] 0,1742 bzw. für Teilpopulation 3: 1,23 [0,97; 1,57] 0,0919).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR (Woche 24)“ das Ergebnis der relevanten Studienpopulation herangezogen werden kann.

Tabelle 4-62: Ergebnisse für „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR (Woche 48)“

| **Studie MONARCH** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Behandlung** | **Ansprechen in Woche 24 (ITT)** | | | | **Ansprechen in Woche 48 (ITT)** | | | |
| | | **gut** | **moderat** | **kein** | | **gut** | **moderat** | **kein** |
| | N | n (%) | n (%) | n (%) | N | n (%) | n (%) | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | | |
| Sarilumab | 87 | 39 (44,8) | 32 (36,8) | 16 (18,4) | 75 | 44 (58,7) | 27 (36,0) | 4 (5,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | | |
| Sarilumab | 35 | 16 (45,7) | 13 (37,1) | 6 (17,1) | 30 | 17 (56,7) | 13 (43,3) | 0 |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | | |
| Sarilumab | 52 | 23 (44,2) | 19 (36,5) | 10 (19,2) | 45 | 27 (60,0) | 14 (31,1) | 4 (8,9) |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Einbeziehung der Erythrozytensedimentationsrate
Patienten mit fehlenden Werten wurden als Non-Responder gezählt.
Quelle: [90]

Die Auswertung zum Endpunkt „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR“ nach 48 Wochen Behandlung mit Sarilumab zeigte im Vergleich zur Auswertung nach 24 Wochen einen weiteren Anstieg der Zahl der Patienten mit gutem Therapieansprechen, während sich die Zahl der Patienten mit moderatem oder keinem Ansprechen verringerte (siehe Tabelle 4-62). Es ist anzunehmen, dass unter den 16 Patienten, die in Woche 24 kein

Ansprechen auf die Behandlung mit Sarilumab zeigten (oder als Non-Responder gezählt wurden), die 12 Patienten sind, die die Behandlung innerhalb der 24-wöchigen Behandlungsphase abbrachen und demzufolge nicht an der offenen Extensionsphase teilnahmen.

Die Auswertung nach 48 Wochen belegt die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Ansprechen gemäß EULAR-Kriterien basierend auf DAS28-ESR“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.4.2 Ansprechen gemäß ACR-Kriterien

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

Tabelle 4-63: Operationalisierung von „Ansprechen gemäß ACR-Kriterien“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die ACR-Kriterien zur Beurteilung eines Therapieansprechens umfassen sieben Parameter:<br>– Anzahl der druckschmerzhaften Gelenke von 68 Gelenken[22] (68TJC): Die Anzahl der druckschmerzhaften Gelenke von insgesamt 68 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt.<br>– Anzahl der geschwollenen Gelenke von 66 Gelenken (66SJC): Die Anzahl der geschwollenen Gelenke von insgesamt 66 Gelenken (siehe Fußnote, mit Ausnahme der Hüftgelenke) wurde durch einen unabhängigen, geschulten Untersucher bestimmt.<br>– CRP: Die Bestimmung des CRP-Werts als Entzündungsparameter wurde anhand einer Nüchternblutprobe in einem Zentrallabor durchgeführt.<br>– Schmerz: Die Patienten schätzten die Intensität ihrer Schmerzen auf einer 100 mm-VAS ein, wobei 0 „kein Schmerz“ und 100 „stärkste vorstellbare Schmerzen“ bedeutet.<br>– Globale Einschätzung des Patienten (PtGA, *Patient Global Assessment*): Die Patienten schätzten ihre Krankheitsaktivität auf einer 100 mm-VAS ein, wobei 0 „keine Krankheitsaktivität“ und 100 „höchstmögliche Krankheitsaktivität“ bedeutet.<br>– Globale Einschätzung des Arztes (PhGA, *Physicians Global Assessment*): Der behandelnde Arzt schätzte die Krankheitsaktivität des Patienten auf einer 100 mm-VAS ein, wobei 0 „keine Krankheitsaktivität“ und 100 „höchstmögliche Krankheitsaktivität“ bedeutet.<br>– Funktionsstatus: Zur Bestimmung des Funktionsstatus der Patienten wurde der HAQ-DI-Fragebogen herangezogen. Die Operationalisierung ist im Detail in Tabelle 4-93 dargestellt.<br>In der Studie MONARCH waren Responderanalysen bezüglich einer 20%igen, 50%igen und 70%igen Verbesserung gemäß ACR- Kriterien a priori geplant.<br>Eine 20%ige Verbesserung (ACR20) erfordert eine mindestens 20%ige Verbesserung der Anzahl der druckschmerzhaften und geschwollenen Gelenke sowie eine mindestens 20%ige Verbesserung für mindestens drei der übrigen fünf Kriterien.<br>Eine 50%ige Verbesserung (ACR50) erfordert eine mindestens 50%ige Verbesserung der Anzahl der druckschmerzhaften und geschwollenen Gelenke sowie eine mindestens 50%ige Verbesserung für mindestens drei der übrigen fünf Kriterien.<br>Eine 70%ige Verbesserung (ACR70) erfordert eine mindestens 70%ige Verbesserung der Anzahl der druckschmerzhaften und geschwollenen Gelenke sowie eine mindestens 70%ige Verbesserung für mindestens drei der übrigen fünf Kriterien.<br>Die Einzelkomponenten wurden zu Beginn der Studie (Woche 0) und anschließend alle vier Wochen in der randomisierten Behandlungsphase sowie in der offenen Extensionsphase erhoben. ACR20, ACR50 und ACR70 wurden in Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) ermittelt. |

[22] Kiefergelenke (n=2), Sternoklavikulargelenke (Brustbein-Schlüsselbein-Gelenke, n=2), Akromioklavikulargelenke (Schultereckgelenke, n=2), Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke der Hände (Fingermittelgelenke, n=8), distale Interphalangealgelenke der Hände (Fingerendgelenke, n=8), Hüftgelenke (n=2), Kniegelenke (n=2), obere Sprungelenke (n=2), untere Sprungelenke (n=2), Metatarsophalangealgelenke (Zehengrundgelenke, n=10), Interphalangealgelenke der großen Zehen (n=2), proximale/distale Interphalangealgelenke der Zehen (Zehenmittel- und -endgelenke, n=8)

*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 4-64: Bewertung des Verzerrungspotenzials für „Ansprechen gemäß ACR-Kriterien“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Ansprechen gemäß ACR-Kriterien“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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.*

### *ACR20*

Tabelle 4-65: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR20 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit ACR20 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 62 (71,3) | 2,04 [1,06; 3,94] **0,0339** | 1,29 [1,03; 1,61] **0,0296** | 0,16 [0,02; 0,30] **0,0242** | |
| Adalimumab (N = 82) | 45 (54,9) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,2788 RR: 0,7334 RD: 0,4592 |
| Sarilumab (N = 35) | 27 (77,1) | 3,40 [1,08; 10,70] **0,0362** | 1,33 [0,97; 1,82] 0,0786 | 0,22 [0,02; 0,42] **0,0343** | |
| Adalimumab (N = 35) | 19 (54,3) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 35 (67,3) | 1,56 [0,67; 3,61] 0,3013 | 1,23 [0,90; 1,68] 0,1938 | 0,11 [-0,08; 0,30] 0,2487 | |
| Adalimumab (N = 47) | 26 (55,3) | | | | |

ACR: *American College of Rheumatology*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 26 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Für die Patienten der Studie MONARCH, die bei Studieneinschluss sämtlich eine hohe Krankheitsaktivität aufwiesen, stellt eine Verbesserung der Krankheitsaktivität um 20 % vermutlich keinen ausreichenden Therapieerfolg dar. Aus diesem Grund wird die Auswertung zum ACR20 als nicht patientenrelevant eingestuft, aber aus Transparenzgründen hier dargestellt.

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich des Anteils der Patienten mit einer 20%igen Verbesserung gemäß ACR-Kriterien (ACR20).

In der relevanten Studienpopulation erreichten 71,3 % der Patienten unter Sarilumab ACR20 gegenüber 54,9 % der Patienten unter Adalimumab. Die Wahrscheinlichkeit, ACR20 zu erreichen, war somit unter Sarilumab im Vergleich zu Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 1,29 [1,03; 1,61] 0,0296).

Bei Auswertung der Patienten in den Teilpopulationen 2 und 3 zeigten sich numerisch sehr vergleichbare Ergebnisse bezüglich des Anteils der Patienten mit ACR20, allerdings waren die Ergebnisse statistisch nicht signifikant (RR [95 % KI] p-Wert für Teilpopulation 2: 1,33 [0,97; 1,82] 0,0786 bzw. für Teilpopulation 3: 1,23 [0,90; 1,68] 0,1938).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation für die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Ansprechen gemäß ACR-Kriterien – ACR20 (Woche 24)“ das Ergebnis der relevanten Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

Tabelle 4-66: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR20 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit ACR20 in Woche 24 (ITT)** | | **Patienten mit ACR20 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 62 (71,3) | 75 | 59 (78,7) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 27 (77,1) | 30 | 24 (80,0) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 35 (67,3) | 45 | 35 (77,8) |

ACR: *American College of Rheumatology*
Patienten mit fehlenden Werten wurden als Non-Responder gezählt.
Quelle: [90]

Die Auswertung des ACR20 nach 48 Wochen Behandlung mit Sarilumab zeigte eine etwa gleichbleibende Anzahl von Patienten mit einer 20%igen Verbesserung gemäß ACR-Kriterien (siehe Tabelle 4-66). Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

### *ACR50*

Tabelle 4-67: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR50 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit ACR50 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 41 (47,1) | 2,38 [1,23; 4,57] **0,0098** | 1,73 [1,14; 2,63] **0,0104** | 0,19 [0,05; 0,34] **0,0091** | |
| Adalimumab (N = 82) | 22 (26,8) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,3852 RR: 0,3116 RD: 0,4393 |
| Sarilumab (N = 35) | 15 (42,9) | 1,66 [0,61; 4,54] 0,3178 | 1,31 [0,70; 2,44] 0,3915 | 0,12 [-0,10; 0,34] 0,2768 | |
| Adalimumab (N = 35) | 11 (31,4) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 26 (50,0) | 3,00 [1,24; 7,30] **0,0156** | 2,03 [1,13; 3,66] **0,0183** | 0,24 [0,05; 0,43] **0,0155** | |
| Adalimumab (N = 47) | 11 (23,4) | | | | |

ACR: *American College of Rheumatology*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 27 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich des Anteils der Patienten mit einer 50%igen Verbesserung gemäß der ACR-Kriterien (ACR50). Eine 50%ige Verringerung der Krankheitsaktivität ist für die Patienten der Studie MONARCH, die bei Studieneinschluss sämtlich eine hohe Krankheitsaktivität aufwiesen, im Gegensatz zur 20%igen Verbesserung als klinisch relevantes Therapieansprechen einzuschätzen.

In der relevanten Studienpopulation erreichten 47,1 % der Patienten unter Sarilumab ACR50 gegenüber 26,8 % der Patienten unter Adalimumab. Die Wahrscheinlichkeit, ACR50 zu

erreichen, war damit unter Sarilumab im Vergleich zu Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 1,73 [1,14; 2,63] 0,0104).

Bei Auswertung der Patienten in Teilpopulation 2 zeigte sich ein numerischer Vorteil für Sarilumab bezüglich des Anteils der Patienten mit ACR50, allerdings war der Behandlungsunterschied nicht statistisch signifikant (RR [95 % KI] p-Wert: 1,31 [0,70; 2,44] 0,3915).

In Teilpopulation 3 war die Wahrscheinlichkeit, ACR50 zu erreichen, unter Sarilumab im Vergleich zu Adalimumab wie in der relevanten Studienpopulation statistisch signifikant erhöht (RR [95 % KI] p-Wert: 2,03 [1,13; 3,66] 0,0183).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Ansprechen gemäß ACR-Kriterien – ACR50 (Woche 24)“ das Ergebnis der relevanten Studienpopulation herangezogen werden kann.

Tabelle 4-68: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR50 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit ACR50 in Woche 24 (ITT)** | | **Patienten mit ACR50 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 41 (47,1) | 75 | 44 (58,7) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 15 (42,9) | 30 | 19 (63,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 26 (50,0) | 45 | 25 (55,6) |

ACR: *American College of Rheumatology*
Patienten mit fehlenden Werten wurden als Non-Responder gezählt.
Quelle: [90]

Die Auswertung des ACR50 nach 48 Wochen Behandlung mit Sarilumab zeigte eine etwa gleichbleibende Anzahl von Patienten mit einer 50%igen Verbesserung gemäß ACR-Kriterien (siehe Tabelle 4-68). Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

### *ACR70*

Tabelle 4-69: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR70 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit ACR70 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 21 (24,1) | 3,34 [1,31; 8,54] **0,0118** | 2,62 [1,17; 5,85] **0,0194** | 0,16 [0,04; 0,29] **0,0114** | |
| Adalimumab (N = 82) | 7 (8,5) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,5016 RR: 0,5100 RD: 0,6189 |
| Sarilumab (N = 35) | 8 (22,9) | 2,37 [0,62; 9,08] 0,2055 | 1,98 [0,66; 5,93] 0,2204 | 0,11 [-0,06; 0,29] 0,1987 | |
| Adalimumab (N = 35) | 4 (11,4) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 13 (25,0) | 4,56 [1,16; 17,86] **0,0296** | 3,43 [1,01; 11,56] **0,0473** | 0,19 [-0,06; 0,43] 0,1288 | |
| Adalimumab (N = 47) | 3 (6,4) | | | | |

ACR: *American College of Rheumatology*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 28 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich des Anteils der Patienten mit einer 70%igen Verbesserung gemäß der ACR-Kriterien (ACR70). Eine 70%ige Verringerung der Krankheitsaktivität ist für die Patienten der Studie MONARCH, die bei Studieneinschluss sämtlich eine hohe Krankheitsaktivität aufwiesen, im Gegensatz zur 20%igen Verbesserung als klinisch relevantes Therapieansprechen einzuschätzen.

In der relevanten Studienpopulation erreichten 24,1 % der Patienten unter Sarilumab ACR70 gegenüber 8,5 % der Patienten unter Adalimumab. Die Wahrscheinlichkeit, ACR70 zu

erreichen, war damit unter Sarilumab im Vergleich zu Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 2,62 [1,17; 5,85] 0,0194).

Bei Auswertung der Patienten in Teilpopulation 2 zeigte sich ein numerischer Vorteil für Sarilumab bezüglich des Anteils der Patienten mit ACR70, allerdings war der Behandlungsunterschied nicht statistisch signifikant (RR [95 % KI] p-Wert: 1,98 [0,66; 5,93] 0,2204).

In Teilpopulation 3 war die Wahrscheinlichkeit, ACR70 zu erreichen, unter Sarilumab im Vergleich zu Adalimumab wie in der relevanten Studienpopulation statistisch signifikant erhöht (RR [95 % KI] p-Wert: 3,43 [1,01; 11,56] 0,0473).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Ansprechen gemäß ACR-Kriterien – ACR70 (Woche 24)“ das Ergebnis der relevanten Studienpopulation herangezogen werden kann.

Tabelle 4-70: Ergebnisse für „Ansprechen gemäß ACR-Kriterien – ACR70 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit ACR70 in Woche 24** | | **Patienten mit ACR70 in Woche 48** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 21 (24,1) | 75 | 30 (40,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 8 (22,9) | 30 | 11 (36,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 13 (25,0) | 45 | 19 (42,2) |
| ACR: *American College of Rheumatology*<br>Quelle: [90] | | | | |

Die Auswertung des ACR70 nach 48 Wochen Behandlung mit Sarilumab zeigte im Vergleich zur Auswertung nach 24 Wochen einen Anstieg der Anzahl von Patienten mit einer 70%igen Verbesserung gemäß ACR-Kriterien (siehe Tabelle 4-70). Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Ansprechen gemäß ACR-Kriterien“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.5 Einzelkomponenten der Scores zur Bestimmung der Krankheitsaktivität –RCT

In Tabelle 4-71 ist zusammengefasst, wie die Einzelkomponenten zur Berechnung der in den vorherigen Abschnitten dargestellten Gesamtscores kombiniert wurden.

Tabelle 4-71: Einzelkomponenten der Scores zur Bestimmung der Krankheitsaktivität

| Einzelkomponenten | DAS28-ESR | CDAI | SDAI | ACR |
|---|---|---|---|---|
| Gelenksymptome: | | | | |
| - Druckschmerzhafte Gelenke (68TJC) | | | | X |
| - Druckschmerzhafte Gelenke (28TJC) | X | X | X | |
| - Geschwollene Gelenke (66SJC) | | | | X |
| - Geschwollene Gelenke (28SJC) | X | X | X | |
| Schmerzen | | | | X |
| Globale Einschätzung des Arztes (PhGA) | | X | X | X |
| Globale Einschätzung des Patienten (PtGA) | X | X | X | X |
| Körperlicher Funktionsstatus - HAQ-DI | | | | X |
| Laborparameter: | | | | |
| - CRP | | | X | X |
| - ESR | X | | | |

#### 4.3.1.3.5.1 Gelenksymptome: Druckschmerzhafte Gelenke

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

Tabelle 4-72: Operationalisierung von „Druckschmerzhafte Gelenke“

| Studie | Operationalisierung |
|---|---|
| MONARCH | – Anzahl der druckschmerzhaften Gelenke von 68 Gelenken (68TJC): Die Anzahl der druckschmerzhaften Gelenke von insgesamt 68 Gelenken[23] wurde durch einen unabhängigen, geschulten Untersucher bestimmt.<br>– Anzahl der druckschmerzhaften Gelenke von 28 Gelenken (28TJC): Die Anzahl der druckschmerzhaften Gelenke von insgesamt 68 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt. Davon wurden 28 Gelenke[24] berücksichtigt.<br>Die Zeitpunkte der Erhebung der jeweiligen Parameter sind im Abschnitt zur Operationalisierung der jeweiligen Scores (ACR 20/50/70; DAS28-ESR, CDAI. SDAI) dargestellt.<br>Die dargestellte Änderung der TJC ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. |

*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.*

[23] Kiefergelenke (n=2), Sternoklavikulargelenke (Brustbein-Schlüsselbein-Gelenke, n=2), Akromioklavikulargelenke (Schultereckgelenke, n=2), Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke der Hände (Fingermittelgelenke, n=8), distale Interphalangealgelenke der Hände (Fingerendgelenke, n=8), Hüftgelenke (n=2), Kniegelenke (n=2), obere Sprungelenke (n=2), untere Sprungelenke (n=2), Metatarsophalangealgelenke (Zehengrundgelenke, n=10), Interphalangealgelenke der großen Zehen (n=2), proximale/distale Interphalangealgelenke der Zehen (Zehenmittel- und -endgelenke, n=8)

[24] Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke (Fingermittelgelenke, n=8), Kniegelenke (n=2)

Tabelle 4-73: Bewertung des Verzerrungspotenzials für „Druckschmerzhafte Gelenke“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse der druckschmerzhaften Gelenke ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-74: Ergebnisse für „Druckschmerzhafte Gelenke – 68TJC (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **68TJC-Basiswert** | | **68TJC in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 27,59 (12,00) | 76 | 8,82 (10,15) | -18,77 (11,07) | -1,281 [-4,314; 1,752] 0,4056 | |
| Adalimumab (N = 82) | 70 | 27,75 (13,45) | 70 | 10,95 (11,10) | -16,81 (12,63) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 27,68 (12,32) | 31 | 8,32 (8,49) | -19,35 (9,79) | -2,201 [-6,308; 1,906] 0,2883 | |
| Adalimumab (N = 35) | 31 | 28,06 (11,98) | 31 | 11,52 (11,84) | -16,55 (10,11) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,6532 |
| Sarilumab (N = 52) | 45 | 27,52 (11,92) | 45 | 9,16 (11,23) | -18,37 (11,96) | -0,532 [-4,997; 3,932] 0,8134 | |
| Adalimumab (N = 47) | 39 | 27,51 (14,67) | 39 | 10,49 (10,60) | -17,02 (14,46) | | |

TJC: *Tender Joint Count*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 29 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-75: Ergebnisse für „Druckschmerzhafte Gelenke – 28TJC (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **28TJC-Basiswert** | | **28TJC in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 16,78 (5,54) | 76 | 4,84 (5,39) | -11,93 (6,08) | -1,452 [-3,126; 0,222] 0,0887 | |
| Adalimumab (N = 82) | 70 | 16,09 (6,50) | 70 | 6,32 (5,37) | -9,78 (7,11) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 17,19 (5,38) | 31 | 4,65 (4,42) | -12,55 (5,11) | -2,364 [-4,750; 0,021] 0,0520 | |
| Adalimumab (N = 35) | 31 | 16,45 (6,05) | 31 | 7,16 (5,78) | -9,29 (6,56) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,9008 |
| Sarilumab (N = 52) | 45 | 16,49 (5,70) | 45 | 4,98 (6,01) | -11,51 (6,70) | -0,706 [-3,055; 1,643] 0,5520 | |
| Adalimumab (N = 47) | 39 | 15,81 (6,90) | 39 | 5,64 (5,00) | -10,17 (7,59) | | |

TJC: *Tender Joint Count*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Nach 24-wöchiger Behandlung zeigte sich sowohl bei den Patienten unter Sarilumab als auch unter Adalimumab eine sehr deutliche Reduktion der Anzahl der druckschmerzhaften Gelenke (TJC, *Tender Joint Counts*).

Bei Studienbeginn wurden durchschnittlich etwa 28 von 68 Gelenken als druckschmerzhaft eingestuft (siehe Tabelle 4-72). In der relevanten Studienpopulation waren nach 24-wöchiger Behandlung etwa 11 Gelenke bei Patienten unter Adalimumab und etwa 9 Gelenke unter Sarilumab druckschmerzhaft. Der numerische Vorteil, der sich für Sarilumab gegenüber Adalimumab gezeigt hat, ist statistisch nicht signifikant (LSMD [95 % KI] p-Wert: -1,281 [-4,314; 1,752] 0,4056). In den Teilpopulationen 2 und 3 wurden vergleichbare Verringerungen der Anzahl druckschmerzhafter Gelenke erzielt.

Bei der Betrachtung von 28 Gelenken wurden zu Studienbeginn durchschnittlich etwa 17 Gelenke als druckschmerzhaft eingestuft (siehe Tabelle 4-75). Nach 24-wöchiger Behandlung traf dies in der relevanten Studienpopulation auf etwa 6 Gelenke unter Adalimumab bzw. 5 Gelenke unter Sarilumab zu. Der numerische Vorteil, der sich für Sarilumab gegenüber Adalimumab gezeigt hat, ist statistisch nicht signifikant (LSMD [95 % KI] p-Wert: -1,452 [-3,126; 0,222] 0,0887). In den Teilpopulationen 2 und 3 wurden vergleichbare Verringerungen der Anzahl druckschmerzhafter Gelenke erzielt.

Die Interaktionstests ergaben keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für die Auswertung der druckschmerzhaften Gelenke nach 68TJC und 28TJC die Ergebnisse der relevanten Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden können.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.5.2 Gelenksymptome: Geschwollene Gelenke

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

Tabelle 4-76: Operationalisierung von „Geschwollene Gelenke “

| Studie | Operationalisierung |
|---|---|
| MONARCH | – Anzahl der geschwollenen Gelenke von 66 Gelenken (66SJC): Die Anzahl der geschwollenen Gelenke von insgesamt 66 Gelenken[25] wurde durch einen unabhängigen, geschulten Untersucher bestimmt.<br>– Anzahl der geschwollenen Gelenke von 28 Gelenken (28SJC): Die Anzahl der geschwollenen Gelenke von insgesamt 66 Gelenken wurde durch einen unabhängigen, geschulten Untersucher bestimmt. Davon wurden 28 Gelenke[26] berücksichtigt.<br>Die Zeitpunkte der Erhebung der jeweiligen Parameter sind im Abschnitt zur Operationalisierung der jeweiligen Scores (ACR 20/50/70; DAS28-ESR, CDAI, SDAI) dargestellt.<br>Die dargestellte Änderung der SJC ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. |

*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.*

[25] Kiefergelenke (n=2), Sternoklavikulargelenke (Brustbein-Schlüsselbein-Gelenke, n=2), Akromioklavikulargelenke (Schultereckgelenke, n=2), Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke der Hände (Fingermittelgelenke, n=8), distale Interphalangealgelenke der Hände (Fingerendgelenke, n=8), Kniegelenke (n=2), obere Sprungelenke (n=2), untere Sprungelenke (n=2), Metatarsophalangealgelenke (Zehengrundgelenke, n=10), Interphalangealgelenke der großen Zehen (n=2), proximale/distale Interphalangealgelenke der Zehen (Zehenmittel- und -endgelenke, n=8)

[26] Schultergelenke (n=2), Ellbogengelenke (n=2), Handgelenke (n=2), Metacarpophalangealgelenke (Fingergrundgelenke, n=10), Interphalangealgelenke der Daumen (n=2), proximale Interphalangealgelenke (Fingermittelgelenke, n=8), Kniegelenke (n=2)

Tabelle 4-77: Bewertung des Verzerrungspotenzials für „Geschwollene Gelenke “ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse der geschwollenen Gelenke ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-78: Ergebnisse für „Geschwollene Gelenke – 66SJC (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **66SJC - Basiswert** | | **66SJC in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 19,04 (9,98) | 76 | 3,96 (5,54) | -15,08 (8,75) | -1,513 [-3,080; 0,054] 0,0584 | |
| Adalimumab (N = 82) | 70 | 19,26 (10,92) | 70 | 5,83 (5,12) | -13,43 (10,38) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 19,77 (11,69) | 31 | 3,03 (4,51) | -16,74 (10,28) | -2,608 [-4,711; -0,506] 0,0159 **-0,622 [-1,124;-0,121]** | 0,0707 |
| Adalimumab (N = 35) | 31 | 19,26 (10,30) | 31 | 5,65 (5,41) | -13,61 (8,53) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | |
| Sarilumab (N = 52) | 45 | 18,53 (8,72) | 45 | 4,60 (6,11) | -13,93 (7,44) | -0,567 [-2,800; 1,666] 0,6155 | |
| Adalimumab (N = 47) | 39 | 19,26 (11,53) | 39 | 5,98 (4,95) | -13,28 (11,76) | | |

SJC: *Swollen Joint Count*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 29 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-79: Ergebnisse für „Geschwollene Gelenke – 28SJC (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **28SJC - Basiswert** | | **28SJC in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 13,36 (5,35) | 76 | 2,66 (3,11) | -10,70 (4,95) | -1,333 [-2,336; -0,329] 0,0095 **-0,431 [-0,755; -0,107]** | |
| Adalimumab (N = 82) | 70 | 13,29 (5,53) | 70 | 4,23 (3,37) | -9,06 (6,28) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 13,19 (5,44) | 31 | 2,26 (2,87) | -10,94 (5,15) | -1,342 [-2,751; 0,067] 0,0616 | |
| Adalimumab (N = 35) | 31 | 13,00 (5,47) | 31 | 3,65 (2,80) | -9,35 (6,10) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,1318 |
| Sarilumab (N = 52) | 45 | 13,47 (5,35) | 45 | 2,93 (3,26) | -10,53 (4,86) | -1,244 [-2,659; 0,170] 0,0840 | |
| Adalimumab (N = 47) | 39 | 13,51 (5,63) | 39 | 4,69 (3,73) | -8,82 (6,49) | | |

SJC: *Swollen Joint Count*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Nach 24-wöchiger Behandlung zeigte sich sowohl bei den Patienten unter Sarilumab als auch unter Adalimumab eine sehr deutliche Reduktion der Anzahl der geschwollenen Gelenke (SJC, *Swollen Joint Counts*).

Bei Studienbeginn wurden durchschnittlich etwa 19 von 66 Gelenken als geschwollen eingestuft (siehe Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-76). In der relevanten Studienpopulation waren nach 24-wöchiger Behandlung noch etwa 5,8 Gelenke bei Patienten unter Adalimumab und etwa 4,0 Gelenke unter Sarilumab geschwollen. Der numerische Vorteil, der sich für Sarilumab gegenüber Adalimumab gezeigt hat, ist statistisch knapp nicht signifikant (LSMD [95 % KI] p-Wert: -1,513 [-3,080; 0,054] 0,0584). In Teilpopulation 2 wurde hingegen eine statistisch signifikante Verringerung der Anzahl geschwollener Gelenke erzielt (LSMD [95 % KI] p-Wert: -2,608 [-4,711; -0,506] 0,0159). Das Konfidenzintervall zu Hedges´g schließt jedoch den Wert -0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt. In Teilpopulation 3 zeigte sich keine statistisch signifikante Verringerung der Anzahl geschwollener Gelenke.

Bei der Betrachtung von 28 Gelenken wurden zu Studienbeginn durchschnittlich etwa 13,3 Gelenke als geschwollen eingestuft (siehe Tabelle 4-79). Nach 24-wöchiger Behandlung traf dies in der relevanten Studienpopulation im Mittel noch auf etwa 4,2 Gelenke unter Adalimumab bzw. 2,7 Gelenke unter Sarilumab zu. Dies stellt einen statistisch signifikanten Vorteil für Sarilumab gegenüber Adalimumab dar (LSMD [95 % KI] p-Wert: -1,333 [-2,336; -0,329] 0,0095). Das Konfidenzintervall zu Hedges´g schließt jedoch den Wert -0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt. In den Teilpopulationen 2 und 3 wurden vergleichbare Verringerungen der Anzahl geschwollener Gelenke erzielt.

Die Interaktionstests ergaben keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für die Auswertung der geschwollenen Gelenke nach 66SJC und 28SJC die Ergebnisse der relevanten Studienpopulation herangezogen werden können.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.5.3 Schmerzen

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

Tabelle 4-80: Operationalisierung von „Schmerzen“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die Patienten schätzten die Intensität ihrer Schmerzen auf einer 100 mm-VAS ein, wobei 0 „kein Schmerz“ und 100 „stärkste vorstellbare Schmerzen“ bedeutet.<br>Zeitpunkte der Erhebung: zu Beginn der Studie (Woche 0) und anschließend alle vier Wochen in der randomisierten Behandlungsphase sowie in der offenen Extensionsphase.<br>Die dargestellte Änderung der mittleren Schmerzen ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0.<br>Der Anteil der Patienten mit einer Verbesserung um mindestens 20,4 mm auf der VAS wurde post hoc für Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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 4-81: Bewertung des Verzerrungspotenzials für „Schmerzen“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Schmerzen“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-82: Ergebnisse für „Schmerz (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **Schmerz – Basiswert [mm]** | | **Schmerz in Woche 24 [mm]** | | **absolute Änderung im Vergleich zum Basiswert [mm]** | **Behandlungs-differenz[a, b] [mm]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 71,82 (19,85) | 76 | 33,07 (21,38) | -38,75 (27,39) | -11,437 [-18,456; -4,419] 0,0016 **-0,524 [-0,846; -0,202]** | |
| Adalimumab (N = 82) | 69 | 71,16 (20,86) | 69 | 43,86 (23,19) | -27,30 (24,06) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 71,26 (17,97) | 31 | 34,71 (23,48) | -36,55 (27,72) | -11,084 [-22,422; 0,254] 0,0552 | |
| Adalimumab (N = 35) | 30 | 66,97 (23,28) | 30 | 44,07 (23,25) | -22,90 (22,89) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,8427 |
| Sarilumab (N = 52) | 45 | 72,20 (21,24) | 45 | 31,93 (20,00) | -40,27 (27,37) | -11,687 [-21,002; -2,371] 0,0145 **-0,537 [-0,965; -0,109]** | |
| Adalimumab (N = 47) | 39 | 74,38 (18,09) | 39 | 43,69 (23,44) | -30,69 (24,68) | | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 29 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24 Wochen Behandlung zeigte sich eine Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verringerung von Schmerzen.

Die Schmerzintensität zu Beginn der Studie wurde von den Patienten der relevanten Studienpopulation durchschnittlich mit etwa 71 mm auf der 100 mm-VAS angegeben (siehe Tabelle 4-82). Bei Patienten mit einem Wert > 70 mm kann man davon ausgehen, dass die Schmerzen als starke Schmerzen empfunden werden [91]. Die Patienten unter Sarilumab erreichten nach 24-wöchiger Behandlung einen durchschnittlichen Wert von 33 mm, während die Patienten unter Adalimumab im Mittel etwa 44 mm angaben. Gemäß Leitlinie der EMA zur klinischen Entwicklung von Arzneimitteln zur Behandlung von Schmerz stellt ein Wert von > 40 mm einen üblichen Schwellenwert für mittelstarke Schmerzen dar [91]. Somit kann man folgern, dass Patienten unter Adalimumab im Mittel nach 24-wöchiger Behandlung die Intensität ihrer Schmerzen als mittelstark empfanden, wohingegen die Patienten unter Sarilumab im Mittel ihre Schmerzen als leichte Schmerzen einstuften.

Zusammenfassend zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung der Schmerzen (LSMD [95% KI] p-Wert: -11,437 [-18,456; -4,419] 0,0016). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g belegt, dass es sich dabei um einen klinisch relevanten Effekt mit kleiner Effektstärke handelt (Hedges´g [95% KI]: -0,524 [-0,846; -0,202]).

Bei den Patienten der Teilpopulation 2 zeigte sich unter Sarilumab im Vergleich zu Adalimumab eine statistisch nicht signifikante Verringerung der Schmerzen (LSMD [95% KI] p-Wert: -11,084 [-22,422; 0,254] 0,0552).

Bei den Patienten der Teilpopulation 3 zeigte sich wiederum unter Sarilumab im Vergleich zu Adalimumab eine statistisch signifikant stärkere Verringerung der Schmerzen (LSMD [95% KI] p-Wert: -11,687 [-21,002; -2,371] 0,0145). Das Konfidenzintervall des Hedges´g schließt jedoch den Wert -0,2 mit ein, so dass keine Aussagen zur klinischen Relevanz dieses Effekts abgeleitet werden können.

Der Interaktionstest ergab keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Schmerz (Woche 24)“ das Ergebnis der relevanten Studienpopulation betrachtet werden kann.

In einer Responderanalyse wurde des Weiteren untersucht, wie viele Patienten eine Verringerung ihrer Schmerzen um mindestens 20,4 mm erreichten. Dieser Schwellenwert wurde für Patienten mit rheumatoider Arthritis als minimale klinische relevante Verbesserung validiert [25]. Bei Betrachtung der Mittelwertdifferenzen zeigt sich jedoch, dass die mittlere Änderung der Schmerzen bereits deutlich über dem definierten Schwellenwert lag (bei den Patienten der relevanten Studienpopulation ca. 38,8 mm bzw. 27,3 mm bei Behandlung mit Sarilumab bzw. Adalimumab).

Dementsprechend erreichten in der relevanten Studienpopulation 52,4 % der Patienten unter Adalimumab und 58,6 % der Patienten unter Sarilumab eine Verbesserung um 20,4 mm (siehe Tabelle 4-83). Diese Gruppendifferenz war statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,07 [0,83; 1,39] 0,5931). In den Teilpopulationen 2 und 3 ergaben sich vergleichbare Ergebnisse.

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Schmerz-Responder (Woche 24)“ das Ergebnis der relevanten Studienpopulation betrachtet werden kann.

Tabelle 4-83: Ergebnisse für „Schmerz-Responder (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Verbesserung ≥ 20,4 mm in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 51 (58,6) | 1,29 [0,69; 2,42] 0,4307 | 1,07 [0,83; 1,39] 0,5931 | 0,06 [-0,09; 0,21] 0,4339 | |
| Adalimumab (N = 82) | 43 (52,4) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,4469 RR: 0,7314 RD: 0,4574 |
| Sarilumab (N = 35) | 19 (54,3) | 1,63 [0,61; 4,33] 0,3285 | 1,13 [0,69; 1,85] 0,6258 | 0,12 [-0,11; 0,35] 0,3212 | |
| Adalimumab (N = 35) | 15 (42,9) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 32 (61,5) | 0,98 [0,42; 2,32] 0,9690 | 1,02 [0,76; 1,37] 0,8849 | 0,00 [-0,18; 0,19] 0,9677 | |
| Adalimumab (N = 47) | 28 (59,6) | | | | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Auswertung der Anzahl der Patienten mit einer Verringerung des Schmerzes um mindestens 20,4 mm nach 48 Wochen Behandlung mit Sarilumab zeigte ein gleichbleibendes Ergebnis im Vergleich zur Auswertung nach Woche 24 (siehe Tabelle 4-84). Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

Tabelle 4-84: Ergebnisse für „Schmerz-Responder (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 20,4 mm in Woche 24 (ITT)** | | **Patienten mit Verbesserung ≥ 20,4 mm in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 51 (58,6) | 75 | 52 (69,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 19 (54,3) | 30 | 19 (63,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 32 (61,5) | 45 | 33 (73,3) |
| Patienten mit fehlenden Werten wurden als Non-Responder gezählt.<br>Quelle: [90] | | | | |

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.5.4 Globale Einschätzung des Arztes (PhGA)

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

Tabelle 4-85: Operationalisierung von „PhGA“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Der behandelnde Arzt schätzte die Krankheitsaktivität des Patienten auf einer 100 mm-VAS ein, wobei 0 „keine Krankheitsaktivität“ und 100 „höchstmögliche Krankheitsaktivität“ bedeutet.<br>Zeitpunkte der Erhebung: zu Beginn der Studie (Woche 0) und anschließend alle vier Wochen in der randomisierten Behandlungsphase sowie in der offenen Extensionsphase.<br>Die dargestellte Änderung des PhGA ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. |

*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 4-86: Bewertung des Verzerrungspotenzials für „PhGA“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | nein | ja | ja | hoch |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Es handelt sich um eine randomisierte, doppelblinde Studie und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde jedoch für den hier deskriptiv dargestellten Endpunkt „Globale Einschätzung des Arztes (PhGA)“ nicht adäquat umgesetzt, da Patienten mit fehlenden Werten nicht berücksichtigt wurden. Für die Analyse des Parameters „PhGA“ ergibt sich somit ein Hinweis für ein mögliches Verzerrungspotential.

*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 4-87: Ergebnisse für „PhGA (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **PhGA –Basiswert [mm]** | | **PhGA in Woche 24 [mm]** | | **absolute Änderung im Vergleich zum Basiswert [mm]** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 76 | 64,67 (15,82) | 76 | 16,80 (16,43) | -47,87 (21,09) |
| Adalimumab (N = 82) | 70 | 67,74 (16,24) | 70 | 28,16 (19,51) | -39,59 (24,03) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 31 | 66,16 (13,67) | 31 | 18,03 (19,33) | -48,13 (22,70) |
| Adalimumab (N = 35) | 31 | 68,06 (18,43) | 31 | 30,84 (21,92) | -37,23 (26,32) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 45 | 63,64 (17,22) | 45 | 15,96 (14,27) | -47,69 (20,17) |
| Adalimumab (N = 47) | 39 | 67,49 (14,51) | 39 | 26,03 (17,37) | -41,46 (22,21) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; PhGA: *Physicians Global Assessment*; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 29 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Die globale Einschätzung der Krankheitsaktivität durch den behandelnden Arzt stellt keinen patientenrelevanten Endpunkt dar und wird deshalb lediglich aus Transparenzgründen deskriptiv dargestellt.

In der relevanten Studienpopulation zeigte sich nach 24-wöchiger Behandlung ein numerischer Vorteil für Sarilumab gegenüber Adalimumab in der Verringerung der Krankheitsaktivität laut globaler Einschätzung des Arztes (siehe Tabelle 4-87). Dieser Vorteil zeigte sich auch in den Teilpopulationen 2 und 3.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.5.5 Globale Einschätzung des Patienten (PtGA)

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

Tabelle 4-88: Operationalisierung von „PtGA“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die Patienten schätzten ihre Krankheitsaktivität auf einer 100 mm-VAS ein, wobei 0 „keine Krankheitsaktivität“ und 100 „höchstmögliche Krankheitsaktivität“ bedeutet.<br>Zeitpunkte der Erhebung: zu Beginn der Studie (Woche 0) und anschließend alle vier Wochen in der randomisierten Behandlungsphase sowie in der offenen Extensionsphase.<br>Die dargestellte Änderung des PtGA ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0.<br>Der Anteil der Patienten mit einer Verbesserung um mindestens 18,4 mm auf der VAS wurde post hoc für Woche 24 (randomisierte Behandlungsphase) und Woche 48 (offene Extensionsphase) bestimmt. |

*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 4-89: Bewertung des Verzerrungspotenzials für „PtGA“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „PtGA“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-90: Ergebnisse für „PtGA (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **PtGA - Basiswert [mm]** | | **PtGA in Woche 24 [mm]** | | **absolute Änderung im Vergleich zum Basiswert [mm]** | **Behandlungs-differenz[a, b] [mm]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 69,20 (17,75) | 76 | 31,82 (20,20) | -37,38 (25,18) | -12,677 [-19,486; -5,868] 0,0003 **-0,597 [-0,917; -0,276]** | |
| Adalimumab (N = 82) | 70 | 68,86 (19,81) | 70 | 45,04 (21,57) | -23,81 (25,51) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 70,42 (16,96) | 31 | 33,65 (22,86) | -36,77 (27,05) | -12,540 [-23,784; -1,297] 0,0294 **-0,550 [-1,043; -0,057]** | |
| Adalimumab (N = 35) | 31 | 63,45 (22,78) | 31 | 45,00 (21,59) | -18,45 (25,00) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,6623 |
| Sarilumab (N = 52) | 45 | 68,36 (18,42) | 45 | 30,56 (18,31) | -37,80 (24,11) | -12,609 [-21,560; -3,568] 0,0063 **-0,606 [-1,036; -0,176]** | |
| Adalimumab (N = 47) | 39 | 73,15 (16,12) | 39 | 45,08 (21,82) | -28,08 (25,43) | | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; PtGA: *Patients Global Assessment*; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 29 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24-wöchiger Behandlung zeigte sich hinsichtlich der globalen Einschätzung der Patienten ihrer Krankheitsaktivität (PtGA) ein Vorteil von Sarilumab gegenüber Adalimumab (siehe Tabelle 4-90).

In der relevanten Studienpopulation ergab sich ein statistisch signifikanter Behandlungsunterschied zugunsten von Sarilumab (LSMD [95 % KI] p-Wert: -12,677 [-19,486; -5,868] 0,0003). Die Betrachtung der standardisierten Mittelwertdifferenz belegt, dass es sich dabei um einen klinisch relevanten Effekt mit kleiner Effektstärke handelt (Hedges´g [95% KI]: -0,597 [-0,917; -0,276]).

Bei den Patienten der Teilpopulationen 2 und 3 zeigte sich ebenfalls unter Sarilumab im Vergleich zu Adalimumab eine statistisch signifikant stärkere Verbesserung gemäß globaler Einschätzung der Patienten (LSMD [95% KI] p-Wert: -12,540 [-23,784; -1,297] 0,0294 für Teilpopulation 2 bzw. -12,609 [-21,560; -3,568] 0,0063 für Teilpopulation 3). Die Konfidenzintervalle der standardisierten Mittelwertdifferenzen schließen jedoch in beiden Fällen den Wert -0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt.

Der Interaktionstest ergab keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „PtGA (Woche 24)“ das Ergebnis der relevanten Studienpopulation betrachtet werden kann.

In einer Responderanalyse wurde des Weiteren untersucht, wie viele Patienten eine Verbesserung beim PtGA um mindestens 18,4 mm erreichten. Dieser Schwellenwert wurde für Patienten mit rheumatoider Arthritis als minimale klinisch relevante Verbesserung validiert [25]. Bei Betrachtung der Mittelwertdifferenzen zeigt sich jedoch, dass die mittlere Änderung im PtGA bereits deutlich über dem definierten Schwellenwert lag (bei den Patienten der relevanten Studienpopulation ca. 37,4 mm bzw. 23,8 mm bei Behandlung mit Sarilumab bzw. Adalimumab).

Demensprechend erreichten in der relevanten Studienpopulation 50,0 % der Patienten unter Adalimumab und 60,9 % der Patienten unter Sarilumab eine Verbesserung um 18,4 mm (siehe Tabelle 4-91). Diese Gruppendifferenz ist statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,18 [0,90; 1,55] 0,2196). In den Teilpopulationen 2 und 3 ergaben sich vergleichbare Ergebnisse.

Die Interaktionstests zeigten für keinen der Effektschätzer keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „PtGA-Responder (Woche 24)“ das Ergebnis der relevanten Studienpopulation betrachtet werden kann.

Tabelle 4-91: Ergebnisse für „PtGA-Responder (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Verbesserung ≥ 18,4 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 53 (60,9) | 1,60 [0,85; 3,01] 0,1411 | 1,18 [0,90; 1,55] 0,2196 | 0,11 [-0,04; 0,26] 0,1423 | |
| Adalimumab (N = 82) | 41 (50,0) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,4685 RR: 0,8098 RD: 0,5203 |
| Sarilumab (N = 35) | 22 (62,9) | 2,11 [0,78; 5,75] 0,1414 | 1,22 [0,78; 1,91] 0,3811 | 0,16 [-0,06; 0,39] 0,1558 | |
| Adalimumab (N = 35) | 16 (45,7) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 31 (59,6) | 1,31 [0,57; 3,00] 0,5188 | 1,14 [0,81; 1,61] 0,4604 | 0,07 [-0,13; 0,26] 0,5072 | |
| Adalimumab (N = 47) | 25 (53,2) | | | | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; OR: Odds Ratio; PtGA: *Patients Global Assessment*; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Auswertung der Anzahl der Patienten mit einer Verbesserung beim PtGA um mindestens 18,4 mm nach 48 Wochen Behandlung mit Sarilumab zeigte ein einen leichten Anstieg im Vergleich zur Auswertung nach Woche 24 (siehe Tabelle 4-92). Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

Tabelle 4-92: Ergebnisse für „PtGA-Responder (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 18,4 mm in Woche 24 (ITT)** | | **Patienten mit Verbesserung ≥ 18,4 mm in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 53 (60,9) | 75 | 56 (74,7) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 22 (62,9) | 30 | 20 (66,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 31 (59,6) | 45 | 36 (80,0) |
| Patienten mit fehlenden Werten wurden als Non-Responder gezählt.<br>Quelle: [90] | | | | |

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.5.6 Körperlicher Funktionsstatus - HAQ-DI

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

Tabelle 4-93: Operationalisierung von „Körperlicher Funktionsstatus – HAQ-DI“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Der Funktionsstatus der Patienten wurde mittels *Health Assessment Questionnaire – Disability Index* (HAQ-DI) ermittelt. Der HAQ-DI ist ein validierter Fragebogen für Patienten mit rheumatoider Arthritis.<br>Der Fragebogen umfasst acht Kategorien: Anziehen, Aufstehen, Essen, Gehen, Körperpflege, Gegenstände erreichen, Greifen, allgemeine tägliche Aktivitäten. Es wird abgefragt, ob eine Tätigkeit ohne, mit geringer oder großer Schwierigkeit oder nicht ohne fremde Hilfe bewältigt werden kann und die Antwort entsprechend mit 0-3 bewertet. Zusätzlich wird der Einsatz von Hilfsmitteln berücksichtigt. Für den Gesamtscore ergibt sich wiederum ein Wert von 0-3, wobei hohe Werte eine stärke Bewegungs-und Funktionseinschränkung bedeuten.<br>Die Patienten füllten den HAQ-DI-Fragebogen zu Beginn der Studie (Woche 0) und anschließend alle vier Wochen in der randomisierten Behandlungsphase sowie in der offenen Extensionsphase aus.<br>Die dargestellte Änderung des HAQ-DI ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. Zusätzlich wurden Responderanalysen bezüglich einer Verbesserung um mindestens 0,22 (a priori festgelegt) sowie 0,375 (post hoc festgelegt) Punkte in Woche 24 und Woche 48 durchgeführt. |

*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 4-94: Bewertung des Verzerrungspotenzials für „Körperlicher Funktionsstatus – HAQ-DI“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Körperlicher Funktionsstatus – HAQ-DI“ ergeben sich keine Hinweise für ein mögliches

Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-95: Ergebnisse für „Änderung des HAQ-DI (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **HAQ-DI-Basiswert** | | **HAQ-DI in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 1,73 (0,56) | 76 | 1,04 (0,66) | -0,68 (0,70) | -0,214 [-0,406; -0,022] 0,0293 **-0,352 [-0,668; -0,036]** | |
| Adalimumab (N = 82) | 70 | 1,70 (0,72) | 70 | 1,28 (0,67) | -0,42 (0,61) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 1,68 (0,65) | 31 | 1,00 (0,68) | -0,68 (0,75) | -0,252 [-0,553; 0,050] 0,1002 | |
| Adalimumab (N = 35) | 31 | 1,52 (0,70) | 31 | 1,21 (0,62) | -0,31 (0,53) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,8384 |
| Sarilumab (N = 52) | 45 | 1,76 (0,49) | 45 | 1,08 (0,64) | -0,69 (0,67) | -0,177 [-0,438; 0,085] 0,1832 | |
| Adalimumab (N = 47) | 39 | 1,84 (0,72) | 39 | 1,34 (0,72) | -0,50 (0,66) | | |

HAQ-DI: *Health Assessment Questionnaire – Disability Index*; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 30 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24 Wochen Behandlung zeigte sich eine Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verbesserung der körperlichen Funktion gemessen mittels HAQ-

DI. Eine Senkung des HAQ-DI entspricht dabei einer Verbesserung des körperlichen Funktionsstatus.

In der relevanten Studienpopulation zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Senkung des HAQ-DI (LSMD [95% KI] p-Wert: -0,214 [-0,406; -0,022] 0,0293, siehe Tabelle 4-95). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g zeigt jedoch, dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt, da das Konfidenzintervall den Wert -0,2 einschließt (Hedges´g [95% KI]: -0,352 [-0,668; -0,036]).

In den Teilpopulationen 2 und 3 zeigte sich ebenfalls eine numerisch stärkere Verbesserung im HAQ-DI, jedoch sind der Gruppenunterschiede statistisch nicht signifikant (LSMD [95% KI] p-Wert in der Teilpopulation 2: -0,252 [-0,553; 0,050] 0,1002 und LSMD [95% KI] p-Wert in der Teilpopulation 3: -0,177 [-0,438; 0,085] 0,1832).

Der Interaktionstest ergab keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Änderung des HAQ-DI (Woche 24)“ die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

Des Weiteren wurde in zwei Responderanalysen der Anteil der Patienten mit einer klinisch relevanten Verbesserung der körperlichen Funktion bestimmt. Dafür wurden die validierten Schwellenwerte -0,22 und -0,375 als Kriterien verwendet [25]. Das IQWiG hat in der Vergangenheit bereits eine Verbesserung des HAQ-DI um mindestens 0,22 Punkte als adäquate individuelle klinisch relevante Verbesserung anerkannt [30].

In der relevanten Studienpopulation erreichten nach 24 Wochen Behandlung 66,7 % der Patienten unter Sarilumab und 48,8 % der Patienten unter Adalimumab eine Verbesserung im HAQ-DI um mindestens 0,22 Punkte (siehe Tabelle 4-96). Die Wahrscheinlichkeit, eine Verbesserung um 0,22 Punkte zu erreichen, war damit in der Sarilumab-Gruppe signifikant höher als in der Adalimumab-Gruppe (RR [95% KI] p-Wert: 1,37 [1,05; 1,78] 0,0213).

In den Teilpopulationen 2 und 3 ergaben sich numerisch vergleichbare Ergebnisse, jedoch sind die Gruppenunterschiede statistisch nicht signifikant (RR [95% KI] p-Wert: 1,35 [0,87; 2,08] 0,1816 in der Teilpopulation 2 und 1,34 [0,96; 1,88] 0,0847 in der Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „HAQ-DI-Responder – MCID 0,22 (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-96: Ergebnisse für „HAQ-DI-Responder – MCID 0,22 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 0,22 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 58 (66,7) | 2,15 [1,14; 4,05] **0,0180** | 1,37 [1,05; 1,78] **0,0213** | 0,18 [0,03; 0,33] **0,0159** | |
| Adalimumab (N = 82) | 40 (48,8) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,7367 RR: 0,9938 RD: 0,8461 |
| Sarilumab (N = 35) | 23 (65,7) | 2,38 [0,88; 6,48] 0,0889 | 1,35 [0,87; 2,08] 0,1816 | 0,19 [-0,04; 0,42] 0,1033 | |
| Adalimumab (N = 35) | 16 (45,7) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 35 (67,3) | 1,91 [0,83; 4,40] 0,1295 | 1,34 [0,96; 1,88] 0,0847 | 0,16 [-0,03; 0,35] 0,1067 | |
| Adalimumab (N = 47) | 24 (51,1) | | | | |

HAQ-DI: *Health Assessment Questionnaire – Disability Index*; ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 16.2.6.13.9 des Appendix 16.2.6 zum Studienbericht zur randomisierten Behandlungsphase der Studie dargestellt [71].

In einer zweiten Responderanalyse wurde eine Verbesserung um 0,375 als Schwellenwert verwendet (siehe Tabelle 4-97).

In dieser Analyse zeigte sich nach 24 Wochen Behandlung ein Anteil von 62,1 % der Patienten unter Sarilumab in der relevanten Studienpopulation, die eine Verbesserung im HAQ-DI um mindestens 0,375 Punkte erreichten. Bei den Patienten unter Adalimumab waren es hingegen 41,5 % der Patienten. Dieser Gruppenunterschied war statistisch signifikant (RR [95% KI] p-Wert: 1,49 [1,11; 2,01] 0,0090).

In den Teilpopulationen 2 und 3 ergaben sich numerisch vergleichbare Ergebnisse, jedoch ist der Gruppenunterschied lediglich in der Teilpopulation 3 statistisch signifikant (RR [95% KI]

p-Wert: 1,37 [0,81; 2,33] 0,2395 in Teilpopulation 2 und 1,46 [1,02; 2,09] 0,0379 in Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „HAQ-DI-Responder – MCID 0,375 (Woche 24)“ ebenfalls die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-97: Ergebnisse für „HAQ-DI-Responder – MCID 0,375 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 0,375 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 54 (62,1) | 2,44 [1,29; 4,62] **0,0061** | 1,49 [1,11; 2,01] **0,0090** | 0,21 [0,06; 0,36] **0,0048** | |
| Adalimumab (N = 82) | 34 (41,5) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,8358<br>RR: 0,8459<br>RD: 0,9789 |
| Sarilumab (N = 35) | 19 (54,3) | 2,59 [0,90; 7,47] 0,0770 | 1,37 [0,81; 2,33] 0,2395 | 0,20 [-0,01; 0,42] 0,0637 | |
| Adalimumab (N = 35) | 12 (34,3) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 35 (67,3) | 2,25 [0,98; 5,19] 0,0568 | 1,46 [1,02; 2,09] **0,0379** | 0,20 [0,01; 0,39] **0,0450** | |
| Adalimumab (N = 47) | 22 (46,8) | | | | |

HAQ-DI: *Health Assessment Questionnaire – Disability Index*; ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bezüglich einer Verbesserung des körperlichen Funktionsstatus um mindestens 0,22 bzw. 0,375 Punkte beim HAQ-DI zeigten lediglich einen leichten Rückgang der Responderzahlen in der relevanten Studienpopulation im Vergleich zur Auswertung nach Woche 24 (siehe Tabelle 4-98 und Tabelle 4-99). Etwa 90 % der Responder in Woche 24 konnten ihre Verbesserung des

Funktionsstatus gemäß HAQ-DI beibehalten (MCID 0,22: 52 Responder in Woche 48 vs. 58 Responder in Woche 24 bzw. MCID 0,375: 49 Responder in Woche 48 vs. 54 Responder in Woche 24).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

Tabelle 4-98: Ergebnisse für „HAQ-DI-Responder – MCID 0,22 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 0,22 in Woche 24 (ITT)** | | **Patienten mit Verbesserung ≥ 0,22 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 58 (66,7) | 75 | 52 (69,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 23 (65,7) | 30 | 22 (73,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 35 (67,3) | 45 | 30 (66,7) |
| HAQ-DI: *Health Assessment Questionnaire – Disability Index*<br>Patienten mit fehlenden Werten wurden als Non-Responder gezählt.<br>Quelle: [90] | | | | |

Tabelle 4-99: Ergebnisse für „HAQ-DI-Responder – MCID 0,375 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 0,375 in Woche 24 (ITT)** | | **Patienten mit Verbesserung ≥ 0,375 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 54 (62,1) | 75 | 49 (65,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 19 (54,3) | 30 | 20 (66,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 35 (67,3) | 45 | 29 (64,4) |
| HAQ-DI: *Health Assessment Questionnaire – Disability Index*<br>Patienten mit fehlenden Werten wurden als Non-Responder gezählt.<br>Quelle: [90] | | | | |

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Körperlicher Funktionsstatus – HAQ-DI“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.5.7 Laborparameter: CRP und ESR

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

Tabelle 4-100: Operationalisierung von „Laborparameter: CRP und ESR“

| Studie | Operationalisierung |
|---|---|
| MONARCH | – CRP: Die Bestimmung des CRP-Werts als Entzündungsparameter wurde anhand einer Nüchternblutprobe in einem Zentrallabor durchgeführt.<br>– ESR: Die Bestimmung des Entzündungsparameters ESR (in mm/h) wurde lokal im jeweiligen Studienzentrum anhand einer Nüchternblutprobe durchgeführt.<br>Die Zeitpunkte der Erhebung der jeweiligen Parameter sind im Abschnitt zur Operationalisierung der jeweiligen Scores (DAS28-ESR, SDAI, ACR20/50/70) dargestellt.<br>Die dargestellte Änderung der Laborparameter ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. |

*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 4-101: Bewertung des Verzerrungspotenzials für „Laborparameter: CRP und ESR“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | nein | ja | ja | hoch |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Es handelt sich um eine randomisierte, doppelblinde Studie und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde jedoch für die hier deskriptiv dargestellten Laborparameter „CRP“ und „ESR“ nicht adäquat umgesetzt, da Patienten mit fehlenden Werten nicht berücksichtigt wurden. Für die Analyse der Laborparameter ergibt sich somit ein Hinweis für ein mögliches Verzerrungspotential.

*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 4-102: Ergebnisse für „CRP (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **CRP –Basiswert [mg/L]** | | **CRP in Woche 24 [mg/L]** | | **absolute Änderung im Vergleich zum Basiswert [mg/L]** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 74 | 1,65 (2,17) | 74 | 0,19 (0,61) | -1,47 (1,98) |
| Adalimumab (N = 82) | 68 | 2,39 (3,05) | 68 | 1,75 (2,56) | -0,64 (2,12) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 30 | 1,84 (2,79) | 30 | 0,19 (0,79) | -1,65 (2,32) |
| Adalimumab (N = 35) | 30 | 2,12 (3,15) | 30 | 1,52 (2,63) | -0,60 (1,32) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 44 | 1,53 (1,64) | 44 | 0,18 (0,45) | -1,34 (1,73) |
| Adalimumab (N = 47) | 38 | 2,61 (2,99) | 38 | 1,94 (2,53) | -0,67 (2,60) |

CRP: C-reaktives Protein; ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 29 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Die Änderung der CRP-Konzentration stellt als Laborparameter keinen patientenrelevanten Endpunkt dar und wird deshalb lediglich aus Transparenzgründen deskriptiv dargestellt.

Sowohl in der relevanten Studienpopulation als auch in den Teilpopulationen 2 und 3 zeigte sich nach 24-wöchiger Behandlung eine deutlich stärkere Senkung des CRP-Wertes bei den mit Sarilumab behandelten Patienten im Vergleich zu den mit Adalimumab behandelten Patienten (siehe Tabelle 4-102).

Tabelle 4-103: Ergebnisse für „ESR (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **ESR-Basiswert [mm/h]** | | **ESR in Woche 24 [mm/h]** | | **absolute Änderung im Vergleich zum Basiswert [mm/h]** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 76 | 49,37 (24,12) | 76 | 13,17 (12,58) | -36,20 (24,91) |
| Adalimumab (N = 82) | 71 | 46,54 (20,72) | 71 | 34,63 (24,59) | -11,90 (22,65) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 31 | 47,29 (21,32) | 31 | 14,00 (11,94) | -33,29 (22,58) |
| Adalimumab (N = 35) | 31 | 44,26 (16,09) | 31 | 31,97 (21,17) | -12,29 (21,57) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 45 | 50,80 (26,02) | 45 | 12,60 (13,10) | -38,20 (26,46) |
| Adalimumab (N = 47) | 40 | 48,30 (23,76) | 40 | 36,70 (27,02) | -11,60 (23,72) |

ESR: Erythrozytensedimentationsrate; ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 29 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Die Änderung der ESR stellt als Laborparameter keinen patientenrelevanten Endpunkt dar und wird deshalb lediglich aus Transparenzgründen deskriptiv dargestellt.

Sowohl in der relevanten Studienpopulation als auch in den Teilpopulationen 2 und 3 zeigte sich nach 24-wöchiger Behandlung eine deutlich stärkere Senkung des ESR-Wertes bei den mit Sarilumab behandelten Patienten im Vergleich zu den mit Adalimumab behandelten Patienten (siehe Tabelle 4-103).

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.6 Morbidität: Morgensteifigkeit – RCT

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

Tabelle 4-104: Operationalisierung von „Körperlicher Funktionsstatus – Morgensteifigkeit“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die Patienten schätzten auf einer 100 mm-VAS ein, wie stark sie durch Morgensteifigkeit in der zurückliegenden Woche beeinträchtigt waren, wobei 0 „kein Problem bzw. keine Einschränkung“ und 100 „großes Problem bzw. große Einschränkung“ bedeutete.<br>Die Morgensteifigkeit wurde zu Beginn der Studie (Woche 0) sowie in der randomisierten Behandlungsphase in Woche 12 und Woche 24 sowie in der offenen Extensionsphase in Woche 48 erhoben.<br>Die dargestellte Änderung der Morgensteifigkeit ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. |

*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 4-105: Bewertung des Verzerrungspotenzials für „Körperlicher Funktionsstatus – Morgensteifigkeit“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Funktionsstatus – Morgensteifigkeit“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-106: Ergebnisse für „Körperlicher Funktionsstatus – Morgensteifigkeit (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **Morgensteifigkeit-Basiswert [mm]** | | **Morgensteifigkeit in Woche 24 [mm]** | | **absolute Änderung im Vergleich zum Basiswert [mm]** | **Behandlungsdifferenz[a, b] [mm]** | **Interaktionstest[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 73,16 (19,48) | 76 | 33,08 (24,22) | -40,08 (29,74) | -10,967 [-18,841; -3,094] 0,0066 **-0,456 [-0,783; -0,129]** | |
| Adalimumab (N = 82) | 70 | 67,43 (23,68) | 70 | 42,97 (25,09) | -24,46 (27,37) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 70,71 (23,87) | 31 | 36,71 (25,68) | -34,00 (31,78) | -7,056 [-19,041; 4,928] 0,2436 | |
| Adalimumab (N = 35) | 31 | 63,03 (25,33) | 31 | 42,48 (23,14) | -20,55 (25,73) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,4186 |
| Sarilumab (N = 52) | 45 | 74,84 (15,83) | 45 | 30,58 (23,12) | -44,27 (27,84) | -12,903 [-23,758; -2,049] 0,0204 **-0,519 [-0,956; -0,082]** | |
| Adalimumab (N = 47) | 39 | 70,92 (21,98) | 39 | 43,36 (26,83) | -27,56 (28,54) | | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 69 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24-wöchiger Behandlung zeigte sich ein Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verminderung der Morgensteifigkeit (siehe Tabelle 4-106).

In der relevanten Studienpopulation führte die Behandlung mit Sarilumab zu einer signifikant stärkeren Verringerung der Morgensteifigkeit als die Behandlung mit Adalimumab (LSMD [95 % KI] p-Wert: -10,967 [-18,841; -3,094] 0,0066). Die klinische Relevanz dieses Gruppenunterschieds kann jedoch aus der vorliegenden Analyse nicht abgeleitet werden, da das Konfidenzintervall der standardisierten Mittelwertdifferenz Hedges´g den Wert -0,2 einschließt (Hedges´g [95 % KI]: -0,456 [-0,783; -0,129]).

In den Teilpopulationen 2 und 3 zeigten sich numerisch vergleichbare Ergebnisse. Allerdings ist lediglich der Gruppenunterschied in der Teilpopulation 3 statistisch signifikant (LSMD [95 % KI] p-Wert: -7,056 [-19,041; 4,928] 0,2436 für Teilpopulation 2 bzw. -12,903 [-23,758; -2,049] 0,0204 für Teilpopulation 3). Das Konfidenzintervall zu Hedges´g schließt jedoch den Wert -0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt.

Der Interaktionstest ergab keine Effektmodifikation für die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Körperlicher Funktionsstatus – Morgensteifigkeit (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Körperlicher Funktionsstatus – Morgensteifigkeit“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.7 Morbidität: Fatigue – RCT

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

Tabelle 4-107: Operationalisierung von „Fatigue“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Das Ausmaß der Fatigue wurde mit Hilfe des Fragebogens FACIT-Fatigue (*Functional Assessment of Chronic Illness Therapy Fatigue*) bestimmt. Der Fragebogen umfasst 13 Fragen bezüglich der Beeinträchtigung durch Müdigkeit und Schwäche. Der FACIT-Fatigue kann Werte zwischen 0 und 52 annehmen, wobei höhere Werte eine geringere Fatigue bedeuten.<br>Die Fatigue wurde zu Beginn der Studie (Woche 0) sowie in der randomisierten Behandlungsphase in Woche 12 und Woche 24 sowie in der offenen Extensionsphase in Woche 48 erhoben.<br>Die dargestellte Änderung der Fatigue ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. Zusätzlich wurde post hoc eine Responderanalyse bezüglich einer Verbesserung um mindestens 4 Punkte in Woche 24 und 48 durchgeführt. |

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

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

Tabelle 4-108: Bewertung des Verzerrungspotenzials für „Fatigue“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Fatigue“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

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

Tabelle 4-109: Ergebnisse für „Fatigue (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **FACIT-Fatigue-Basiswert** | | **FACIT-Fatigue-Wert in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 23,49 (9,66) | 76 | 35,30 (9,61) | 11,81 (10,74) | 3,711 [0,814; 6,608] 0,0124 **0,415 [0,091; 0,740]** | |
| Adalimumab (N = 82) | 70 | 23,80 (10,92) | 70 | 31,97 (10,07) | 8,17 (10,65) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 22,68 (10,45) | 31 | 34,97 (8,99) | 12,29 (11,29) | 4,074 [-0,203; 8,352] 0,0616 | |
| Adalimumab (N = 35) | 31 | 26,08 (11,22) | 31 | 32,40 (9,76) | 6,32 (9,84) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,8570 |
| Sarilumab (N = 52) | 45 | 24,05 (9,16) | 45 | 35,54 (10,11) | 11,49 (10,46) | 3,234 [-0,882; 7,350] 0,1219 | |
| Adalimumab (N = 47) | 39 | 21,99 (10,46) | 39 | 31,63 (10,42) | 9,64 (11,15) | | |

FACIT-Fatigue: *Functional Assessment of Chronic Illness Therapy Fatigue;* ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 39 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24 Wochen Behandlung zeigte sich eine Überlegenheit von Sarilumab gegenüber Adalimumab hinsichtlich der Verringerung der Fatigue (siehe Tabelle 4-109).

In der relevanten Studienpopulation erreichten die Patienten nach 24-wöchiger Behandlung in beiden Studienarmen im Mittel höhere Werte für den FACIT-Fatigue-Fragebogen als zu Studienbeginn und somit eine Verringerung ihrer Fatigue. Dabei zeigte sich eine signifikant stärkere Verbesserung bei den Patienten unter Sarilumab (LSMD [95 % KI] p-Wert: 3,711 [0,814; 6,608] 0,0124). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges´g gibt jedoch keinen Aufschluss über die klinische Relevanz dieses Behandlungsunterschieds, da das Konfidenzintervall den Wert 0,2 einschließt (Hedges´g [95 % KI]: 0,415 [0,091; 0,740]).

In den Teilpopulationen 2 und 3 zeigten sich lediglich numerische Vorteile für Sarilumab gegenüber Adalimumab (LSMD [95 % KI] p-Wert: 4,074 [-0,203; 8,352] 0,0616 in Teilpopulation 2 bzw. 3,234 [-0,882; 7,350] 0,1219 in Teilpopulation 3).

Der Interaktionstest ergab keine Effektmodifikation für die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Fatigue (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Des Weiteren wurde eine Responderanalyse durchgeführt, um den Anteil der Patienten, die eine Verbesserung im FACIT-Fatigue um mindestens 4 Punkte erreicht haben, zu bestimmen. Dieser Schwellenwert wurde für Patienten mit rheumatoider Arthritis als klinisch relevante Verbesserung definiert [55].

In der relevanten Studienpopulation erzielten nach 24-wöchiger Behandlung 66,7 % der Patienten unter Sarilumab eine Verbesserung im FACIT-Fatigue um mindestens 4 Punkte, während es im Adalimumab-Arm 52,4 % der Patienten waren. Bei Betrachtung des RR ist dieser Gruppenunterschied statistisch knapp nicht signifikant (RR [95 % KI] p-Wert: 1,27 [0,99; 1,64] 0,0631).

In der Teilpopulation 2 zeigte sich ein deutlicherer Vorteil für Sarilumab. Während im Sarilumab-Arm 68,6 % der Patienten eine Verbesserung um mindestens 4 Punkte erreichten, waren es im Adalimumab-Arm lediglich 40,0 %. Dieser Gruppenunterschied war statistisch signifikant (RR [95 % KI] p-Wert: 1,62 [1,00; 2,62] 0,0494).

In der Teilpopulation 3 zeigte sich hingegen kein signifikanter Unterschied zwischen den Behandlungsarmen hinsichtlich der Verbesserung um mindestens 4 Punkte im FACIT-Fatigue (RR [95 % KI] p-Wert: 1,05 [0,77; 1,42] 0,7696).

Die Interaktionstests hinsichtlich einer Effektmodifikation durch die Anzahl der csDMARD-Vortherapien ergaben für das RR keinen Beleg für eine Effektmodifikation. Somit kann für den Endpunkt „Fatigue-Responder (Woche 24)“ das Ergebnis der relevanten Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden.

Tabelle 4-110: Ergebnisse für „Fatigue-Responder (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 4 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 58 (66,7) | 1,86 [0,99; 3,50] 0,0533 | 1,27 [0,99; 1,64] 0,0631 | 0,15 [-0,001; 0,30] 0,0522 | |
| Adalimumab (N = 82) | 43 (52,4) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,0887 RR: 0,1323 RD: 0,1034 |
| Sarilumab (N = 35) | 24 (68,6) | 3,48 [1,25; 9,68] **0,0170** | 1,62 [1,00; 2,62] **0,0494** | 0,27 [0,05; 0,50] **0,0185** | |
| Adalimumab (N = 35) | 14 (40,0) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 34 (65,4) | 1,10 [0,47; 2,57] 0,8185 | 1,05 [0,77; 1,42] 0,7696 | 0,03 [-0,17; 0,22] 0,7967 | |
| Adalimumab (N = 47) | 29 (61,7) | | | | |

FACIT-Fatigue: *Functional Assessment of Chronic Illness Therapy Fatigue*; ITT: Intent-to treat; n/N: Anzahl der Patienten mit Ereignis/Anzahl der Patienten mit verfügbaren Daten; KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Auswertung nach 48 Wochen Behandlung mit Sarilumab bezüglich einer Verbesserung der Fatigue um mindestens 4 Punkte beim FACIT-Fatigue zeigte lediglich einen leichten Rückgang der Responderzahlen in der relevanten Studienpopulation im Vergleich zur Auswertung nach Woche 24 (siehe Tabelle 4-111). Etwa 93 % der Responder in Woche 24 konnten ihre Verbesserung der Fatigue gemäß FACIT-Fatigue beibehalten (54 Responder in Woche 48 vs. 58 Responder in Woche 24).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab.

Tabelle 4-111: Ergebnisse für „Fatigue-Responder (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 4 in Woche 24 (ITT)** | | **Patienten mit Verbesserung ≥ 4 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 58 (66,7) | 75 | 54 (72,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 24 (68,6) | 30 | 22 (73,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 34 (65,4) | 45 | 32 (71,1) |

FACIT-Fatigue: *Functional Assessment of Chronic Illness Therapy Fatigue*
Patienten mit fehlenden Werten wurden als Non-Responder gezählt.
Quelle: [90]

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Fatigue“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.8 Morbidität: Änderung des allgemeinen Gesundheitszustandes – RCT

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

Tabelle 4-112: Operationalisierung von „Änderung des allgemeinen Gesundheitszustandes“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die Änderung des allgemeinen Gesundheitszustands wurde über die VAS des EQ-5D ermittelt. Der EQ-5D ist ein generischer Fragebogen, welcher neben fünf Fragen zur Ermittlung der Lebensqualität eine VAS zur Bestimmung des allgemeinen Gesundheitszustands umfasst. Die VAS reicht von 0 („schlechtester möglicher Zustand“) bis 100 („bestmöglicher Zustand“).<br>Der allgemeine Gesundheitszustand mittels EQ-5D-VAS wurde zu Beginn der Studie (Woche 0) sowie in der randomisierten Behandlungsphase in Woche 12 und Woche 24 sowie in der offenen Extensionsphase in Woche 48 erhoben.<br>Die dargestellte Änderung des allgemeinen Gesundheitszustands ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. |

*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 „Änderung des allgemeinen Gesundheitszustandes“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Änderung des allgemeinen Gesundheitszustandes“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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 4-114: Ergebnisse für „Änderung des allgemeinen Gesundheitszustandes (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **EQ-5D-VAS-Basiswert** | | **EQ-5D-VAS in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 76 | 42,96 (21,54) | 76 | 68,20 (20,24) | 25,24 (26,39) | 9,244 [2,683; 15,806] 0,0061 | |
| Adalimumab (N = 82) | 68 | 40,37 (20,91) | 68 | 58,18 (20,45) | 17,81 (26,20) | **0,461 [0,134; 0,788]** | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 31 | 43,16 (21,73) | 31 | 62,97 (24,53) | 19,81 (28,26) | 3,237 [-7,740; 14,214] 0,5574 | |
| Adalimumab (N = 35) | 29 | 43,55 (22,96) | 29 | 58,48 (20,77) | 14,93 (22,09) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,1128 |
| Sarilumab (N = 52) | 45 | 42,82 (21,66) | 45 | 71,80 (15,99) | 28,98 (24,64) | 13,885 [5,771; 21,998] 0,0010 | |
| Adalimumab (N = 47) | 39 | 38,00 (19,22) | 39 | 57,95 (20,47) | 19,95 (28,98) | **0,746 [0,310; 1,182]** | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung; VAS: visuelle Analogskala

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 68 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24-wöchiger Behandlung zeigte sich ein Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verbesserung des allgemeinen Gesundheitszustandes gemessen mittels EQ-5D-VAS (siehe Tabelle 4-114).

In der relevanten Studienpopulation erreichten die Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine signifikant stärkere Verbesserung ihres allgemeinen Gesundheitszustandes (LSMD [95 % KI] p-Wert: 9,244 [2,683; 15,806] 0,0061). Das Konfidenzintervall zu Hedges´g schließt jedoch den Wert 0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt (Hedges´g [95 % KI]: 0,461 [0,134; 0,788]).

In der Teilpopulation 2 zeigte sich ein geringerer Unterschied zu Gunsten von Sarilumab zwischen den Behandlungsgruppen hinsichtlich der Verbesserung des allgemeinen Gesundheitszustands, der Behandlungsunterschied ist nicht statistisch signifikant (LSMD [95 % KI] p-Wert: 3,237 [-7,740; 14,214] 0,5574).

In der Teilpopulation 3 zeigte sich wiederum ein sehr ähnliches Ergebnis wie in der relevanten Studienpopulation, mit statistisch signifikantem Gruppenunterschied (LSMD [95 % KI] p-Wert: 13,885 [5,771; 21,998] 0,0010). Die Betrachtung des Hedges´g belegt, dass es sich um einen klinisch relevanten Effekt von kleiner Effektstärke handelt (Hedges´g [95 % KI]: 0,746 [0,310; 1,182]).

Der Interaktionstest auf Effektmodifikation durch die Anzahl der csDMARD-Vortherapien ergab für den Endpunkt „Änderung des allgemeinen Gesundheitszustandes (Woche 24)“ keinen Beleg für eine Effektmodifikation. Für die Ableitung des Zusatznutzens wird dementsprechend für den Endpunkt „Änderung des allgemeinen Gesundheitszustandes (Woche 24)“ nur die relevante Studienpopulation herangezogen.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Änderung des allgemeinen Gesundheitszustandes“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.9 Lebensqualität – RCT

##### 4.3.1.3.9.1 SF-36

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

Tabelle 4-115: Operationalisierung von „Lebensqualität – SF-36“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die Veränderung der Lebensqualität wurde mit dem generischen *Medical-Outcome-Study-36-Item-Short-Form*-Gesundheitsfragenbogen (SF-36, Version 2) ermittelt.<br>Der SF-36 beinhaltet acht Dimensionen: körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperlicher Schmerz, allgemeine Gesundheitswahrnehmung, Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden. In der Studie wurden zunächst die Rohwerte der einzelnen Dimensionen bestimmt und in eine 0–100-Skala transformiert, wobei höhere Werte ein besseres Befinden bedeuten. Die einzelnen Dimensionen wurden anschließend von 0–100-Scores in normbasierte 0–100-Scores transformiert (allgemeine US-Population von 1998). Anschließend wurden die körperliche und die psychische Summenskala (PCS und MCS, *physical* und *mental component summary scores*) bestimmt. Die PCS berücksichtigt die Dimensionen körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperlicher Schmerz und allgemeine Gesundheitswahrnehmung. Die MCS berücksichtigt die Dimensionen Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden.<br>Die Lebensqualität mittels SF-36 wurde zu Beginn der Studie (Woche 0) sowie in der randomisierten Behandlungsphase in Woche 12 und Woche 24 sowie in der offenen Extensionsphase in Woche 48 erhoben.<br>Die dargestellte Änderung der Lebensqualität bezüglich PCS und MCS ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. Zusätzlich wurden post hoc Responderanalysen bezüglich einer Verbesserung der PCS um mindestens 2,5 sowie 5,1 und 7,2 Punkte in Woche 24 und Woche 48 durchgeführt. |

*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 4-116: Bewertung des Verzerrungspotenzials für „Lebensqualität – SF-36“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen. Für die Analysen des Endpunktes „Lebensqualität – SF-36“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential.

*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.*

### *Körperliche Summenskala (SF-36-PCS)*

Tabelle 4-117: Ergebnisse für „Änderung des SF-36-PCS (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **SF-36-PCS-Basiswert** | | **SF-36-PCS in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 72 | 29,92 (5,61) | 72 | 38,99 (7,87) | 9,07 (7,44) | 3,640 [1,399; 5,882] 0,0016 **0,532 [0,204; 0,860]** | |
| Adalimumab (N = 82) | 70 | 30,53 (6,09) | 70 | 35,96 (7,47) | 5,43 (6,68) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 29 | 30,08 (4,86) | 29 | 37,68 (7,71) | 7,61 (7,16) | 2,516 [-1,044; 6,077] 0,1627 | |
| Adalimumab (N = 35) | 31 | 32,14 (6,87) | 31 | 37,00 (8,38) | 4,86 (6,97) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,3626 |
| Sarilumab (N = 52) | 43 | 29,82 (6,11) | 43 | 39,87 (7,95) | 10,05 (7,54) | 4,379 [1,377; 7,381] 0,0048 **0,637 [0,200; 1,073]** | |
| Adalimumab (N = 47) | 39 | 29,24 (5,12) | 39 | 35,14 (6,66) | 5,89 (6,50) | | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; PCS: *Physical Component Score*; SF-36: *Short-Form 36;* SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 35 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Die Auswertung der körperlichen Summenskala des SF-36 (SF-36-PCS) ergab nach 24-wöchiger Behandlung einen Vorteil von Sarilumab gegenüber Adalimumab (siehe Tabelle 4-117). Die Einzeldomänen, die zur SF-36-PCS zusammengefasst wurden, sind in Tabelle 4-125 bis Tabelle 4-128 deskriptiv dargestellt.

In der relevanten Studienpopulation verbesserte sich die Lebensqualität gemäß SF-36-PCS unter Sarilumab signifikant stärker als unter Adalimumab (LSMD [95 % KI] p-Wert: 3,640 [1,399; 5,882] 0,0016). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges´g zeigt, dass es sich dabei um einen klinisch relevanten Effekt von kleiner Effektstärke handelt (Hedges´g [95 % KI]: 0,532 [0,204; 0,860]).

Bei den Patienten der Teilpopulation 2 zeigte sich ebenfalls ein numerischer Vorteil für Sarilumab, jedoch ist der Behandlungsunterschied zu Adalimumab nicht statistisch signifikant (LSMD [95 % KI] p-Wert: 2,516 [-1,044; 6,077] 0,1627).

In der Teilpopulation 3 führte die Behandlung mit Sarilumab wiederum zu einer signifikant stärkeren Verbesserung der Lebensqualität gemäß SF-36-PCS (LSMD [95 % KI] p-Wert: 4,379 [1,377; 7,381] 0,0048). Das Konfidenzintervall zu Hedges´g schließt jedoch den Wert 0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt.

Der Interaktionstest ergab keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Änderung des SF-36-PCS (Woche 24)“ die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

Für Patienten mit rheumatoider Arthritis wurden verschiedene MCIDs (2,5, 5,1 und 7,2 Punkte) für die körperliche Summenskala des SF-36 definiert [62]. Der Anteil der Patienten, die diese Schwellenwerte nach 24-wöchiger Behandlung erreichten, wurde in den folgenden Responderanalysen bestimmt.

In der relevanten Studienpopulation erzielten 63,2 % der Patienten unter Sarilumab eine Verbesserung der SF-36-PCS um mindestens 2,5 Punkte, während es unter Adalimumab 51,2 % der Patienten waren (siehe Tabelle 4-118). Bei Betrachtung des RR ist dieser Gruppenunterschied statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,23 [0,95; 1,59] 0,1138).

Auch in den Teilpopulationen 2 und 3 zeigten sich ebenfalls numerische Vorteile zugunsten von Sarilumab bezüglich des Anteils der Patienten, die eine Verbesserung der SF-36-PCS um mindestens 2,5 Punkte erreichten. Die Behandlungsunterschiede sind statistisch nicht signifikant (RR [95 % KI] p-Wert: 0,98 [0,59; 1,62] 0,9313 für Teilpopulation 2 und 1,34 [0,96; 1,87] 0,0840 für Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den

Endpunkt „SF-36-PCS-Responder – MCID 2,5 (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Tabelle 4-118: Ergebnisse für „SF-36-PCS-Responder – MCID 2,5 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 2,5 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 55 (63,2) | 1,79 [0,95; 3,37] 0,0720 | 1,23 [0,95; 1,59] 0,1138 | 0,13 [-0,02; 0,28] 0,0814 | |
| Adalimumab (N = 82) | 42 (51,2) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,3863 RR: 0,3059 RD: 0,3140 |
| Sarilumab (N = 35) | 20 (57,1) | 1,28 [0,48; 3,46] 0,6180 | 0,98 [0,59; 1,62] 0,9313 | 0,03 [-0,21; 0,27] 0,7997 | |
| Adalimumab (N = 35) | 18 (51,4) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 35 (67,3) | 2,28 [0,97; 5,38] 0,0587 | 1,34 [0,96; 1,87] 0,0840 | 0,19 [-0,01; 0,38] 0,0560 | |
| Adalimumab (N = 47) | 24 (51,1) | | | | |

ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; PCS: *Physical Component Score*; RR: Risk Ratio; RD: Risikodifferenz; SF-36: *Medical-Outcome-Study-36-Item-Short-Form*

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

In der nachfolgenden Tabelle 4-119 ist die Responderanalyse für eine Verbesserung der SF-36-PCS um mindestens 5,1 Punkte dargestellt.

Tabelle 4-119: Ergebnisse für „SF-36-PCS-Responder – MCID 5,1 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 5,1 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 50 (57,5) | 1,88 [1,00; 3,53] 0,0509 | 1,30 [0,97; 1,75] 0,0803 | 0,15 [0,00; 0,29] **0,0497** | |
| Adalimumab (N = 82) | 36 (43,9) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,5885 RR: 0,2671 RD: 0,4493 |
| Sarilumab (N = 35) | 18 (51,4) | 1,50 [0,53; 4,19] 0,4397 | 0,95 [0,54; 1,69] 0,8682 | 0,07 [-0,16; 0,30] 0,5397 | |
| Adalimumab (N = 35) | 15 (42,9) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 32 (61,5) | 2,15 [0,94; 4,95] 0,0707 | 1,41 [0,96; 2,07] 0,0817 | 0,19 [-0,01; 0,38] 0,0635 | |
| Adalimumab (N = 47) | 21 (44,7) | | | | |

ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; PCS: *Physical Component Score*; RR: Risk Ratio; RD: Risikodifferenz; SF-36: *Medical-Outcome-Study-36-Item-Short-Form*

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

In der relevanten Studienpopulation erzielten nach 24 Wochen Behandlung 57,5 % der Patienten unter Sarilumab eine Verbesserung um mindestens 5,1 Punkte, während es unter Adalimumab 43,9 % der Patienten waren. Dieser Behandlungsunterschied ist bei Betrachtung des RR statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,30 [0,97; 1,75] 0,0803).

Auch in den Teilpopulationen 2 und 3 zeigten sich ebenfalls numerische Vorteile zugunsten von Sarilumab bezüglich des Anteils der Patienten, die eine Verbesserung der SF-36-PCS um mindestens 5,1 Punkte erreichten. Die Behandlungsunterschiede sind statistisch nicht signifikant (RR [95 % KI] p-Wert: 0,95 [0,54; 1,69] 0,8682 für Teilpopulation 2 und 1,41 [0,96; 2,07] 0,0817 für Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den

Endpunkt „SF-36-PCS-Responder – MCID 5,1 (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

In der nachfolgenden Tabelle 4-120 ist die Responderanalyse für eine Verbesserung der SF-36-PCS um mindestens 7,2 Punkte dargestellt.

In der relevanten Studienpopulation erzielten nach 24 Wochen Behandlung 49,4 % der Patienten unter Sarilumab eine Verbesserung um mindestens 7,2 Punkte, während es unter Adalimumab 35,4 % der Patienten waren. Dieser Behandlungsunterschied ist bei Betrachtung des RR statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,38 [0,96; 1,97] 0,0788).

In den Teilpopulationen 2 und 3 sind die Anteile der Patienten, die eine Verbesserung der SF-36-PCS um mindestens 7,2 Punkte erzielten, sehr vergleichbar mit den Ergebnissen der relevanten Studienpopulation. Die Behandlungsunterschiede sind statistisch ebenfalls nicht signifikant (RR [95 % KI] p-Wert: 1,09 [0,60; 1,99] 0,7799 in der Teilpopulation 2 bzw. 1,49 [0,92; 2,42] 0,1021 in der Teilpopulation 3).

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „SF-36-PCS-Responder – MCID 7,2 (Woche 24)“ die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

Tabelle 4-120: Ergebnisse für „SF-36-PCS-Responder – MCID 7,2 (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 7,2 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 43 (49,4) | 1,93 [1,02; 3,65] **0,0443** | 1,38 [0,96; 1,97] 0,0788 | 0,16 [0,01; 0,30] **0,0368** | |
| Adalimumab (N = 82) | 29 (35,4) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,7245 RR: 0,4203 RD: 0,6685 |
| Sarilumab (N = 35) | 17 (48,6) | 1,71 [0,61; 4,79] 0,3063 | 1,09 [0,60; 1,99] 0,7799 | 0,12 [-0,10; 0,34] 0,2665 | |
| Adalimumab (N = 35) | 13 (37,1) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 26 (50,0) | 2,15 [0,92; 5,02] 0,0766 | 1,49 [0,92; 2,42] 0,1021 | 0,19 [-0,01; 0,38] 0,0598 | |
| Adalimumab (N = 47) | 16 (34,0) | | | | |

ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; PCS: *Physical Component Score*; RR: Risk Ratio; RD: Risikodifferenz; SF-36: *Medical-Outcome-Study-36-Item-Short-Form*

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Auswertung nach 48 Wochen Behandlung mit Sarilumab bezüglich einer Verbesserung der körperlichen Summenskala des SF-36 um mindestens 2,5 Punkte zeigte einen leichten Rückgang der Responderzahlen im Vergleich zu den Auswertungen nach Woche 24, wohingegen die Auswertungen bezüglich einer Verbesserung um mindestens 5,1 oder 7,2 Punkte etwa gleichbleibende Responderzahlen im Vergleich zu den Auswertungen nach Woche 24 ergaben (siehe Tabelle 4-121 bis Tabelle 4-123).

Die Auswertungen nach 48 Wochen belegen somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab auch in Bezug auf die Verbesserung der gesundheitsbezogenen Lebensqualität.

Tabelle 4-121: Ergebnisse für „SF-36-PCS-Responder – MCID 2,5 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 2,5 in Woche 24 (ITT)** | | **Patienten mit Verbesserung ≥ 2,5 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 55 (63,2) | 75 | 50 (66,7) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 20 (57,1) | 30 | 19 (63,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 35 (67,3) | 45 | 31 (68,9) |

SF-36: *Medical-Outcome-Study-36-Item-Short-Form*; PCS: *Physical Component Score*
Patienten mit fehlenden Werten wurden in der Auswertung als Non-Responder berücksichtigt.
Quelle: [90]

Tabelle 4-122: Ergebnisse für „SF-36-PCS-Responder – MCID 5,1 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 5,1 in Woche 24 (ITT)** | | **Patienten mit Verbesserung ≥ 5,1 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 50 (57,5) | 75 | 48 (64,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 18 (51,4) | 30 | 19 (63,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 32 (61,5) | 45 | 29 (64,4) |

SF-36: *Medical-Outcome-Study-36-Item-Short-Form*; PCS: *Physical Component Score*
Patienten mit fehlenden Werten wurden in der Auswertung als Non-Responder berücksichtigt.
Quelle: [90]

Tabelle 4-123: Ergebnisse für „SF-36-PCS-Responder – MCID 7,2 (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 7,2 in Woche 24 (ITT)** | | **Patienten mit Verbesserung ≥ 2,5 in Woche 48 (ITT)** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 43 (49,4) | 75 | 42 (56,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 17 (48,6) | 30 | 18 (60,0) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 26 (50,0) | 45 | 24 (53,3) |

SF-36: *Medical-Outcome-Study-36-Item-Short-Form*; PCS: *Physical Component Score*
Patienten mit fehlenden Werten wurden in der Auswertung als Non-Responder berücksichtigt.
Quelle: [90]

### *Psychische Summenskala (SF-36-MCS)*

Tabelle 4-124: Ergebnisse für „Änderung des SF-36-MCS (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **SF-36-MCS-Basiswert** | | **SF-36-MCS in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 72 | 37,40 (11,97) | 72 | 47,11 (10,45) | 9,71 (11,40) | 2,437 [-0,806; 5,681] 0,1397 | |
| Adalimumab (N = 82) | 70 | 34,71 (11,95) | 70 | 43,33 (11,16) | 8,61 (12,64) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 29 | 36,23 (9,19) | 29 | 47,05 (9,32) | 10,82 (8,38) | 5,321 [0,829; 9,814] 0,0210 **0,602 [0,094; 1,110)]** | 0,1403 |
| Adalimumab (N = 35) | 31 | 36,49 (11,19) | 31 | 42,02 (10,44) | 5,53 (11,25) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | |
| Sarilumab (N = 52) | 43 | 38,19 (13,58) | 43 | 47,15 (11,26) | 8,96 (13,10) | 0,210 [-4,505; 4,924] 0,9297 | |
| Adalimumab (N = 47) | 39 | 33,30 (12,48) | 39 | 44,37 (11,73) | 11,07 (13,28) | | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MCS: *Mental Component Score*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SF-36: *Medical-Outcome-Study-36-Item-Short-Form;* SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 37 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Der SF-36 ist grundsätzlich weniger geeignet, die psychischen Aspekte der Lebensqualität von Patienten mit rheumatoider Arthritis zu erfassen, da die Domänen der psychischen Gesundheit des SF-36 weniger stark mit einer Veränderung der Krankheitsaktivität assoziiert sind [61, 62].

In der Studie MONARCH zeigte die Auswertung der psychischen Summenskala des SF-36 (SF-36-MCS) nach 24 Wochen Behandlung in beiden Behandlungsarmen eine Verbesserung von vergleichbarem Ausmaß (siehe Tabelle 4-124). Die Einzeldomänen, die zur SF-36-MCS zusammengefasst wurden, sind in Tabelle 4-129 bis Tabelle 4-132 deskriptiv dargestellt.

Sowohl in der relevanten Studienpopulation (LSMD [95 % KI] p-Wert: 2,437 [-0,806; 5,681] 0,1397) als auch in der Teilpopulation 3 ergaben sich keine statistisch signifikanten Unterschiede zugunsten einer der beiden Therapien. In der Teilpopulation 2 zeigte sich ein statistisch signifikanter Vorteil für Sarilumab (LSMD [95 % KI] p-Wert: 5,321 [0,829; 9,814] 0,0210). Das Konfidenzintervall zu Hedges´g schließt jedoch den Wert -0,2 mit ein, so dass nicht ausgeschlossen werden kann, dass der Effektschätzer in einem sicher irrelevanten Bereich liegt.

Der Interaktionstest zeigte keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Änderung des SF-36-MCS (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

***Einzeldomänen des SF-36***

In den nachfolgenden Tabellen sind die Auswertungen der einzelnen Domänen des SF-36 deskriptiv dargestellt.

In allen acht Domänen zeigte sich nach 24-wöchiger Behandlung eine deutliche Verbesserung in beiden Behandlungsarmen und allen Teilpopulationen. In der Regel fiel die Verbesserung unter Sarilumab jeweils numerisch stärker aus als unter Adalimumab.

Tabelle 4-125: Ergebnisse für „SF-36 - Körperliche Funktionsfähigkeit (*Physical Function*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | | **Basiswert** | | **Wert in Woche 24** | **absolute Änderung im Vergleich zum Basiswert** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 76 | 30,86 (19,92) | 76 | 55,50 (24,10) | 24,65 (21,15) |
| Adalimumab (N = 82) | 70 | 31,75 (22,72) | 70 | 45,48 (23,71) | 13,72 (22,13) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 31 | 27,58 (19,32) | 31 | 50,97 (24,03) | 23,39 (19,30) |
| Adalimumab (N = 35) | 31 | 35,41 (24,46) | 31 | 48,55 (23,74) | 13,14 (20,04) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 45 | 33,11 (20,23) | 45 | 58,63 (23,91) | 25,52 (22,51) |
| Adalimumab (N = 47) | 39 | 28,85 (21,10) | 39 | 43,03 (23,71) | 14,19 (23,90) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in der Tabelle 36 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-126: Ergebnisse für „SF-36 - Körperliche Rollenfunktion (*Role Physical*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Basiswert** | | **Wert in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 76 | 34,32 (18,12) | 76 | 54,69 (21,99) | 20,37 (24,41) |
| Adalimumab (N = 82) | 70 | 29,11 (20,24) | 70 | 47,14 (20,90) | 18,04 (21,72) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 31 | 32,66 (18,09) | 31 | 51,01 (24,16) | 18,35 (25,10) |
| Adalimumab (N = 35) | 31 | 32,46 (23,57) | 31 | 48,39 (22,30) | 15,93 (20,59) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 45 | 35,46 (18,25) | 45 | 57,22 (20,25) | 21,76 (24,10) |
| Adalimumab (N = 47) | 39 | 26,44 (16,99) | 39 | 46,15 (19,95) | 19,71 (22,70) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in der Tabelle 36 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-127: Ergebnisse für „SF-36 - Körperlicher Schmerz (*Bodily Pain*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Basiswert** | | **Wert in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 75 | 24,61 (15,64) | 75 | 53,27 (20,03) | 28,65 (24,35) |
| Adalimumab (N = 82) | 70 | 27,81 (17,79) | 70 | 45,07 (18,92) | 17,26 (20,56) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 31 | 24,45 (15,28) | 31 | 54,68 (19,59) | 30,23 (20,42) |
| Adalimumab (N = 35) | 31 | 33,55 (19,73) | 31 | 45,61 (20,81) | 12,06 (19,36) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 44 | 24,73 (16,07) | 44 | 52,27 (20,50) | 27,55 (26,96) |
| Adalimumab (N = 47) | 39 | 23,26 (14,81) | 39 | 44,64 (17,54) | 21,38 (20,79) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in der Tabelle 36 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-128: Ergebnisse für „SF-36 - Allgemeine Gesundheitswahrnehmung (*Global Health*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Basiswert** | | **Wert in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 75 | 35,69 (17,02) | 75 | 50,56 (20,06) | 14,87 (19,17) |
| Adalimumab (N = 82) | 70 | 33,25 (16,29) | 70 | 46,60 (17,03) | 13,34 (17,69) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 30 | 34,59 (13,43) | 30 | 45,53 (18,97) | 10,94 (16,77) |
| Adalimumab (N = 35) | 31 | 36,29 (17,37) | 31 | 46,06 (16,38) | 9,77 (17,86) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 45 | 36,42 (19,16) | 45 | 53,91 (20,28) | 17,49 (20,37) |
| Adalimumab (N = 47) | 39 | 30,84 (15,18) | 39 | 47,03 (17,73) | 16,19 (17,25) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in der Tabelle 36 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-129: Ergebnisse für „SF-36 - Vitalität (*Vitality*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Basiswert** | | **Wert in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 74 | 33,53 (17,91) | 74 | 54,48 (19,81) | 20,95 (23,44) |
| Adalimumab (N = 82) | 70 | 32,23 (18,63) | 70 | 46,88 (18,48) | 14,64 (21,09) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 30 | 30,21 (16,50) | 30 | 53,96 (17,33) | 23,75 (18,96) |
| Adalimumab (N = 35) | 31 | 35,08 (17,21) | 31 | 46,98 (17,96) | 11,90 (19,25) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 44 | 35,80 (18,65) | 44 | 54,83 (21,52) | 19,03 (26,10) |
| Adalimumab (N = 47) | 39 | 29,97 (19,62) | 39 | 46,79 (19,12) | 16,83 (22,45) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in der Tabelle 37 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-130: Ergebnisse für „SF-36 - Soziale Funktionsfähigkeit (*Social Function*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Basiswert** | | **Wert in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 76 | 46,05 (26,08) | 76 | 70,89 (24,01) | 24,84 (28,69) |
| Adalimumab (N = 82) | 70 | 43,04 (26,63) | 70 | 60,00 (25,32) | 16,96 (27,99) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 31 | 46,37 (23,54) | 31 | 69,76 (23,67) | 23,39 (23,88) |
| Adalimumab (N = 35) | 31 | 49,60 (25,92) | 31 | 57,66 (25,55) | 8,06 (24,70) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 45 | 45,83 (27,95) | 45 | 71,67 (24,48) | 25,83 (31,80) |
| Adalimumab (N = 47) | 39 | 37,82 (26,36) | 39 | 61,86 (25,32) | 24,04 (28,73) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in der Tabelle 37 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-131: Ergebnisse für „SF-36 - Emotionale Rollenfunktion (*Role Emotional*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Basiswert** | | **Wert in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 76 | 48,36 (26,46) | 76 | 68,42 (24,16) | 20,07 (28,91) |
| Adalimumab (N = 82) | 70 | 41,43 (28,09) | 70 | 58,21 (24,43) | 16,79 (26,33) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 31 | 45,97 (23,26) | 31 | 66,67 (22,26) | 20,70 (26,52) |
| Adalimumab (N = 35) | 31 | 44,89 (29,16) | 31 | 56,72 (22,91) | 11,83 (25,44) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 45 | 50,00 (28,59) | 45 | 69,63 (25,57) | 19,63 (30,73) |
| Adalimumab (N = 47) | 39 | 38,68 (27,27) | 39 | 59,40 (25,80) | 20,73 (26,69) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in der Tabelle 37 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Tabelle 4-132: Ergebnisse für „SF-36 - Psychisches Wohlbefinden (*Mental Health*) (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Basiswert** | | **Wert in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 74 | 49,53 (21,37) | 74 | 66,55 (16,90) | 17,03 (20,75) |
| Adalimumab (N = 82) | 70 | 46,79 (20,32) | 70 | 63,29 (19,65) | 16,50 (23,18) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | |
| Sarilumab (N = 35) | 30 | 45,17 (17,79) | 30 | 65,83 (14,15) | 20,67 (14,78) |
| Adalimumab (N = 35) | 31 | 49,52 (21,07) | 31 | 61,94 (17,92) | 12,42 (21,44) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 44 | 52,50 (23,23) | 44 | 67,05 (18,69) | 14,55 (23,83) |
| Adalimumab (N = 47) | 39 | 44,62 (19,71) | 39 | 64,36 (21,09) | 19,74 (24,25) |

ITT: *Intent-to-treat*; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in der Tabelle 37 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Lebensqualität – SF-36“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.9.2 RAID

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

Tabelle 4-133: Operationalisierung von „Lebensqualität – RAID“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Die krankheitsspezifische Lebensqualität wurde mit dem RAID-Fragebogen (*Rheumatoid Arthritis Impact of Disease*) erhoben. Der Fragebogen besteht aus 7 numerischen Ratingskalen (NRS) von 0 bis 10, welche die Domänen Schmerzen, Beeinträchtigung der Funktionsfähigkeit, Müdigkeit/Erschöpfung, Schlaf, körperliches Wohlbefinden, psychisches Wohlbefinden und Situationsbewältigung umfassen. Der Score setzt sich aus den gewichteten Werten der 7 NRS zusammen und kann wiederum Werte zwischen 0 und 10 annehmen. Höhere Werte des RAID bilden einen schlechteren Gesundheitszustand ab.<br>Die Lebensqualität mittels RAID wurde zu Beginn der Studie (Woche 0) sowie in der randomisierten Behandlungsphase in Woche 12 und Woche 24 sowie in der offenen Extensionsphase in Woche 48 erhoben.<br>Die dargestellte Änderung des RAID ergibt sich aus der Differenz des jeweiligen Wertes in Woche 24 und des Ausgangswerts in Woche 0. Zusätzlich wurde post hoc eine Responderanalyse bezüglich einer Verbesserung um mindestens 3 Punkte in Woche 24 und Woche 48 durchgeführt. |

*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 4-134: Bewertung des Verzerrungspotenzials für „Lebensqualität – RAID“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen. Für die Analysen des Endpunktes „Lebensqualität – RAID“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential.

*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 4-135: Ergebnisse für „Änderung des RAID (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **RAID-Basiswert** | | **RAID in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a, b]** | **Interaktions-test[b] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | | | |
| Sarilumab (N = 87) | 75 | 6,65 (1,92) | 75 | 3,14 (2,03) | -3,51 (2,60) | -1,346 [-2,022; -0,671] 0,0001 **-0,651 [-0,978; -0,325]** | |
| Adalimumab (N = 82) | 69 | 6,49 (2,31) | 69 | 4,43 (2,09) | -2,06 (2,62) | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | | | |
| Sarilumab (N = 35) | 30 | 6,73 (2,13) | 30 | 3,42 (2,03) | -3,31 (2,77) | -1,306 [-2,397; -0,216] 0,0197 **-0,609 [-1,117; -0,101]** | |
| Adalimumab (N = 35) | 30 | 5,87 (2,55) | 30 | 4,47 (2,23) | -1,40 (2,60) | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | | | 0,8304 |
| Sarilumab (N = 52) | 45 | 6,60 (1,78) | 45 | 2,95 (2,03) | -3,64 (2,50) | -1,321 [-2,228; -0,415] 0,0048 **-0,632 [-1,066; -0,198]** | |
| Adalimumab (N = 47) | 39 | 6,97 (2,02) | 39 | 4,40 (2,00) | -2,57 (2,56) | | |

ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; RAID: *Rheumatoid Arthritis Impact of Disease*; SD: Standardabweichung

[a] LSMD, 95% KI und p-Wert basieren auf dem a priori definierten MMRM (*Mixed Model for Repeated Measures*)-Modell. In das Modell wurden als feste Effekte Behandlung und Region, der Messwiederholungsfaktor Visite und die Interaktion Behandlung*Visite aufgenommen. Der Baseline-Wert wurde dabei als stetige Kovariate in das Modell aufgenommen.

[b] Für die Auswertung der Teilpopulationen wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 70 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Nach 24-wöchiger Behandlung zeigte sich ein deutlicher Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verbesserung der Lebensqualität gemessen mittels des krankheitsspezifischen Fragebogens RAID.

In der relevanten Studienpopulation erreichten die Patienten unter Sarilumab nach 24 Wochen Behandlung eine statistisch signifikant stärkere Verbesserung ihrer Lebensqualität als die Patienten unter Adalimumab (LSMD [95 % KI] p-Wert: -1,346 [-2,022;-0,671] 0,0001). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges´g zeigt, dass es sich dabei um einen klinisch relevanten Effekt von kleiner Effektstärke handelt (Hedges´g [95 % KI]: -0,651 [-0,978; -0,325]).

In den Teilpopulationen 2 und 3 zeigte sich wie in der relevanten Studienpopulation ein statistisch signifikanter Vorteil für Sarilumab (LSMD [95 % KI] p-Wert: -1,306 [-2,397; -0,216] 0,0197 in Teilpopulation 2 und -1,321 [-2,228; -0,415] 0,0048 in Teilpopulation 3). Anhand der standardisierten Mittelwertdifferenz Hedges´g kann jedoch keine Aussage zur klinischen Relevanz dieser Effekte abgeleitet werden, da das Konfidenzintervall in beiden Fällen den Wert -0,2 einschließt.

Der Interaktionstest ergab keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den Endpunkt „Änderung des RAID (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Des Weiteren wurde eine Responderanalyse hinsichtlich einer Verbesserung des RAID um mindestens 3 Punkte durchgeführt. Dieser Schwellenwert wurde als minimale klinisch relevante Verbesserung für Patienten mit rheumatoider Arthritis definiert [70].

In der relevanten Studienpopulation erzielten 49,4 % der Patienten unter Sarilumab eine Verbesserung des RAID um mindestens 3 Punkte, während es unter Adalimumab 30,5 % der Patienten waren. Dieser Gruppenunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 1,62 [1,10; 2,37] 0,0142).

In der Teilpopulation 2 erreichten 51,4 % der mit Sarilumab behandelten Patienten eine Verbesserung des RAID um mindestens 3 Punkte, während es unter Adalimumab 20,0 % der Patienten waren. Dieser Gruppenunterschied ist ebenfalls statistisch signifikant (RR [95 % KI] p-Wert: 2,35 [1,15; 4,83] 0,0201).

Bei den Patienten der Teilpopulation 3 zeigte sich ebenfalls ein numerischer Vorteil für Sarilumab hinsichtlich der Verbesserung des RAID um mindestens 3 Punkte. Der Behandlungsunterschied im Vergleich zu Adalimumab ist jedoch statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,26 [0,80; 2,00] 0,3186).

Der Interaktionstest bezüglich einer Effektmodifikation durch die Anzahl der csDMARD-Vortherapien ergab für das RR keinen Beleg für eine Effektmodifikation. Somit kann für den Endpunkt „RAID-Responder (Woche 24)“ die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden.

Tabelle 4-136: Ergebnisse für „RAID-Responder (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 3 in Woche 24** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (ITT) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 43 (49,4) | 2,35 [1,23; 4,52] **0,0103** | 1,62 [1,10; 2,37] **0,0142** | 0,20 [0,05; 0,34] **0,0077** | |
| Adalimumab (N = 82) | 25 (30,5) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,0651 RR: 0,1520 RD: 0,1349 |
| Sarilumab (N = 35) | 18 (51,4) | 5,71 [1,68; 19,43] **0,0055** | 2,35 [1,15; 4,83] **0,0201** | 0,30 [0,10; 0,50] **0,0034** | |
| Adalimumab (N = 35) | 7 (20,0) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 25 (48,1) | 1,43 [0,63; 3,25] 0,3947 | 1,26 [0,80; 2,00] 0,3186 | 0,08 [-0,12; 0,29] 0,4154 | |
| Adalimumab (N = 47) | 18 (38,3) | | | | |

ITT: Intent-to treat; KI: Konfidenzintervall; OR: Odds Ratio; RAID: *Rheumatoid Arthritis Impact of Disease*; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe, Region sowie die Interaktionen Subgruppe*Behandlung und Region*Behandlung aufgenommen.

Patienten mit fehlenden Werten wurden als Non-Responder gezählt.

Quelle: [72]

Die Auswertung nach 48 Wochen Behandlung mit Sarilumab bezüglich einer Verbesserung der Lebensqualität gemäß RAID um mindestens 3 Punkte zeigte lediglich einen leichten Rückgang der Responderzahlen im Vergleich zur Auswertung nach Woche 24 (siehe Tabelle 4-137). Etwa 90 % der Responder in Woche 24 konnten ihre Verbesserung der Lebensqualität gemäß RAID beibehalten (39 Responder in Woche 48 vs. 43 Responder in Woche 24).

Die Auswertung nach 48 Wochen belegt somit die langanhaltende gute Wirksamkeit der Behandlung mit Sarilumab auch in Bezug auf die gesundheitsbezogene Lebensqualität.

Tabelle 4-137: Ergebnisse für „RAID-Responder (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Verbesserung ≥ 3 in Woche 24** | | **Patienten mit Verbesserung ≥ 3 in Woche 48** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 43 (49,4) | 75 | 39 (52,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 18 (51,4) | 30 | 15 (50,0) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 25 (48,1) | 45 | 24 (53,3) |

RAID: *Rheumatoid Arthritis Impact of Disease*
Patienten mit fehlenden Werten wurden in der Auswertung als Non-Responder berücksichtigt.
Quelle: [90]

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Krankheitsspezifische Lebensqualität – RAID“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

### 4.3.1.3.10 Sicherheit/Verträglichkeit – RCT

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

Tabelle 4-138: Operationalisierung von „Sicherheit/Verträglichkeit“

| Studie | Operationalisierung |
|---|---|
| MONARCH | Unerwünschte Ereignisse (UE) waren alle Erkrankungen, Symptome oder medizinisch relevante Parameter, die im zeitlichen Zusammenhang mit der Studie neu auftraten oder sich verschlimmerten, unabhängig davon, ob ein ursächlicher Zusammenhang zur Studienmedikation bestand oder nicht. Alle UEs wurden vom Zeitpunkt der Unterzeichnung der Einverständniserklärung bis zum Studienende dokumentiert.<br>Hier dargestellt sind die UE, die im Behandlungszeitraum auftraten, d.h. von der ersten bis zur letzten Dosis der Studienmedikation + 60 Tage (*treatment-emergent adverse events*).<br>Alle Ereignisse wurden nach MedDRA (Version 18.1) codiert und dokumentiert. Neben der Gesamtrate an UE wurden schwerwiegende unerwünschte Ereignisse (SUE), UE mit Todesfolge, UE, die zum dauerhaften Abbruch der Behandlung führten sowie UE von besonderem Interesse (UEBI) aufgeführt.<br>Als UEBI waren folgende Ereignisse präspezifiziert:<br>• *Leukopenien* wurden über den MedDRA SMQ „Hämatopoetische Leukopenien“ erfasst. Neutropenien vom Grad 3 (absolute Neutrophilenzahl < 1.000/µl) und Grad 4 (< 500/µl) wurden als UE gewertet. Patienten mit einer Grad 3-Neutropenie und Anzeichen einer Infektion sowie Patienten mit einer Grad 4-Neutropenie mussten gemäß Protokoll die Behandlung dauerhaft abbrechen. Als schwerwiegende Neutropenie galt das Vorliegen einer Neutropenie vom Grad 4 über mindestens 6 Tage.<br>• *Thrombozytopenien* wurden über den MedDRA SMQ „Hämatopoetische Thrombozytopenien“ erfasst. Thrombozytopenien < 100.000/µl wurden als UE gewertet.<br>• *Infektionen* wurden über die MedDRA SOC „Infektionen und parasitäre Erkrankungen“ erfasst.<br>• *Leberwerterhöhungen* wurden über den MedDRA SMQ „Durch Arzneimittel bedingte Erkrankungen der Leber – Umfassende Suche“ erfasst.<br>• *Divertikulitis und gastrointestinale Ulzerationen* wurden über die MedDRA SMQs „Gastrointestinale Perforation“ und „Gastrointestinaler Ulkus“ erfasst.<br>• Eine *Erhöhung der Lipidwerte* wurde über den MedDRA SMQ „Dyslipidaemie“ erfasst.<br>• *Überempfindlichkeit* wurde über den MedDRA SMQ „Hypersensibilität“ erfasst. *Anaphylaxie* wurde über den MedDRA SMQ „Anaphylaktische Reaktion“ erfasst.<br>• *Reaktionen der Injektionsstelle* wurden über den MedDRA HLT „Reaktionen der Injektionsstelle“ erfasst.<br>• *Krebserkrankungen* wurden über den MedDRA SMQ „Maligne oder unspezifizierte Tumoren“ erfasst.<br>• *Lupus-ähnliches Syndrom* wurde über den MedDRA SMQ „Systemischer Lupus erythematodes“ erfasst.<br>• *Demyelinisierende Erkrankungen* wurden über den MedDRA SMQ „Demyelinisierung“ erfasst.<br>Alle mutmaßlichen kardiovaskulären unerwünschten Ereignisse sowie alle Todesfälle wurden an den Ausschuss zur Beurteilung kardiovaskulärer Ereignisse gemeldet, welcher basierend auf verblindeten Daten eine Beurteilung dieser Fälle vornahm. |

*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 4-139: Bewertung des Verzerrungspotenzials für „Sicherheit/Verträglichkeit“ 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 |
|---|---|---|---|---|---|---|
| MONARCH | niedrig | ja | ja | ja | ja | niedrig |

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential der Studie MONARCH ist niedrig. Für die Analyse des Endpunktes „Sicherheit/Verträglichkeit“ ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine randomisierte, doppelblinde Studie, das ITT-Prinzip wurde adäquat umgesetzt und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen.

*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.*

### *Unerwünschte Ereignisse*

Tabelle 4-140: Ergebnisse für „Gesamtrate unerwünschter Ereignisse (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 58 (66,7) | 0,93 [0,49;1,77] 0,8217 | 0,98 [0,79; 1,20] 0,8216 | -0,02 [-0,16; 0,13] 0,8216 | |
| Adalimumab (N = 82) | 56 (68,3) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,2828 RR: 0,2721 RD: 0,2713 |
| Sarilumab (N = 35) | 19 (54,3) | 0,62 [0,23; 1,64] 0,3318 | 0,83 [0,56; 1,22] 0,3345 | -0,11 [-0,34; 0,12] 0,3272 | |
| Adalimumab (N = 35) | 23 (65,7) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 39 (75,0) | 1,27 [0,52; 3,11] 0,5944 | 1,07 [0,84; 1,37] 0,5962 | 0,05 [-0,13; 0,23] 0,5945 | |
| Adalimumab (N = 47) | 33 (70,2) | | | | |

KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 42 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Die Gesamtrate der unerwünschten Ereignisse war nach 24 Wochen Behandlung in beiden Behandlungsgruppen vergleichbar (siehe Tabelle 4-140). In der relevanten Studienpopulation wurden bei 66,7 % der mit Sarilumab behandelten Patienten sowie bei 68,3 % der mit Adalimumab behandelten Patienten unerwünschte Ereignisse dokumentiert. Somit ergab sich bei einem RR von 0,98 kein statistisch signifikanter Gruppenunterschied ([95 % KI] p-Wert: [0,79; 1,20] 0,8216).

Die Gesamtraten unerwünschter Ereignisse bei den Patienten der Teilpopulationen 2 und 3 waren mit denen der relevanten Studienpopulation vergleichbar. Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Gesamtrate unerwünschter Ereignisse (Woche 24)“

die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

In Tabelle 4-141 sind für die relevante Studienpopulation diejenigen unerwünschten Ereignisse dargestellt, die entweder häufig auftraten, das heißt bei mindestens 10 % der Patienten in einem Behandlungsarm, oder die in den Behandlungsarmen unterschiedlich häufig auftraten, das heißt mit einem Gruppenunterschied von mindestens 5 Prozentpunkten.

Tabelle 4-141: Deskriptive Darstellung ausgewählter unerwünschter Ereignisse in der relevanten Studienpopulation nach MedDRA SOC und PT (Woche 24)

| **Studie MONARCH** | | |
|---|---|---|
| **Behandlung** | **Adalimumab (N = 82)** | **Sarilumab (N = 87)** |
| Systemorganklasse[a]<br>- *Preferred Term*[a] | n (%) | n (%) |
| Infektionen und parasitäre Erkrankungen | 23 (28,0) | 25 (28,7) |
| Erkrankungen des Blutes und des Lymphsystems | 2 (2,4) | 15 (17,2) |
| - Neutropenien | 1 (1,2) | 14 (16,1) |
| Erkrankungen des Nervensystems | 9 (11,0) | 3 (3,4) |
| - Kopfschmerzen | 5 (6,1) | 1 (1,1) |
| Erkrankungen des Gastrointestinaltrakts | 11 (13,4) | 9 (10,3) |
| Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen | 11 (13,4) | 7 (8,0) |
| Allgemeine Erkrankungen und Beschwerden am Verabreichungsort | 9 (11,0) | 9 (10,3) |
| Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen | 10 (12,2) | 7 (8,0) |

MedDRA: *Medical Dictionary for Regulatory Activities*

[a] Auswahl von UE mit einer Häufigkeit von ≥ 10 % in einer Behandlungsgruppe oder mit einem Unterschied von ≥ 5 %-Punkten zwischen den Studienarmen

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 43 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Am häufigsten wurden in beiden Behandlungsarmen Infektionen verzeichnet. In der Mehrzahl der Fälle handelte es sich dabei um Atemwegsinfekte. Die Häufigkeit von Infektionen war in beiden Behandlungsarmen vergleichbar. Infektionen waren als unerwünschte Ereignisse von besonderem Interesse präspezifiziert und sind in Tabelle 4-149 und Tabelle 4-150 detaillierter dargestellt.

Weitere häufige unerwünschte Ereignisse bei den mit Sarilumab behandelten Patienten waren Erkrankungen des Blutes und des Lymphsystems. Dabei handelte es sich fast ausschließlich um Neutropenien. Neutropenien waren als unerwünschte Ereignisse von besonderem Interesse präspezifiziert und sind in Tabelle 4-153 und Tabelle 4-154 detaillierter dargestellt.

Des Weiteren traten bei den mit Adalimumab behandelten Patienten häufiger Erkrankungen des Nervensystems auf. Weitere häufige Ereignisse, für die sich keine relevanten Unterschiede zwischen den Behandlungsgruppen zeigten, waren Erkrankungen des Gastrointestinaltrakts, Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen, allgemeine Erkrankungen und Beschwerden am Verabreichungsort sowie Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen.

Für den Großteil der unerwüschten Ereignisse wurde der Schweregrad als leicht oder moderat eingestuft [73].

Tabelle 4-142: Ergebnisse für „Gesamtrate unerwünschter Ereignisse (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 58 (66,7) | 75 | 50 (66,7) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 19 (54,3) | 30 | 17 (56,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 39 (75,0) | 45 | 33 (73,3) |
| Quelle: [92] | | | | |

Die Gesamtrate der UE in der offenen Extensionsphase (Woche 25 bis 48) betrug in der relevanten Studienpopulation 66,7% und blieb damit unverändert im Vergleich zu den ersten 24 Wochen der Behandlung (siehe Tabelle 4-142).

In Tabelle 4-143 sind diejenigen unerwünschten Ereignisse nach MedDRA SOC und PT dargestellt, die unter Sarilumab innerhalb der offenen Extensionsphase bei mindestens 10 % der Patienten auftraten. Die häufigsten UE waren, wie schon in der 24-wöchigen randomisierten Behandlungsphase, Infektionen.

Die Auswertung nach 48 Wochen bestätigt somit das Sicherheitsprofil von Sarilumab, das sich in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat.

Tabelle 4-143: Deskriptive Darstellung ausgewählter unerwünschter Ereignisse in der relevanten Studienpopulation nach MedDRA SOC und PT (Woche 25 - 48)

| **Studie MONARCH** | | |
|---|---|---|
| | **Patienten mit Ereignis bis Woche 24 (N = 87)** | **Patienten mit Ereignis in Woche 25 bis 48 (N = 75)** |
| Systemorganklasse[a]<br>- *Preferred Term*[a] | n (%) | n (%) |
| Infektionen und parasitäre Erkrankungen | 25 (28,7) | 22 (29,3) |
| Erkrankungen des Blutes und des Lymphsystems | 15 (17,2) | 9 (12,0) |
| - Neutropenien | 14 (16,1) | 8 (10,7) |
| Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen | 7 (8,0) | 9 (12,0) |
| Allgemeine Erkrankungen und Beschwerden am Verabreichungsort | 9 (10,3) | 10 (13,3) |
| Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen | 7 (8,0) | 11 (14,7) |

MedDRA: *Medical Dictionary for Regulatory Activities*

[a] Auswahl von UE mit einer Häufigkeit von ≥ 10 % in der Behandlungsgruppe mit 200 mg Sarilumab im Zeitraum von Woche 25 bis 48

Quellen: [73, 92]

### *Schwerwiegende unerwünschte Ereignisse*

Tabelle 4-144: Ergebnisse für „Gesamtrate schwerwiegender unerwünschter Ereignisse (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[b] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 4 (4,6) | 0,74 [0,19; 2,88] 0,6655 | 0,75 [0,21; 2,72] 0,6657 | -0,02 [-0,08; 0,05] 0,6656 | |
| Adalimumab (N = 82) | 5 (6,1) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: 0,1339 RR: n.a. RD: 0,4826 |
| Sarilumab (N = 35) | 4 (11,4) | OR: 1,38 [0,28; 6,74] 0,6918 Peto-OR: 1,37 [0,29; 6,45] 0,6924 | 1,33 [0,32; 5,58] 0,6922 | 0,03 [-0,11; 0,17] 0,6905 | |
| Adalimumab (N = 35) | 3 (8,6) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 0 | Peto-OR: 0,12 [0,01; 1,94] 0,1349 | n.a. | -0,04 [-0,18; 0,10] 0,5533 | |
| Adalimumab (N = 47) | 2 (4,3) | | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

[b] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet und in der korrespondierenden Subgruppe wurde Peto-OR zusätzlich zum OR angegeben. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 42 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Schwerwiegende unerwünschte Ereignisse traten innerhalb der 24-wöchigen Behandlungsphase lediglich bei 4 der mit Sarilumab behandelten bzw. 5 der mit Adalimumab behandelten Patienten der relevanten Studienpopulation auf (siehe Tabelle 4-144). Es ergab sich bei einem RR von 0,75 kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen ([95 % KI] p-Wert: [0,21; 2,72] 0,6657).

Bei den schwerwiegenden unerwünschten Ereignissen der Patienten unter Sarilumab handelte es sich um zwei Infektionen (eine Bursitis und eine Mastitis), eine Neutropenie sowie einen Fall einer Herzerkrankung mit Todesfolge [72].

Bei den schwerwiegenden unerwünschten Ereignissen der Patienten unter Adalimumab handelte es sich um einen Schlaganfall, eine Lungenembolie, eine Darmobstruktion, eine Verschlechterung der Arthritis und eine Leberwerterhöhung [72].

Die Interaktionstests ergaben keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Gesamtrate schwerwiegender unerwünschter Ereignisse (Woche 24)“ die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

Tabelle 4-145: Ergebnisse für „Gesamtrate schwerwiegender unerwünschter Ereignisse (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 4 (4,6) | 75 | 3 (4,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 4 (11,4) | 30 | 0 |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 0 | 45 | 3 (6,7) |
| Quelle: [92] | | | | |

Die Gesamtrate der SUE in der offenen Extensionsphase (Woche 25 bis 48) betrug in der relevanten Studienpopulation unter Sarilumab lediglich 4,0 % und blieb damit ebenfalls konstant im Vergleich zu den ersten 24 Wochen der Behandlung (siehe Tabelle 4-145). Bei den Patienten mit einem schwerwiegenden unerwünschten Ereignis in der Extensionsphase handelte es sich um einen Patienten mit Anämie, einen Fall von Thoraxschmerzen nicht kardialen Ursprungs sowie einen Patienten mit Schilddrüsenvergrößerung.

Die Auswertung nach 48 Wochen bestätigt das Sicherheitsprofil von Sarilumab, das sich in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat. Es kam nicht zu einer Zunahme schwerwiegender unerwünschter Ereignisse.

### *Therapieabbrüche aufgrund unerwünschter Ereignisse*

Tabelle 4-146: Ergebnisse für „Therapieabbrüche aufgrund unerwünschter Ereignisse (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 7 (8,0) | 2,30 [0,57; 9,27] 0,2392 | 2,20 [0,59; 8,26] 0,2421 | 0,04 [-0,03; 0,11] 0,2209 | |
| Adalimumab (N = 82) | 3 (3,7) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,9215 RR: 0,9330 RD: 0,7199 |
| Sarilumab (N = 35) | 2 (5,7) | 2,06 [0,18; 24,3] 0,5634 | 2,00 [0,19; 21,4] 0,5647 | 0,03 [-0,07; 0,12] 0,5549 | |
| Adalimumab (N = 35) | 1 (2,9) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 5 (9,6) | 2,39 [0,44; 13,1] 0,3129 | 2,26 [0,46; 11,2] 0,3169 | 0,05 [-0,05; 0,15] 0,2889 | |
| Adalimumab (N = 47) | 2 (4,3) | | | | |

KI: Konfidenzintervall; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 42 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Therapieabbrüche aufgrund von unerwünschten Ereignissen traten nach 24 Wochen Behandlung in der relevanten Studienpopulation bei 8,0 % der mit Sarilumab behandelten und 3,7 % der mit Adalimumab behandelten Patienten auf (siehe Tabelle 4-146). Der Gruppenunterschied war bei einem RR von 2,2 statistisch nicht signifikant ([95 % KI] p-Wert: [0,59; 8,26] 0,2421). Die Raten der Therapieabbrüche aufgrund unerwünschter Ereignisse bei den Patienten der Teilpopulationen 2 und 3 waren mit denen der relevanten Studienpopulation vergleichbar.

Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass zur Bewertung des Zusatznutzens für den

Endpunkt „Therapieabbrüche aufgrund unerwünschter Ereignisse (Woche 24)“ die relevante Studienpopulation herangezogen werden kann.

Bei den mit Sarilumab behandelten Patienten stellten Neutropenien die häufigste Ursache für den Abbruch der Behandlung dar. Gemäß Protokoll mussten Patienten mit einer Grad 3-Neutropenie und Anzeichen einer Infektion sowie Patienten mit einer Grad 4-Neutropenie die Behandlung dauerhaft abbrechen. Neutropenien waren auch als unerwünschte Ereignisse von besonderem Interesse präspezifiziert und sind in Tabelle 4-153 und Tabelle 4-154 detaillierter dargestellt.

Tabelle 4-147: Ergebnisse für „Therapieabbrüche aufgrund unerwünschter Ereignisse (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 7 (8,0) | 75 | 7 (9,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 2 (5,7) | 30 | 0 |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 5 (9,6) | 45 | 7 (15,6) |
| Quelle: [92] | | | | |

Die Auswertung der Therapieabbrüche aufgrund unerwünschter Ereignisse in der offenen Extensionsphase (Woche 25 bis 48) ergab eine gleichbleibende Rate von Therapieabbrüchen im Vergleich zur doppelblinden Behandlungsphase (bis Woche 24). 9,3 % der Patienten der relevanten Studienpopulation beendeten die Therapie nach Woche 24 aufgrund von unerwünschten Ereignissen (siehe Tabelle 4-147). Die Auswertung nach 48 Wochen bestätigt somit die Verträglichkeit von Sarilumab, die sich in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat.

### *Unerwünschte Ereignisse mit Todesfolge*

Tabelle 4-148: Ergebnisse für „Unerwünschte Ereignisse mit Todesfolge (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer** | | | **Interaktionstest (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[a] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 1 (1,1) | Peto-OR: 6,98 [0,14; 352,19] 0,3316 | n.a. | 0,01 [-0,03; 0,06][b] 0,5255[b] | |
| Adalimumab (N = 82) | 0 | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: n.a. RR: n.a. RD: n.a. |
| Sarilumab (N = 35) | 1 (2,9) | Peto-OR: 7,39 [0,15; 372,38] 0,3173 | n.a. | 0,03 [-0,073; 0,150][b] 0,5276[b] | |
| Adalimumab (N = 35) | 0 | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 0 | n.a. | n.a. | 0,00 [-0,076; 0,071] [b] 1,0000 | |
| Adalimumab (N = 47) | 0 | | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

[b] Konfidenzintervall und p-Wert wurden kombinatorisch exakt berechnet.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 42 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

In der 24-wöchigen randomisierten Behandlungsphase trat ein Todesfall in der Gruppe der mit Sarilumab behandelten Patienten auf. Es handelte sich dabei um eine Frau, die infolge einer akuten Herzinsuffizienz, Papillarmuskelruptur und Aortendissektion verstarb. Es gab keine Todesfälle im Adalimumab-Arm.

In der offenen Extensionsphase (Woche 25 bis 48) trat in der relevanten Studienpopulation kein weiterer Todesfall auf.

### *Unerwünschte Ereignisse von besonderem Interesse*

*Infektionen*

Tabelle 4-149: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Infektionen (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 25 (28,7) | 1,03 [0,53; 2,02] 0,9212 | 1,02 [0,63; 1,66] 0,9212 | 0,01 [-0,13; 0,14] 0,9212 | |
| Adalimumab (N = 82) | 23 (28,0) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,9328 RR: 0,9326 RD: 0,9330 |
| Sarilumab (N = 35) | 10 (28,6) | 1,00 [0,35; 2,84] 1,0000 | 1,00 [0,47; 2,11] 1,0000 | 0,00 [-0,21; 0,21] 1,0000 | |
| Adalimumab (N = 35) | 10 (28,6) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 15 (28,8) | 1,06 [0,44; 2,56] 0,8960 | 1,04 [0,55; 1,97] 0,8961 | 0,01 [-0,17; 0,19] 0,8959 | |
| Adalimumab (N = 47) | 13 (27,7) | | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 47 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Die Therapie mit TNF- und IL-6R-Antagonisten ist mit einem erhöhten Infektionsrisiko verbunden. Aus diesem Grund wurden Infektionen in der Studie MONARCH als unerwünschte Ereignisse von besonderem Interesse präspezifiziert und dokumentiert.

Infektionen traten in der 24-wöchigen doppelblinden Behandlungsphase bei 28,7 % der mit Sarilumab behandelten bzw. 28,0 % der mit Adalimumab behandelten Patienten der relevanten Studienpopulation auf (siehe Tabelle 4-149). Das Risiko für das Auftreten einer

Infektion war somit in den Behandlungsgruppen vergleichbar (RR [95 % KI] p-Wert: 1,02 [0,63; 1,66] 0,9212). In den meisten Fällen handelte es sich um Atemwegsinfekte mit mildem bis moderatem Verlauf [72].

Die Raten der Infektionen in den Teilpopulationen 2 und 3 waren mit denen der relevanten Studienpopulation vergleichbar. Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse - Infektionen (Woche 24)“ die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

Lediglich bei zwei Patienten im Sarilumab-Arm wurden die Infektionen als schwerwiegende Infektionen eingestuft (siehe nachfolgende Tabelle 4-150). Bei den Infektionen handelte sich um eine Bursitis und eine Mastitis. Beide schwerwiegenden Infektionen waren nicht mit einer Neutropenie assoziiert.

Tabelle 4-150: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Infektionen (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[b] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 2 (2,3) | Peto-OR: 7,06 [0,44; 113,92] 0,1685 | n.a. | 0,02 [-0,02; 0,07] 0,2774 | |
| Adalimumab (N = 82) | 0 | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: n.a. RR: n.a. RD: n.a. |
| Sarilumab (N = 35) | 2 (5,7) | Peto-OR: 7,61 [0,47; 124,15] 0,1543 | n.a. | 0,06 [-0,05; 0,19][c] 0,2070[c] | |
| Adalimumab (N = 35) | 0 | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 0 | n.a. | n.a. | 0,00 [-0,076; 0,071] [b] 1,0000 | |
| Adalimumab (N = 47) | 0 | | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

[b] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

[c] Konfidenzintervall und p-Wert wurden kombinatorisch exakt berechnet.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 47 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

In der offenen Extensionsphase (Woche 25 bis 48) traten bei 29,3 % der Patienten der relevanten Studienpopulation Infektionen auf (siehe Tabelle 4-151). Die Rate der Infektionen war damit vergleichbar mit der Rate der Infektionen innerhalb der 24-wöchigen doppelblinden Behandlungsphase. Keine der aufgetretenen Infektionen nach Woche 24 wurde als schwerwiegend eingestuft (siehe Tabelle 4-152). Die Auswertung nach 48 Wochen bestätigt somit das Sicherheitsprofil von Sarilumab, das sich in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat. Es kam nicht zu einer Zunahme der Zahl von

Infektionen oder schwerwiegenden Infektionen unter Sarilumab. Dies wird auch durch die EMA für die Gesamtschau der Zulassungsstudien von Sarilumab bestätigt, die belegt, dass die Rate der Infektionen und schwerwiegenden Infektionen im Zeitverlauf stabil bleibt [89].

Tabelle 4-151: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Infektionen (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 25 (28,7) | 75 | 22 (29,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 10 (28,6) | 30 | 6 (20,0) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 15 (28,8) | 45 | 16 (35,6) |
| Quelle: [92] | | | | |

Tabelle 4-152: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Infektionen (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | N | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 2 (2,3) | 75 | 0 |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 2 (5,7) | 30 | 0 |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 0 | 45 | 0 |
| Quelle: [92] | | | | |

*Leukopenien*

Tabelle 4-153: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Leukopenien (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[b] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 14 (16,1) | 7,67 [1,68; 35,1] **0,0088** | 6,60 [1,54; 28,3] **0,0112** | 0,14 [0,05; 0,22] **0,0016** | |
| Adalimumab (N = 82) | 2 (2,4) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: 0,5576 RR: n.a. RD: 0,6862 |
| Sarilumab (N = 35) | 4 (11,4) | Peto-OR: 8,09 [1,09; 60,03] **0,0409** | n.a. | 0,11 [-0,01; 0,24] 0,0674 | |
| Adalimumab (N = 35) | 0 | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 10 (19,2) | OR: 5,36 [1,10; 26,2] **0,0383** | 4,52 [1,03; 19,8] **0,0454** | 0,15 [0,03; 0,27] **0,0169** | |
| Adalimumab (N = 47) | 2 (4,3) | Peto-OR: 4,02 [1,21; 13,39] **0,0233** | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

[b] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet und in der korrespondierenden Subgruppe wurde Peto-OR zusätzlich zum OR angegeben. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 46 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Die Verringerung der Zahl der Neutrophilen im Blut (Neutropenie) gehört zu den bekannten Risiken bei der Anwendung von IL-6R-Antagonisten [1, 93]. Aus diesem Grund wurde das Auftreten von Neutropenien in der Studie MONARCH als unerwünschtes Ereignis von besonderem Interesse dokumentiert und ausgewertet. In der Studie wurden sämtliche

Leukopenien erfasst, in den allermeisten Fällen handelte es sich jedoch um Neutropenien (siehe Tabelle 4-141).

In der relevanten Studienpopulation trat innerhalb von 24 Wochen Behandlung bei 16,1 % der Patienten unter Sarilumab eine Leukopenie auf, unter Adalimumab waren im Gegensatz nur 2,4 % der Patienten von einer Leukopenie betroffen (siehe Tabelle 4-153). Der Gruppenunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 6,60 [1,54; 28,3] 0,0112).

Bei den Patienten in den Teilpopulationen 2 und 3 zeigte sich ebenfalls ein statistisch signifikant erhöhtes Risiko für das Auftreten einer Leukopenie (Peto-OR [95 % KI] p-Wert: 8,09 [1,09; 60,03] 0,0409 in Teilpopulation 2 und RR [95 % KI] p-Wert: 4,52 [1,03; 19,8] 0,0454 sowie Peto-OR [95 % KI] p-Wert: 4,02 [1,21; 13,39] 0,0233 in Teilpopulation 3).

Die Interaktionstests ergaben keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, somit kann zur Bewertung des Zusatznutzens für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse – Leukopenien (Woche 24)“ die relevante Studienpopulation herangezogen werden.

Die in der Studie MONARCH aufgetretenen Neutropenien waren von transientem Verlauf und normalisierten sich innerhalb weniger Tage[27]. Das Auftreten von Neutropenien war nicht mit dem erhöhten Auftreten von Infektionen assoziiert. Von den aufgetretenen Leukopenien wurde nur ein Fall im Sarilumab-Arm als schwerwiegend eingestuft (siehe Tabelle 4-154). Bei diesem Patienten lag eine Überdosierung von Sarilumab aufgrund der Nichteinhaltung des Behandlungsintervalls von zwei Wochen vor. Die Neutrophilenzahl normalisierte sich innerhalb von 10 Tagen nach der letzten Sarilumab-Dosis [71].

[27] Für einen Patienten in der Studie konnte die Normalisierung der Werte nicht bestätigt werden, da nach Absetzen der Therapie keine Laborwerte verfügbar waren (*lost to follow-up*).

Tabelle 4-154: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Leukopenien (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer** | | | **Interaktionstest (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[a] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 1 (1,1) | Peto-OR: 6,98 [0,14; 352,19] 0,3316 | n.a. | 0,01 [-0,03; 0,06][b] 0,5255[b] | |
| Adalimumab (N = 82) | 0 | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: n.a. RR: n.a. RD: n.a. |
| Sarilumab (N = 35) | 1 (2,9) | Peto-OR: 7,39 [0,15; 372,38] 0,3173 | n.a. | 0,03 [-0,07; 0,15][b] 0,5276[b] | |
| Adalimumab (N = 35) | 0 | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 0 | n.a. | n.a. | 0,00 [-0,076; 0,071] [b] 1,0000 | |
| Adalimumab (N = 47) | 0 | | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Peto-OR

[b] Konfidenzintervall und p-Wert wurden kombinatorisch exakt berechnet.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 46 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

In der offenen Extensionsphase (Woche 25 bis 48) trat bei 10,7 % der Patienten der relevanten Studienpopulation eine Leukopenie auf (siehe Tabelle 4-155). Die Rate der Leukopenien war damit etwas niedriger als die Rate der Leukopenien innerhalb der 24-wöchigen doppelblinden Behandlungsphase. Keine der aufgetretenen Leukopenien nach Woche 24 wurde als schwerwiegend eingestuft (siehe Tabelle 4-156). Die Auswertung nach 48 Wochen bestätigt somit das Sicherheitsprofil von Sarilumab, das sich in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat. Es kam nicht zu einer Zunahme der Zahl von Leukopenien oder schwerwiegenden Leukopenien unter Sarilumab. Dies wird auch durch die EMA für die Gesamtschau der Zulassungsstudien von Sarilumab bestätigt, die belegt, dass die Rate der Neutropenien und schwerwiegenden Neutropenien im Zeitverlauf nicht ansteigt [89].

Tabelle 4-155: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Leukopenien (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 14 (16,1) | 75 | 8 (10,7) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 4 (11,4) | 30 | 3 (10,0) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 10 (19,2) | 45 | 5 (11,1) |
| Quelle: [92] | | | | |

Tabelle 4-156: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Leukopenien (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 1 (1,1) | 75 | 0 |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 1 (2,9) | 30 | 0 |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 0 | 45 | 0 |
| Quelle: [92] | | | | |

*Leberwerterhöhungen*

Tabelle 4-157: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Leberwerterhöhungen (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 6 (6,9) | 1,95 [0,47; 8,11] 0,3571 | 1,89 [0,49; 7,32] 0,3589 | 0,03 [-0,03; 0,10] 0,3440 | |
| Adalimumab (N = 82) | 3 (3,7) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: 0,5784 RR: 0,5816 RD: 0,5511 |
| Sarilumab (N = 35) | 3 (8,6) | 3,19 [0,31; 32,8] 0,3276 | 3,00 [0,32; 27,9] 0,3322 | 0,06 [-0,05; 0,17] 0,3009 | |
| Adalimumab (N = 35) | 1 (2,9) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 3 (5,8) | 1,38 [0,22; 8,74] 0,7326 | 1,36 [0,23; 7,87] 0,7329 | 0,02 [-0,07; 0,10] 0,7296 | |
| Adalimumab (N = 47) | 2 (4,3) | | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 46 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Eine Erhöhung der Lebertransaminasen ist ein Risiko, welches unter Behandlung mit IL-6R- und TNF-Antagonisten besteht. Aus diesem Grund wurde das Auftreten von Leberwerterhöhungen in der Studie MONARCH als unerwünschtes Ereignis von besonderem Interesse dokumentiert und ausgewertet.

In der relevanten Studienpopulation trat in der 24-wöchigen Behandlungsphase bei 6,9 % der Patienten im Sarilumab-Arm und bei 3,7 % der Patienten im Adalimumab-Arm eine Erhöhung der Lebertransaminasen auf. Der Gruppenunterschied ist bei einem RR von 1,89 statistisch nicht signifikant ([95 % KI] p-Wert: [0,49; 7,32] 0,3589).

Die Ereignisraten in den Teilpopulationen 2 und 3 waren mit denen der relevanten Studienpopulation vergleichbar. Die Interaktionstests ergaben für keinen der Effektschätzer eine Effektmodifikation durch die Anzahl der csDMARD-Therapien, so dass für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse - Leberwerterhöhungen (Woche 24)“ die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

Lediglich ein Fall einer Erhöhung der Lebertransaminasen im Adalimumab-Arm in der Teilpopulation 2 wurde als schwerwiegend eingestuft. Es handelte sich dabei um Laborwertänderungen ohne klinische Zeichen einer Hepatitis oder Leberfunktionsstörung. Bei den mit Sarilumab behandelten Patienten traten keine schwerwiegenden Leberwerterhöhungen auf.

Tabelle 4-158: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Leberwerterhöhungen (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer** | | | **Interaktionstest (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[a]<br>[95% KI]<br>p-Wert | RR<br>[95% KI]<br>p-Wert | RD<br>[95% KI]<br>p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 0 | Peto-OR: 0,13<br>[0,0; 6,43]<br>0,3030 | n.a. | n.a. | |
| Adalimumab (N = 82) | 1 (1,1) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a.<br>Peto-OR: n.a.<br>RR: n.a.<br>RD: n.a. |
| Sarilumab (N = 35) | 0 | n.a. | n.a. | n.a. | |
| Adalimumab (N = 35) | 0 | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 0 | Peto-OR: 0,12<br>[0,0; 6,16]<br>0,2929 | n.a. | n.a. | |
| Adalimumab (N = 47) | 1 (2,1) | | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Peto-OR

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 46 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

In der offenen Extensionsphase (Woche 25 bis 48) trat bei 9,3 % der Patienten der relevanten Studienpopulation eine Leberwerterhöhung auf (siehe Tabelle 4-159). Die Rate der Leberwerterhöhungen war damit nach Woche 24 etwa gleichbleibend. Die Auswertung nach 48 Wochen bestätigt somit das Sicherheitsprofil von Sarilumab, das sich in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat. Es kam nicht zu einer Zunahme der Zahl von Leberwerterhöhungen unter Sarilumab.

Tabelle 4-159: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Leberwerterhöhungen (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 6 (6,9) | 75 | 7 (9,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 3 (8,6) | 30 | 2 (6,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 3 (5,8) | 45 | 5 (11,1) |
| Quelle: [92] | | | | |

Tabelle 4-160: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – schwerwiegende Leberwerterhöhungen (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 0 | 75 | 0 |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 0 | 30 | 0 |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 0 | 45 | 0 |
| Quelle: [92] | | | | |

*Lipidwerterhöhungen*

Tabelle 4-161: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Lipidwerterhöhung (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[b] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 0 | Peto-OR: 0,12 [0,02; 0,89] **0,0376** | n.a. | -0,05 [-0,10; 0,002] 0,0619 | |
| Adalimumab (N = 82) | 4 (4,9) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: 0,9483 RR: n.a. RD: n.a. |
| Sarilumab (N = 35) | 0 | Peto-OR: 0,14 [0,003; 6,82] 0,3173 | n.a. | -0,03 [-0,15; 0,07][c] 0,5276[c] | |
| Adalimumab (N = 35) | 1 (2,9) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 0 | Peto-OR: 0,12 [0,01; 1,15] 0,0657 | n.a. | -0,06 [-0,18; 0,01][c] 0,0765[c] | |
| Adalimumab (N = 47) | 3 (6,4) | | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

[b] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

[c] Konfidenzintervall und p-Wert wurden kombinatorisch exakt berechnet.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 46 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Ein weiteres Risiko unter Behandlung mit IL-6R- und TNF-Antagonisten ist eine Erhöhung der Lipidwerte. Aus diesem Grund wurde das Auftreten von Lipidwerterhöhungen in der Studie MONARCH als unerwünschtes Ereignis von besonderem Interesse dokumentiert und ausgewertet.

In der relevanten Studienpopulation wurde innerhalb von 24 Wochen Behandlung lediglich im Adalimumab-Arm bei 4,9 % der Patienten eine Lipidwerterhöhung festgestellt (siehe Tabelle 4-161). Der Gruppenunterschied zwischen Sarilumab und Adalimumab ist statistisch signifikant (Peto-OR [95 % KI] p-Wert: 0,12 [0,02; 0,89] 0,0376).

In den Teilpopulationen 2 und 3 zeigte sich aufgrund der geringen Ereignisraten kein signifikanter Unterschied zwischen den Behandlungsarmen. Der Interaktionstest auf Basis des Peto-OR ergab keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien, so dass für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse - Lipidwerterhöhung (Woche 24)“ die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden kann.

Schwerwiegende Lipidwerterhöhungen traten in der Studie MONARCH nicht auf.

In der offenen Extensionsphase (Woche 25 bis 48) trat bei einem Patienten der relevanten Studienpopulation eine Lipidwerterhöhung auf (siehe Tabelle 4-162). Die Rate der Lipidwerterhöhungen zeigte damit nach Woche 24 keine relevante Veränderung. Die Auswertung nach 48 Wochen bestätigt somit das Sicherheitsprofil von Sarilumab, das sich in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat. Es kam nicht zu einer relevanten Zunahme der Zahl von Lipidwerterhöhungen unter Sarilumab.

Tabelle 4-162: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse - Lipidwerterhöhung (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 0 | 75 | 1 (1,3) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 0 | 30 | 1 (3,3) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 0 | 45 | 0 |
| Quelle: [92] | | | | |

*Überempfindlichkeit*

Tabelle 4-163: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – Überempfindlichkeit (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[b] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 1 (1,1) | 0,23 [0,02; 2,09] 0,1895 | 0,24 [0,03; 2,08] 0,1926 | -0,04 [-0,09; 0,02] 0,1586 | |
| Adalimumab (N = 82) | 4 (4,9) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: 0,4996 RR: n.a. RD: 0,4952 |
| Sarilumab (N = 35) | 0 | Peto-OR: 0,13 [0,01; 2,14] 0,1543 | n.a. | -0,06 [-0,13; 0,01] 0,1051 | |
| Adalimumab (N = 35) | 2 (5,7) | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 1 (1,9) | OR: 0,44 [0,04; 5,12] 0,5108 | 0,45 [0,04; 4,91] 0,5118 | -0,02 [-0,09; 0,05] 0,5069 | |
| Adalimumab (N = 47) | 2 (4,3) | Peto-OR: 0,46 [0,05; 4,5] 0,5012 | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

[b] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 46 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Ein generelles Risiko bei der Applikation von Arzneimitteln ist das Auftreten von Überempfindlichkeitsreaktionen. Aus diesem Grund wurde in der Studie MONARCH Überempfindlichkeit als unerwünschtes Ereignis von besonderem Interesse erfasst.

Überempfindlichkeit trat bei 1,1 % der Sarilumab-Patienten und 4,9 % der Adalimumab-Patienten auf (siehe Tabelle 4-163). Alle Ereignisse waren von mildem bis moderatem Schweregrad. In der Studie traten keine anaphylaktischen Reaktionen auf. Der Gruppenunterschied ist mit einem RR von 0,24 statistisch nicht signifikant ([95 % KI] p-Wert: [0,03; 2,08] 0,1926).

In den Teilpopulationen 2 und 3 zeigten sich ebenfalls keine signifikanten Unterschiede.

Die Interaktionstests ergaben keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien. Für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse - Überempfindlichkeit (Woche 24)“ kann somit die relevante Studienpopulation zur Bewertung des Zusatznutzens herangezogen werden.

In der offenen Extensionsphase (Woche 25 bis 48) trat bei 4,0 % der Patienten der relevanten Studienpopulation eine Überempfindlichkeitsreaktion auf (siehe Tabelle 4-164). Die Rate der Überempfindlichkeitsreaktionen stieg damit nach Woche 24 geringfügig an. Dennoch traten Überempfindlichkeitsreaktionen insgesamt selten auf. Die Auswertung nach 48 Wochen bestätigt somit das Sicherheitsprofil von Sarilumab, das sich in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat.

Tabelle 4-164: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – Überempfindlichkeit (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 1 (1,1) | 75 | 3 (4,0) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 0 | 30 | 0 |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 1 (1,9) | 45 | 3 (6,7) |
| Quelle: [92] | | | | |

*Reaktionen der Injektionsstelle*

Tabelle 4-165: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – Reaktionen an der Injektionsstelle (Woche 24)“ aus RCT mit dem zu bewertenden Arzneimittel

| **Studie MONARCH** | | | | | |
|---|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis** | **Effektschätzer[a]** | | | **Interaktionstest[a] (Vortherapie mit 1 / ≥ 2 csDMARDs)** |
| (*Safety*-Population) | n (%) | OR[b] [95% KI] p-Wert | RR [95% KI] p-Wert | RD [95% KI] p-Wert | p-Wert |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | | |
| Sarilumab (N = 87) | 7 (8,0) | 1,71 [0,48; 6,09] 0,4092 | 1,65 [0,50; 5,45] 0,4107 | 0,03 [-0,04; 0,11] 0,4005 | |
| Adalimumab (N = 82) | 4 (4,9) | | | | |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | | OR: n.a. Peto-OR: 0,2316 RR: n.a. RD: 0,5721 |
| Sarilumab (N = 35) | 2 (5,7) | Peto-OR: 7,61 [0,47; 124,15] 0,1543 | n.a. | 0,06 [-0,06; 0,17] 0,3234 | |
| Adalimumab (N = 35) | 0 | | | | |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | | |
| Sarilumab (N = 52) | 5 (9,6) | OR: 1,14 [0,29; 4,58] 0,8489 | 1,13 [0,32; 4,00] 0,8490 | 0,01 [-0,10; 0,13] 0,8484 | |
| Adalimumab (N = 47) | 4 (8,5) | Peto-OR: 1,14 [0,29; 4,47] 0,8493 | | | |

KI: Konfidenzintervall; n.a.: nicht auswertbar; OR: Odds Ratio; RR: Risk Ratio; RD: Risikodifferenz

[a] Effektschätzer, korrespondierende 95%-Konfidenzintervalle, p-Werte und Interaktions-p-Werte basieren auf der SAS-Prozedur PROC GLIMMIX. In das Modell wurden als feste Effekte Behandlung, Subgruppe sowie die Interaktion Subgruppe*Behandlung aufgenommen.

[b] Bei Auftreten einer relativen Häufigkeit ≤ 1% in einem Behandlungsarm wurde Peto-OR anstelle von OR und RR berechnet. In der Spalte 'Interaktionstest' ist für die Heterogenität bzgl. Peto-OR der p-Wert von Cochrane’s Q-Test angegeben.

Quelle: [73]

Die Ergebnisse für die Gesamtpopulation der Studie MONARCH sind in Tabelle 46 des Studienberichts zur randomisierten Behandlungsphase der Studie dargestellt [71].

Reaktionen an der Injektionsstelle traten in der relevanten Studienpopulation bei 8,0 % der mit Sarilumab behandelten Patienten und 4,9 % der mit Adalimumab behandelten Patienten auf (siehe Tabelle 4-165). Dabei handelte es sich vor allem um Erytheme und Schwellungen

der Injektionsstelle von milder bis moderater Intensität. Der Gruppenunterschied ist bei einem RR von 1,65 statistisch nicht signifikant ([95 % KI] p-Wert: [0,50; 5,45] 0,4107).

In den Teilpopulationen 2 und 3 zeigten sich ebenfalls keine signifikanten Unterschiede.

Die Interaktionstests ergaben keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien. Für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse - Überempfindlichkeit (Woche 24)“ kann somit zur Bewertung des Zusatznutzens die relevante Studienpopulation herangezogen werden.

In der offenen Extensionsphase (Woche 25 bis 48) trat bei 10,7 % der Patienten der relevanten Studienpopulation eine Reaktion an der Injektionsstelle auf (siehe Tabelle 4-166). Die Auswertung nach 48 Wochen bestätigt damit das Ergebnis der 24-wöchigen doppelblinden Behandlungsphase.

Tabelle 4-166: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse – Reaktionen der Injektionsstelle (Woche 48)“

| **Studie MONARCH** | | | | |
|---|---|---|---|---|
| **Behandlung** | **Patienten mit Ereignis bis Woche 24** | | **Patienten mit Ereignis in Woche 25 bis 48** | |
| | **N** | n (%) | N | n (%) |
| Relevante Studienpopulation (Patienten mit MTX-Intoleranz) | | | | |
| Sarilumab | 87 | 7 (8,0) | 75 | 8 (10,7) |
| Patienten in Teilpopulation 2 (Vortherapie mit einem csDMARD) | | | | |
| Sarilumab | 35 | 2 (5,7) | 30 | 2 (6,7) |
| Patienten in Teilpopulation 3 (Vortherapie mit ≥ 2 csDMARDs) | | | | |
| Sarilumab | 52 | 5 (9,6) | 45 | 6 (13,3) |
| Quelle: [92] | | | | |

*Weitere unerwünschte Ereignisse von besonderem Interesse*

In der Studie MONARCH waren Thrombozytopenien, Divertikulitis und gastrointestinale Ulzerationen, Krebserkrankungen, demyelinisierende Erkrankungen sowie das Lupus-ähnliche Syndrom als weitere unerwünschte Ereignisse von besonderem Interesse präspezifiziert. Diese Ereignisse traten in der relevanten Studienpopulation jedoch nicht auf.

*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.*

Der Bewertung liegt nur eine randomisierte kontrollierte klinische Studie zugrunde, aus diesem Grund ist eine Meta-Analyse nicht durchführbar.

Die Ergebnisse zum Endpunkt „Sicherheit/Verträglichkeit“ sind auf den deutschen Versorgungskontext übertragbar (Angaben zur Übertragbarkeit der Studienergebnisse siehe Abschnitt 4.3.1.2.1).

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.

#### 4.3.1.3.11 Subgruppenanalysen – RCT

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

*Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.*

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

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

Die Ergebnisse der in der Studie MONARCH a priori geplanten Subgruppenanalysen sind für den Nachweis des Zusatznutzens von Sarilumab nicht relevant, da die a priori geplanten und im Studienbericht dargestellten Subgruppenanalysen die gesamte Studienpopulation der Studie MONARCH betreffen. Diese umfasst auch Patienten, die für die Nutzenbewertung nicht relevant sind (siehe Abschnitt 4.3.1.2.1). Somit ermöglichen die a priori geplanten Subgruppenanalysen keine Aussage zur relevanten Studienpopulation und werden im vorliegenden Dossier nicht dargestellt [71].

Um das Vorliegen einer klinisch relevanten Effektmodifikation in der relevanten Studienpopulation ausschließen zu können, wurden die a priori geplanten Subgruppenanalysen post hoc für die relevante Studienpopulation berechnet (siehe Tabelle 4-167).

Weiterhin wurden für die Nutzenbewertung von Sarilumab post hoc Subgruppenanalysen für weitere Endpunkte, für die keine Subgruppenanalysen a priori geplant waren, nach Geschlecht (Männer / Frauen), Alter (< 65 Jahre / ≥ 65 Jahre) und Region (Region 1: Westliche Staaten / Region 2: Süd-Amerika/ Region 3: andere) durchgeführt (siehe Tabelle 4-168). Für diese post-hoc-Analysen wurde ausschließlich die relevante Studienpopulation herangezogen. Eine Subgruppenanalyse nach Krankheitsschwere wurde als nicht sinnvoll erachtet, da alle Patienten der Studie MONARCH gemäß der im Studienprotokoll definierten Einschlusskriterien eine hohe Krankheitsaktivität aufwiesen.

Für die post hoc geplanten Subgruppenanalysen wurden ausschließlich die binären Endpunkte herangezogen, da diese die Basis für die Bewertung des Zusatznutzens bilden. Die Subgruppenanalysen wurden nur dann berechnet, wenn die Anzahl der Ereignisse basierend auf der relevanten Studienpopulation als ausreichend angesehen wurden. Dieses Vorgehen steht in Übereinstimmung mit dem Vorschlag des IQWiG, Subgruppenanalysen bei binären Daten nur dann durchzuführen, falls mindestens 10 Ereignisse in einer der Subgruppen aufgetreten sind [74].

Für die binären Endpunkte sind die Interaktions-p-Werte auf Basis der Risk Ratios dargestellt, da diese die Basis für die Bewertung des Zusatznutzens bilden. In Fällen, in denen kein RR und damit kein entsprechender Interaktions-p-Wert berechenbar war, wurde der Interaktions-p-Wert auf Basis des Peto-OR herangezogen.

Für die Bewertung möglicher Effektmodifikatoren wurde gemäß IQWiG-Methodik ein Interaktions-p-Wert < 0,05 als Beleg für eine Effektmodifikation durch ein Merkmal gewertet [74]. Wurden nach dieser Methodik zwischen den Subgruppen signifikant unterschiedliche Effekte ($p < 0,05$) ermittelt, werden die Ergebnisse im Folgenden auf Subgruppenebene aufgeführt, verglichen und bewertet.

### *Interaktions-p-Werte der Subgruppenanalysen*

Tabelle 4-167: In der Studie MONARCH a priori spezifizierte Subgruppenanalysen, post-hoc-Auswertung für die relevante Studienpopulation (Woche 24)

| Subgruppenanalysen | p-Werte für Interaktion[a] | |
|---|---|---|
| | **Endpunkt: Änderung des DAS28-ESR** | **Endpunkt: DAS28-ESR-Remission (DAS28-ESR < 2,6)** |
| Geschlecht (Männer / Frauen) | 0,6527 | 0,2148 |
| Rasse (kaukasisch bzw. weiß / andere) | n.a.[b] | n.a.[b] |
| Region (Region 1: Westliche Staaten / Region 2: Süd-Amerika/ Region 3: andere)[c] | 0,5649 | 0,7693[d] |
| Alter (< 65 Jahre / ≥ 65 Jahre) | 0,6180 | 0,9744 |
| Gewicht (< 60 kg / ≥ 60 kg und < 100 kg / ≥ 100 kg) | 0,1540 | 0,6793[d] |
| BMI (< 25 / ≥ 25 und < 30 / ≥ 30 kg/m$^2$) | 0,8910 | 0,4090[d] |
| Rheumafaktor (positiv / negativ) | 0,7949 | 0,9348 |
| Anti-CCP-Antikörper (positiv/ negativ) | 0,4397 | 0,9521 |
| Baseline CRP (≤ 15 mg/L / > 15 mg/L) | **0,0003** | 0,1818 |
| Baseline ESR (≤ Median / > Median) | 0,5360 | 0,3436 |
| Dauer der RA (≤ Median / > Median) | 0,5751 | 0,8047 |
| Dauer der RA (≤ 3 Jahre / > 3 Jahre) | 0,7411 | 0,7414 |
| Anzahl der vorangegangenen DMARDs (≤ 1 / 2 / ≥ 3) | 0,7897 | 0,9440 |
| Raucherstatus (Nicht-Raucher / Raucher / Ex-Raucher) | 0,7393 | 0,8849 |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate; BMI: *Body Mass Index*; CCP: cyclisches citrulliniertes Peptid; CRP: C-reaktives Protein; ESR: Erythrozytensedimentationsrate; RA: rheumatoide Arthritis; DMARD: krankheitsmodifizierendes Antirheumatikum

[a] Sofern nicht anders angegeben beziehen sich die p-Werte für Interaktion beim Endpunkt „Änderung des DAS28-ESR" auf die LSMD und beim Endpunkt „DAS28-ESR-Remission" auf das Risk Ratio.

[b] Nicht auswertbar aufgrund der zu geringen Patientenzahl in einer Subgruppe.

[c] Region 1 (Westliche Länder): Tschechien, Deutschland, Ungarn, Israel, Spanien, USA; Region 2 (Südamerika): Chile, Peru; Region 3 (andere): Südkorea, Polen, Südafrika, Rumänien, Russland, Ukraine

[d] p-Wert für Interaktion bezogen auf das Peto-Odds Ratio

Quellen: [72]

Tabelle 4-168: Post hoc geplante Subgruppenanalysen für die relevante Studienpopulation der Studie MONARCH, Morbidität (Woche 24)

| Endpunkt | p-Werte für Interaktion[a] | | |
|---|---|---|---|
| | **Geschlecht (Männer / Frauen)** | **Alter (< 65 Jahre / ≥ 65 Jahre)** | **Region (Region 1 / 2 / 3)[b]** |
| DAS28-ESR < 3,2 (niedrige Krankheitsaktivität) | 0,7547 | 0,5454 | 0,8432 |
| CDAI ≤ 10 (niedrige Krankheitsaktivität) | 0,7817 | 0,4196 | 0,7518 |
| SDAI ≤ 11 (niedrige Krankheitsaktivität) | 0,7297 | 0,2156 | 0,9638 |
| Ansprechen gemäß EULAR-Kriterien | 0,9677 | 0,8299 | 0,8238 |
| Ansprechen gemäß ACR-Kriterien (ACR20) | 0,8697 | 0,3795 | 0,1144 |
| Ansprechen gemäß ACR-Kriterien (ACR50) | 0,9779 | 0,1075 | 0,6685 |
| Ansprechen gemäß ACR-Kriterien (ACR70) | 0,9568 | 0,5078 | 0,1548[c] |
| Schmerzen (Verbesserung um ≥ 20,4 mm) | 0,4731 | 0,4302 | 0,2987 |
| PtGA (Verbesserung um ≥ 18,4 mm) | 0,9505 | 0,5156 | 0,3076 |
| HAQ-DI (Verbesserung um ≥ 0,22) | 0,8913 | 0,3615 | 0,2297 |
| HAQ-DI (Verbesserung um ≥ 0,375) | 0,8267 | 0,2116 | 0,2125 |
| FACIT-Fatigue (Verbesserung um ≥ 4) | 0,6445 | 0,8397 | 0,5700 |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate; CDAI: *Clinical Disease Activity Index*; SDAI: *Simplified Disease Activity Index*; ACR: *American College of Rheumatology*; EULAR: *European League Against Rheumatism*; FACIT: *Functional Assessment of Chronic Illness Therapy*; HAQ-DI: *Health Assessment Questionnaire – Disability Index*; PtGA: *Patients Global Assessment*

[a] Sofern nicht anders angegeben beziehen sich die p-Werte für Interaktion auf das jeweilige Risk Ratio.

[b] Region 1 (Westliche Länder): Tschechien, Deutschland, Ungarn, Israel, Spanien, USA; Region 2 (Südamerika): Chile, Peru; Region 3 (andere): Südkorea, Polen, Südafrika, Rumänien, Russland, Ukraine

[c] p-Wert für Interaktion bezogen auf das Peto-Odds Ratio

Quellen: [72]

Tabelle 4-169: Post hoc geplante Subgruppenanalysen für die relevante Studienpopulation der Studie MONARCH, Lebensqualität (Woche 24)

| Endpunkt | p-Werte für Interaktion[a] | | |
|---|---|---|---|
| | **Geschlecht (Männer / Frauen)** | **Alter (< 65 Jahre / ≥ 65 Jahre)** | **Region (Region 1 / 2 / 3)[b]** |
| SF-36-PCS (Verbesserung um ≥ 2,5) | 0,4064 | 0,1643 | 0,3130 |
| SF-36-PCS (Verbesserung um ≥ 5,1) | 0,6878 | 0,2363 | 0,4711 |
| SF-36-PCS (Verbesserung um ≥ 7,2) | 0,4799 | 0,2677 | 0,0971 |
| RAID (Verbesserung um ≥ 3) | 0,7599 | 0,8273 | 0,0737 |

SF-36-PCS: *Physical Component Summary Score* des SF-36; RAID: *Rheumatoid Arthritis Impact of Disease*

[a] Sofern nicht anders angegeben beziehen sich die p-Werte für Interaktion auf das jeweilige Risk Ratio.

[b] Region 1 (Westliche Länder): Tschechien, Deutschland, Ungarn, Israel, Spanien, USA; Region 2 (Südamerika): Chile, Peru; Region 3 (andere): Südkorea, Polen, Südafrika, Rumänien, Russland, Ukraine

Quelle: [72]

Tabelle 4-170: Post hoc geplante Subgruppenanalysen für die relevante Studienpopulation der Studie MONARCH, Sicherheit (Woche 24)

| Endpunkt | p-Werte für Interaktion[a] | | |
|---|---|---|---|
| | **Geschlecht (Männer / Frauen)** | **Alter (< 65 Jahre / ≥ 65 Jahre)** | **Region (Region 1 / 2 / 3)[b]** |
| Gesamtrate UE | 0,5894 | 0,0819 | 0,6193 |
| Infektionen | 0,9890 | 0,2816 | 0,7583 |
| Leukopenien | 0,4357[c] | 0,4936[c] | 0,0518[c] |

UE: unerwünschte Ereignisse

[a] Sofern nicht anders angegeben beziehen sich die p-Werte für Interaktion auf das jeweilige Risk Ratio.

[b] Region 1 (Westliche Länder): Tschechien, Deutschland, Ungarn, Israel, Spanien, USA; Region 2 (Südamerika): Chile, Peru; Region 3 (andere): Südkorea, Polen, Südafrika, Rumänien, Russland, Ukraine

[c] p-Wert für Interaktion bezogen auf das Peto-Odds Ratio

Quellen: [73]

***Detaildarstellung der Subgruppenanalysen mit p-Wert < 0,05***

Tabelle 4-171: Subgruppenanalyse nach Baseline-CRP für den Endpunkt Änderung des DAS28-ESR (Woche 24)

| **Studie MONARCH** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Behandlung** | **DAS28-ESR-Basiswert** | | **DAS28-ESR in Woche 24** | | **absolute Änderung im Vergleich zum Basiswert** | **Behandlungs-differenz[a]** | **Interaktions-test[a]** |
| (ITT) | N | MW (SD) | N | MW (SD) | MD (SD) | LSMD [95% KI] p-Wert **Hedges´g [95% KI]** | p-Wert |
| Baseline CRP ≤ 15 mg/L | | | | | | | |
| Sarilumab (N = 35) | 49 | 6,77 (0,57) | 49 | 3,54 (1,40) | -3,24 (1,33) | -0,651 [-1,184; -0,118] 0,0172 **-0,503 [-0,915; -0,091]** | **0,0003** |
| Adalimumab (N = 35) | 41 | 6,59 (0,74) | 41 | 4,23 (1,22) | -2,36 (1,36) | | |
| Baseline CRP > 15 mg/L | | | | | | | |
| Sarilumab (N = 52) | 27 | 7,03 (0,68) | 27 | 2,92 (1,18) | -4,11 (0,95) | -2,230 [-2,816; -1,645] <0,0001 **-2,002 [-2,527; -1,476]** | |
| Adalimumab (N = 47) | 30 | 7,02 (0,86) | 30 | 5,25 (1,14) | -1,77 (1,27) | | |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate; ITT: *Intent-to-treat*; KI: Konfidenzintervall; LSMD: *Least Squares mean difference* - modellbasierte adjustierte Mittelwertdifferenz; MW: Mittelwert; MD: Mittelwert der Differenzen Woche 24 minus Basiswert; SD: Standardabweichung

[a] Für die Auswertung wurde eine Subgruppenanalyse mittels MMRM-Modell durchgeführt. Das Modell enthielt als feste Effekte Behandlung, Region und Subgruppe, den Messwiederholungsfaktor Visite sowie die Interaktionsterme Behandlung*Visite, Behandlung*Subgruppe und Behandlung*Subgruppe*Visite. Der Baseline-Wert wurde als stetige Kovariate in das Modell aufgenommen.

Quelle: [72]

Die Subgruppenanalyse ergab für den Endpunkt „Änderung des DAS28-ESR“ einen Beleg für eine Effektmodifikation durch die Höhe des CRP-Wertes zu Beginn der Studie. Die Patienten mit einem CRP > 15 mg/L zeigten nach 24 Wochen Behandlung mit Sarilumab eine stärkere Senkung des DAS28-ESR als die Patienten mit einem CRP ≤ 15 mg/L (siehe Tabelle 4-171). Dennoch bestand in beiden Subgruppen ein signifikanter Vorteil gegenüber der Behandlung mit Adalimumab. Es handelt sich also lediglich um eine quantitative Effektmodifikation. Bei der Subgruppenanalyse des Endpunktes „DAS28-ESR-Remission“ zeigte sich zudem keine Effektmodifikation durch die Höhe des CRP-Wertes zu Beginn der Studie (p = 0,1818).

***Fazit zu Subgruppenanalysen***

Für den Nachweis des Zusatznutzens von Sarilumab sind die vorstehend dargestellten Subgruppenanalysen nicht von Bedeutung, da sich keine klinisch relevante Effektmodifikation ergab. Somit kann die Bewertung des Zusatznutzens auf Basis der gesamten relevanten Studienpopulation erfolgen.

### 4.3.1.3.12 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

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

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

Für das vorliegende Nutzenbewertungsdossier wurden drei Patientenpopulationen der Studie MONARCH ausgewertet. Die relevante Studienpopulation umfasst alle Patienten der Studie mit einer ärztlich bestätigten MTX-Intoleranz. Für diese Population wurde eine Subgruppenanalyse nach Anzahl der csDMARD-Vortherapien durchgeführt, um zwischen Patienten zu unterscheiden, die vor Eintritt in die Studie mit einem csDMARD (Teilpopulation 2) oder mindestens zwei csDMARDs (Teilpopulation 3) behandelt wurden. Für jeden Endpunkt wurden die Interaktions-p-Werte betrachtet. Ein Interaktions-p-Wert < 0,05 wurde als Beleg für eine Effektmodifikation gewertet. Da für keinen der Endpunkte eine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien bestand, können die Ergebnisse der relevanten Studienpopulation für die Ableitung des Zusatznutzens für die Teilanwendungsgebiete 2 und 3 gleichermaßen herangezogen werden. Im Folgenden sind somit die Ergebnisse für die patientenrelevanten Endpunkte der relevanten Studienpopulation zusammengefasst.

In den Teilpopulationen 2 und 3 zeigten sich in der Regel Ergebnisse, die hinsichtlich der Effektrichtung mit den Ergebnissen der relevanten Studienpopulation vergleichbar waren. Allerdings waren die Gruppenunterschiede in den Teilpopulationen im Gegensatz zur relevanten Studienpopulation aufgrund der geringeren Patientenzahlen weniger häufig statistisch signifikant.

***Morbidität***

*Änderung der Krankheitsaktivität*

In der Studie MONARCH wurde die Krankheitsaktivität der Patienten mittels DAS28-ESR, und CDAI ermittelt. Post hoc wurde zusätzlich eine Auswertung des SDAI vorgenommen. Die Änderung der Krankheitsaktivität nach 24 Wochen gemessen mittels DAS28-ESR war der primäre Endpunkt der Studie.

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verminderung der Krankheitsaktivität unabhängig vom verwendeten Messinstrument.

Bezüglich des DAS28-ESR zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung der Krankheitsaktivität (LSMD [95% KI] p-Wert: -1,253 [-1,660; -0,846] < 0,0001). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g belegt, dass es sich dabei um einen klinisch relevanten Effekt mit mittlerer Effektstärke handelt (Hedges´g [95% KI]: -0,990 [-1,311; -0,668]).

Auch bezüglich des CDAI zeigte sich nach 24-wöchiger Behandlung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung (LSMD [95% KI] p-Wert: -5,005 [-7,998; -2,011] 0,0012). Dabei handelt es sich ebenfalls um einen klinisch relevanten Effekt mit kleiner Effektstärke (Hedges´g [95% KI]: -0,536 [-0,856; -0,215]).

Des Weiteren war bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung des SDAI zu beobachten (LSMD [95% KI] p-Wert: -5,922 [-9,116; -2,727] 0,0003). Dabei handelt es sich ebenfalls um einen klinisch relevanten Effekt mit kleiner Effektstärke (Hedges´g [95% KI]: -0,596 [-0,917; -0,274]).

*Remission*

In der Studie MONARCH wurden verschiedene Operationalisierungen für den Endpunkt Remission ausgewertet.

Das Kriterium DAS28-ESR < 2,6 erfüllten 29,9 % der Patienten unter Sarilumab, wohingegen nur 6,1 % der Patienten unter Adalimumab diesen Schwellenwert unterschritten. Damit war unter Sarilumab die Wahrscheinlichkeit, eine Remission gemäß DAS28-ESR zu erreichen, gegenüber Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 4,39 [1,78; 10,85] 0,0014).

Bei Anwendung der strengeren CDAI- und SDAI- sowie der booleschen ACR/EULAR-Kriterien für eine Remission erreichten in beiden Behandlungsarmen deutlich weniger Patienten die Remission als bei Anwendung des DAS28. Für alle drei Remissions-Kriterien zeigte sich nach 24 Wochen Behandlung ein numerischer Vorteil für Sarilumab gegenüber Adalimumab. Der Behandlungsunterschied war jedoch in keinem Fall signifikant (RR [95 % KI] p-Wert für CDAI-Remission: 3,37 [0,73; 15,45] 0,1182; für SDAI-Remission: 3,75 [0,83; 16,88] 0,0844; für ACR/EULAR-Remission: 2,20 [0,43; 11,16] 0,3423).

*Niedrige Krankheitsaktivität*

Das Erreichen einer niedrigen Krankheitsaktivität stellt neben der Remission ein relevantes Therapieziel dar, insbesondere für Patienten mit einer langjährigen schweren RA, wie sie in die MONARCH Studie eingeschlossen wurden. Nach 24-wöchiger Behandlung zeigte sich

mit allen drei Messinstrumenten ein deutlicher Vorteil für Sarilumab gegenüber Adalimumab bei der Zahl der Patienten, die eine niedrige Krankheitsaktivität erreichten.

Nach 24 Wochen Behandlung erzielten 44,8 % der mit Sarilumab behandelten Patienten und 9,8 % der mit Adalimumab behandelten Patienten einen DAS28-ESR < 3,2. Dieser Behandlungsunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 4,22 [2,10; 8,46] < 0,0001).

Die Wahrscheinlichkeit, nach 24 Wochen einen CDAI ≤ 10 oder SDAI ≤ 11 zu erreichen, war unter der Behandlung mit Sarilumab ebenfalls signifikant größer als unter Behandlung mit Adalimumab (RR [95 % KI] p-Wert für CDAI: 2,39 [1,44; 3,97] 0,0008 und RR [95 % KI] p-Wert für SDAI: 2,29 [1,40; 3,74] 0,0010). Die Behandlungseffekte sind über alle drei Messinstrumente konsistent.

*Therapieansprechen*

Zur Quantifizierung des Therapieansprechens wurden die Kriterien der EULAR, basierend auf dem DAS28-ESR, und die Kriterien des ACR angewendet. Nach 24-wöchiger Behandlung zeigte sich über alle Auswertungen ein Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich des Anteils der Patienten mit einem Ansprechen auf die Behandlung.

Die Wahrscheinlichkeit für ein gutes bis moderates Ansprechen gemäß EULAR-Kriterien war unter Sarilumab im Vergleich zu Adalimumab signifikant erhöht (RR [95 % KI] p-Wert: 1,22 [1,02; 1,46] 0,0328). Dabei erreichten 44,8 % der mit Sarilumab behandelten Patienten und 9,8 % der mit Adalimumab behandelten Patienten ein gutes Ansprechen. Demgegenüber zeigten 32,9 % der mit Adalimumab behandelten Patienten kein Ansprechen gemäß EULAR-Kriterien, während es unter Sarilumab nur 18,4 % waren.

Die ACR-Kriterien für das Therapieansprechen definieren Schwellenwerte für eine 20%ige, 50%ige und 70%ige Verbesserung der Krankheitsaktivität (ACR20, ACR50, ACR70). Für die Patienten der Studie MONARCH, die bei Studieneinschluss sämtlich eine hohe Krankheitsaktivität aufwiesen, stellt eine Verbesserung um 20 % jedoch vermutlich keinen ausreichenden Therapieerfolg dar. Aus diesem Grund wird die Auswertung zum ACR20 als nicht patientenrelevant eingestuft, aber aus Transparenzgründen mit dargestellt. Eine 50%ige oder gar 70%ige Verringerung der Krankheitsaktivität ist hingegen als klinisch relevantes Therapieansprechen einzuschätzen.

Die Wahrscheinlichkeit, ACR20 zu erreichen, war unter Sarilumab im Vergleich zu Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 1,29 [1,03; 1,61] 0,0296). Die Wahrscheinlichkeit, ACR50 zu erreichen, war ebenfalls unter Sarilumab im Vergleich zu Adalimumab statistisch signifikant erhöht (RR [95 % KI] p-Wert: 1,73 [1,14; 2,63] 0,0104). 47,1 % der Patienten unter Sarilumab erreichten ACR50 gegenüber 26,8 % der Patienten unter Adalimumab. Auch die Wahrscheinlichkeit, ACR70 zu erreichen, war unter Sarilumab statistisch signifikant größer als unter Adalimumab (RR [95 % KI] p-Wert: 2,62 [1,17; 5,85]

0,0194). 24,1 % der Patienten unter Sarilumab und 8,5 % der Patienten unter Adalimumab erreichten ACR70.

*Gelenksymptome*

Die Anzahl der druckschmerzhaften Gelenke (TJC, *Tender Joint Counts*) wurde basierend auf 68 und 28 definierten Gelenken bestimmt. Bei der Auswertung der druckschmerzhaften Gelenke ausgehend sowohl von 68 als auch 28 Gelenken zeigte sich nach 24 Wochen Behandlung jeweils ein numerischer Vorteil für Sarilumab gegenüber Adalimumab, jedoch keine statistisch signifikanten Unterschiede (LSMD [95 % KI] p-Wert für 68TJC: -1,281 [-4,314; 1,752] 0,4056 und LSMD [95 % KI] p-Wert für 28TJC: -1,452 [-3,126; 0,222] 0,0887).

Die Anzahl der geschwollenen Gelenke (SJC, *Swollen Joint Counts*) wurde basierend auf 66 und 28 definierten Gelenken bestimmt. Bei der Auswertung der geschwollenen Gelenke ausgehend von 66 Gelenken zeigte sich nach 24 Wochen Behandlung ein numerischer Vorteil für Sarilumab gegenüber Adalimumab, jedoch kein statistisch signifikanter Unterschied (LSMD [95 % KI] p-Wert: -1,513 [-3,080; 0,054] 0,0584). Bei der Auswertung der geschwollenen Gelenke ausgehend von 28 Gelenken zeigte sich jedoch ein statistisch signifikanter Vorteil für Sarilumab gegenüber Adalimumab (LSMD [95 % KI] p-Wert: -1,333 [-2,336; -0,329] 0,0095). Anhand der standardisierten Mittelwertdifferenz Hedges`g kann allerdings keine Aussage zur klinischen Relevanz dieses Effekts abgeleitet werden, da das Konfidenzintervall den Wert -0,2 einschließt (Hedges´g [95% KI]: -0,431 [-0,755; -0,107]).

*Schmerzen*

Nach 24 Wochen Behandlung zeigte sich eine Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verringerung von Schmerzen.

Die Schmerzintensität zu Beginn der Studie wurde von den Patienten durchschnittlich mit etwa 71 mm auf der 100 mm-VAS angegeben, so dass man davon ausgehen kann, dass die Schmerzen zu Beginn der Studie als starke Schmerzen empfunden wurden. Nach 24-wöchiger Behandlung zeigte sich bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Verringerung der Schmerzen (LSMD [95% KI] p-Wert: -11,437 [-18,456; -4,419] 0,0016). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges`g belegt, dass es sich dabei um einen klinisch relevanten Effekt mit kleiner Effektstärke handelt (Hedges´g [95% KI]: -0,524 [-0,846; -0,202]).

In einer Responderanalyse wurde des Weiteren untersucht, wie viele Patienten eine minimale klinische relevante Verringerung ihrer Schmerzen um mindestens 20,4 mm erreichten. Dieses Kriterium erfüllten 52,4 % der Patienten unter Adalimumab und 58,6 % der Patienten unter Sarilumab. Dieser Behandlungsunterschied für die minimale klinische relevante Verbesserung ist statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,07 [0,83; 1,39] 0,5931).

*Globale Einschätzung des Patienten (PtGA)*

Nach 24-wöchiger Behandlung zeigte sich hinsichtlich der globalen Einschätzung der Patienten zu ihrer Krankheitsaktivität (PtGA) ein Vorteil von Sarilumab gegenüber Adalimumab.

In der relevanten Studienpopulation ergab sich ein statistisch signifikanter Behandlungsunterschied zugunsten von Sarilumab (LSMD [95 % KI] p-Wert: -12,677 [-19,486; -5,868] 0,0003). Die Betrachtung der standardisierten Mittelwertdifferenz belegt, dass es sich dabei um einen klinisch relevanten Effekt mit kleiner Effektstärke handelt (Hedges´g [95% KI]: -0,597 [-0,917; -0,276]).

In einer Responderanalyse wurde des Weiteren untersucht, wie viele Patienten eine minimale klinisch relevante Verbesserung beim PtGA um mindestens 18,4 mm erreichten. Dieses Kriterium erfüllten 50,0 % der Patienten unter Adalimumab und 60,9 % der Patienten unter Sarilumab. Dieser Behandlungsunterschied für eine minimale klinisch relevante Verbesserung ist statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,18 [0,90; 1,55] 0,2196).

*Körperlicher Funktionsstatus – HAQ-DI*

Nach 24 Wochen Behandlung zeigte sich eine deutliche Überlegenheit von Sarilumab gegenüber Adalimumab bezüglich der Verbesserung der körperlichen Funktion gemessen mittels HAQ-DI.

Die Patienten unter Sarilumab erreichten im Vergleich zu den Patienten unter Adalimumab eine statistisch signifikant stärkere Senkung des HAQ-DI nach 24 Wochen (LSMD [95% KI] p-Wert: -0,214 [-0,406; -0,022] 0,0293). Die klinische Relevanz dieser mittleren Verbesserung kann jedoch nicht abgeleitet werden kann, da das Konfidenzintervall der standardisierten Mittelwertdifferenz den Wert -0,2 einschließt (Hedges´g [95% KI]: -0,352 [-0,668; -0,036]).

Des Weiteren wurden zwei Responderanalysen bezüglich einer minimalen klinisch relevanten Verbesserung um 0,22 und 0,375 Punkte durchgeführt.

66,7 % der Patienten unter Sarilumab und 48,8 % der Patienten unter Adalimumab erreichten eine Verbesserung im HAQ-DI um mindestens 0,22 Punkte. Die Wahrscheinlichkeit, eine Verbesserung um 0,22 Punkte zu erreichen, war damit in der Sarilumab-Gruppe signifikant höher als in der Adalimumab-Gruppe (RR [95% KI] p-Wert: 1,37 [1,05; 1,78] 0,0213).

Eine Verbesserung im HAQ-DI um mindestens 0,375 Punkte erreichten 62,1 % der Patienten unter Sarilumab gegenüber 41,5 % der Patienten unter Adalimumab. Dieser Gruppenunterschied war ebenfalls statistisch signifikant (RR [95% KI] p-Wert: 1,49 [1,11; 2,01] 0,0090).

*Morgensteifigkeit*

Nach 24-wöchiger Behandlung zeigte sich ein Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verminderung der Morgensteifigkeit.

Die Behandlung mit Sarilumab führte zu einer signifikant stärkeren Verringerung der Morgensteifigkeit als die Behandlung mit Adalimumab (LSMD [95 % KI] p-Wert: -10,967 [-18,841; -3,094] 0,0066). Die klinische Relevanz dieses Gruppenunterschieds kann jedoch nicht abgeleitet werden, da das Konfidenzintervall der standardisierten Mittelwertdifferenz Hedges´g den Wert -0,2 einschließt (Hedges´g [95 % KI]: -0,456 [-0,783; -0,129]).

*Fatigue*

Nach 24 Wochen Behandlung zeigte sich eine Überlegenheit von Sarilumab gegenüber Adalimumab hinsichtlich der Verringerung der Fatigue gemessen mittels FACIT-Fatigue Fragebogen.

Bei Auswertung des FACIT-Fatigue zeigte sich eine signifikant stärkere Verbesserung bei den Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab (LSMD [95 % KI] p-Wert: 3,711 [0,814; 6,608] 0,0124). Die klinische Relevanz dieser mittleren Verbesserung kann jedoch nicht abgeleitet werden kann, da das Konfidenzintervall der standardisierten Mittelwertdifferenz den Wert 0,2 einschließt (Hedges´g [95 % KI]: 0,415 [0,091; 0,740]).

Des Weiteren wurde eine Responderanalyse durchgeführt, um den Anteil der Patienten, die eine minimale klinisch relevante Verbesserung im FACIT-Fatigue um mindestens 4 Punkte erreicht haben, zu bestimmen. Dieses Kriterium erfüllten in Woche 24 66,7 % der Patienten unter Sarilumab, während es im Adalimumab-Arm 52,4 % der Patienten waren. Bei Betrachtung des RR ist dieser Gruppenunterschied statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,27 [0,99; 1,64] 0,0631).

*Änderung des allgemeinen Gesundheitszustandes*

Nach 24-wöchiger Behandlung zeigte sich ein Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verbesserung des allgemeinen Gesundheitszustandes gemessen mittels EQ-5D-VAS.

In der relevanten Studienpopulation erreichten die Patienten unter Sarilumab im Vergleich zu den Patienten unter Adalimumab eine signifikant stärkere Verbesserung ihres allgemeinen Gesundheitszustandes (LSMD [95 % KI] p-Wert: 9,244 [2,683; 15,806] 0,0061). Die klinische Relevanz dieses Behandlungsunterschieds bleibt jedoch unklar, da das Konfidenzintervall der standardisierten Mittelwertdifferenz den Wert 0,2 einschließt (Hedges´g [95 % KI]: 0,461 [0,134; 0,788]).

***Gesundheitsbezogene Lebensqualität***

*SF-36*

Die Auswertung der körperlichen Summenskala des SF-36 (SF-36-PCS) ergab nach 24-wöchiger Behandlung einen numerischen Vorteil von Sarilumab gegenüber Adalimumab.

Die mittlere Lebensqualität gemäß SF-36-PCS verbesserte sich unter Sarilumab signifikant stärker als unter Adalimumab (LSMD [95 % KI] p-Wert: 3,640 [1,399; 5,882] 0,0016). Die

Betrachtung der standardisierten Mittelwertdifferenz Hedges´g zeigt, dass es sich dabei um einen klinisch relevanten Effekt von kleiner Effektstärke handelt (Hedges´g [95 % KI]: 0,532 [0,204; 0,860]).

Für Patienten mit rheumatoider Arthritis wurden für die körperliche Summenskala des SF-36 verschiedene MCIDs (2,5, 5,1 und 7,2 Punkte) als klinisch relevante Verbesserung definiert. Der Anteil der Patienten, die diese Schwellenwerte nach 24-wöchiger Behandlung erreichten, wurde in Responderanalysen bestimmt.

63,2 % der Patienten unter Sarilumab erzielten nach 24 Wochen eine Verbesserung der SF-36-PCS um mindestens 2,5 Punkte, während es unter Adalimumab 51,2 % der Patienten waren. Bei Betrachtung des RR ist dieser Gruppenunterschied statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,23 [0,95; 1,59] 0,1138).

Eine Verbesserung der SF-36-PCS um mindestens 5,1 Punkte erreichten nach 24-wöchiger Behandlung 57,5 % der Patienten unter Sarilumab, während es unter Adalimumab 43,9 % der Patienten waren. Dieser Behandlungsunterschied ist statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,30 [0,97; 1,75] 0,0803).

Des Weiteren erzielten 49,4 % der Patienten unter Sarilumab eine Verbesserung der SF-36-PCS um mindestens 7,2 Punkte, während es unter Adalimumab 35,4 % der Patienten waren. Dieser Behandlungsunterschied ist bei Betrachtung des RR statistisch nicht signifikant (RR [95 % KI] p-Wert: 1,38 [0,96; 1,97] 0,0788).

Die Auswertung der psychischen Summenskala des SF-36 (SF-36-MCS) nach 24 Wochen Behandlung zeigte in beiden Behandlungsarmen eine Verbesserung von vergleichbarem Ausmaß. Es ergaben sich keine statistisch signifikanten Unterschiede zugunsten einer der beiden Therapien (LSMD [95 % KI] p-Wert: 2,437 [-0,806; 5,681] 0,1397).

*RAID*

Nach 24-wöchiger Behandlung zeigte sich ein deutlicher Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der Verbesserung der Lebensqualität gemessen mittels des krankheitsspezifischen Fragebogens RAID.

Die Patienten unter Sarilumab erreichten nach 24 Wochen Behandlung eine statistisch signifikant stärkere Verbesserung ihrer Lebensqualität als die Patienten unter Adalimumab (LSMD [95 % KI] p-Wert: -1,346 [-2,022; -0,671] 0,0001). Die Betrachtung der standardisierten Mittelwertdifferenz Hedges´g zeigt, dass es sich dabei um einen klinisch relevanten Effekt von kleiner Effektstärke handelt (Hedges´g [95 % KI]: -0,651 [-0,978; -0,325]).

Des Weiteren wurde eine Responderanalyse hinsichtlich der minimalen klinisch relevanten Verbesserung des RAID um mindestens 3 Punkte durchgeführt. Dieses Kriterium erfüllten 49,4 % der Patienten unter Sarilumab, während es unter Adalimumab 30,5 % der Patienten

waren. Dieser Gruppenunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 1,62 [1,10; 2,37] 0,0142).

***Sicherheit/Verträglichkeit***

Zusammenfassend zeigte sich ein vergleichbares Sicherheitsprofil von Sarilumab und Adalimumab innerhalb der 24-wöchigen randomisierten Behandlungsphase.

Die Gesamtrate der unerwünschten Ereignisse war in beiden Behandlungsgruppen nach 24 Wochen vergleichbar. Bei 66,7 % der mit Sarilumab behandelten Patienten sowie bei 68,3 % der mit Adalimumab behandelten Patienten wurden unerwünschte Ereignisse dokumentiert. Somit ergab sich bei einem RR von 0,98 kein statistisch signifikanter Gruppenunterschied ([95 % KI] p-Wert: [0,79; 1,20] 0,8216). Am häufigsten wurden in beiden Behandlungsarmen Infektionen verzeichnet. In der Mehrzahl der Fälle handelte es sich dabei um Atemwegsinfekte. Weitere häufige unerwünschte Ereignisse bei den mit Sarilumab behandelten Patienten waren Erkrankungen des Blutes und des Lymphsystems. Dabei handelte es sich fast ausschließlich um Neutropenien.

Schwerwiegende unerwünschte Ereignisse traten innerhalb der 24-wöchigen randomisierten Behandlungsphase selten auf. Es ergab sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (RR [95 % KI] p-Wert: 0,75 [0,21; 2,72] 0,6657).

Therapieabbrüche aufgrund von unerwünschten Ereignissen traten innerhalb der 24-wöchigen randomisierten Behandlungsphase bei 8,0 % der mit Sarilumab behandelten und 3,7 % der mit Adalimumab behandelten Patienten auf. Der Gruppenunterschied war bei einem RR von 2,2 statistisch nicht signifikant ([95 % KI] p-Wert: [0,59; 8,26] 0,2421). Die häufigste Ursache für Therapieabbrüche unter Sarilumab waren Neutropenien.

Es gab einen Todesfall im Sarilumab-Arm innerhalb der 24-wöchigen Behandlungsphase. Dabei handelte es sich um ein akutes Herzversagen nach Aortendissektion.

Als unerwünschte Ereignisse von besonderem Interesse wurden Infektionen, Leukopenien, Thrombozytopenien, Leberwerterhöhungen, Divertikulitis, gastrointestinale Ulzerationen, Erhöhung der Lipidwerte, Überempfindlichkeit, Anaphylaxie, Reaktionen an der Injektionsstelle, Krebserkrankungen, Lupus-ähnliches Syndrom und demyelinisierende Erkrankung präspezifiziert. Für die meisten dieser Ereignisse ist bekannt, dass die Behandlung mit einem TNF- oder IL-6R-Antagonisten das Risiko für diese Ereignisse erhöht.

Infektionen traten in der doppelblinden Behandlungsphase bei 28,7 % der mit Sarilumab behandelten bzw. 28,0 % der mit Adalimumab behandelten Patienten auf. Das Risiko für das Auftreten einer Infektion war somit in den Behandlungsgruppen vergleichbar (RR [95 % KI] p-Wert: 1,02 [0,63; 1,66] 0,9212). Bei zwei Patienten im Sarilumab-Arm wurden die Infektionen als schwerwiegende Infektionen eingestuft. In beiden Fällen war die Infektion nicht mit einer Neutropenie assoziiert.

Leukopenien traten innerhalb von 24 Wochen unter Sarilumab bei 16,1 % der Patienten auf, unter Adalimumab waren im Gegensatz nur 2,4 % der Patienten betroffen. Bis auf einen Fall handelte es sich dabei um Neutropenien. Der Gruppenunterschied ist statistisch signifikant (RR [95 % KI] p-Wert: 6,60 [1,54; 28,3] 0,0112). Ein Fall im Sarilumab-Arm wurde als schwerwiegend eingestuft. Bei diesem Patienten lag eine Überdosierung von Sarilumab aufgrund der Nichteinhaltung des Behandlungsintervalls von zwei Wochen vor.

Eine Erhöhung der Lebertransaminasen trat bei 6,9 % der Patienten im Sarilumab-Arm und bei 3,7 % der Patienten im Adalimumab-Arm auf. Der Gruppenunterschied ist bei einem RR von 1,89 statistisch nicht signifikant ([95 % KI] p-Wert: [0,49; 7,32] 0,3589).

Eine Lipidwerterhöhung wurde lediglich im Adalimumab-Arm bei 4,9 % der Patienten festgestellt, unter Sarilumab traten keine Lipidwerterhöhungen auf. Der Gruppenunterschied zwischen Sarilumab und Adalimumab ist statistisch signifikant (Peto-OR [95 % KI] p-Wert: 0,12 [0,02; 0,89] 0,0376).

Überempfindlichkeit trat bei 1,1 % der Sarilumab-Patienten und 4,9 % der Adalimumab-Patienten auf. Der Gruppenunterschied ist mit einem RR von 0,24 statistisch nicht signifikant ([95 % KI] p-Wert: [0,03; 2,08] 0,1926). In der Studie traten keine anaphylaktischen Reaktionen auf.

Reaktionen an der Injektionsstelle traten bei 8,0 % der mit Sarilumab behandelten Patienten und 4,9 % der mit Adalimumab behandelten Patienten auf. Der Gruppenunterschied ist bei einem RR von 1,65 statistisch nicht signifikant ([95 % KI] p-Wert: [0,50; 5,45] 0,4107).

Thrombozytopenien, Divertikulitis und gastrointestinale Ulzerationen, Krebserkrankungen, demyelinisierende Erkrankungen sowie das Lupus-ähnliche Syndrom traten in der 24-wöchigen Behandlungsphase nicht auf.

***Auswertungen nach 48 Wochen (offene Extensionsphase)***

Die Ergebnisse der offenen Extensionsphase zeigen eine langanhaltende Wirksamkeit von Sarilumab in Bezug auf die Verbesserung der Morbidität und Lebensqualität sowie ein im Vergleich zur 24-wöchigen doppelblinden Behandlungsphase gleichbleibendes Sicherheitsprofil. Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bestätigen somit die oben zusammengefassten Ergebnisse der 24-wöchigen doppelblinden Behandlungsphase.

*Morbiditätsendpunkte*

In der offenen Extensionsphase kam es unter Sarilumab zu einer weiteren Verringerung der Krankheitsaktivität gemessen mittels DAS28-ESR. Auch die Anzahl der Patienten, die eine Remission erreichten, stieg nach Woche 24 weiter an, unabhängig von den angewendeten Remissionskriterien (DAS28-ESR, CDAI, SDAI sowie den booleschen ACR/EULAR-Kriterien). Die Zahl der Patienten, die nach 48 Wochen Behandlung eine niedrige Krankheitsaktivität, gemessen mittels DAS28-ESR, CDAI oder SDAI, erreichten, nahm ebenfalls in der offenen Extensionsphase weiter zu. Der Anteil der Patienten mit einem guten

Therapieansprechen gemäß EULAR-Kriterien stieg ebenso wie der Anteil der Patienten mit einem 50%igen oder 70%igen Therapieansprechen gemäß ACR-Kriterien weiter an.

Die Auswertungen nach 48 Wochen hinsichtlich einer klinisch relevanten Verbesserung der Schmerzen und der globalen Einschätzung der Patienten (PtGA) zeigten gleichbleibende Ergebnisse im Vergleich zur Auswertung nach Woche 24. Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bezüglich einer klinisch relevanten Verbesserung des körperlichen Funktionsstatus gemäß HAQ-DI sowie der Fatigue gemäß FACIT-Fatigue zeigten lediglich einen leichten Rückgang der jeweiligen Responderzahlen im Vergleich zu den Auswertungen nach 24 Wochen.

*Gesundheitsbezogene Lebensqualität*

Die Auswertungen hinsichtlich einer klinisch relevanten Verbesserung der Lebensqualität unter Sarilumab zeigten nach 48 Wochen für den RAID-Fragebogen einen leichten Rückgang der Responderzahlen im Vergleich zu den Auswertungen nach 24 Wochen und gleichbleibende Ergebnisse für die körperliche Summenskala des SF-36.

*Sicherheit/Verträglichkeit*

Die Auswertung der unerwünschten Ereignisse bis Woche 48 bestätigte das Sicherheitsprofil von Sarilumab, das sich bereits in der 24-wöchigen doppelblinden Behandlungsphase gezeigt hat.

Die Gesamtraten der UE und der SUE sowie der Therapieabbrüche aufgrund von UE in der offenen Extensionsphase (Woche 25 bis 48) blieben im Vergleich zu den ersten 24 Wochen der Behandlung konstant. Es kam nicht zu einer Zunahme schwerwiegender unerwünschter Ereignisse. Todesfälle traten bei den Patienten der relevanten Studienpopulation in der offenen Extensionsphase nicht auf.

Die Zahl der Infektionen blieb bis Woche 48 konstant und es war keine Zunahme schwerwiegender Infektionen zu beobachten. Die Rate der Leukopenien war in der offenen Extensionsphase etwas niedriger als die Rate der Leukopenien innerhalb der 24-wöchigen doppelblinden Behandlungsphase. Keine der aufgetretenen Leukopenien nach Woche 24 wurde als schwerwiegend eingestuft. Auch die EMA hat für die Gesamtschau der Zulassungsstudien von Sarilumab bestätigt, dass sowohl die Raten der Neutropenien und schwerwiegenden Neutropenien als auch die Raten der Infektionen und schwerwiegenden Infektionen im Zeitverlauf nicht ansteigen [89].

Bezüglich der weiteren UE von besonderem Interesse (Leberwerterhöhungen, Lipidwerterhöhungen, Überempfindlichkeitsreaktionen sowie Reaktionen an der Injektionsstelle) zeigten sich keine relevanten Änderungen nach Woche 24.

***Subgruppenanalysen***

Die post hoc durchgeführten Subgruppenanalysen ergaben keine klinisch relevante Effektmodifikation. Somit kann die Bewertung des Zusatznutzens auf Basis der gesamten relevanten Studienpopulation erfolgen.

### 4.3.2 Weitere Unterlagen

#### 4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

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

##### 4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche.* ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

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

Es werden keine indirekten Vergleiche durchgeführt.

##### 4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

*Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial.* ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

Es werden keine indirekten Vergleiche durchgeführt.

#### 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 4-172: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

#### 4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in 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 4-173: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

| **Anzahl Studien** | **Studie** | **Intervention** | **<Vergleichs-therapie 1>** | **<Vergleichs-therapie 2>** | **<Vergleichs-therapie 3>** |
|---|---|---|---|---|---|
| 1 | <Studie 1> | ● | | ● | ● |
| 2 | <Studie 2> | ● | | ● | |
| | <Studie 3> | ● | | ● | |
| 1 | <Studie 4> | | ● | ● | ● |
| etc. | etc. | etc. | etc. | | |

*Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.*

Es werden keine indirekten Vergleiche durchgeführt.

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

Tabelle 4-174: 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 4-175: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

| 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.*

Es werden keine indirekten Vergleiche durchgeführt.

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

Tabelle 4-176: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Es werden keine indirekten Vergleiche durchgeführt.

*Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:*

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

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

Es werden keine indirekten Vergleiche durchgeführt.

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

#### 4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT.* ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.11.***

Es werden keine indirekten Vergleiche durchgeführt.

#### 4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

##### 4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien.* ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

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

Es werden keine nicht randomisierten vergleichenden Studien dargestellt.

##### 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 4-177: 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.*

Es werden keine nicht randomisierten vergleichenden Studien dargestellt.

#### 4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

##### 4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-178: 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 4-179: 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.*

Es werden keine nicht randomisierten vergleichenden Studien dargestellt.

*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.*

Es werden keine nicht randomisierten vergleichenden Studien dargestellt.

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 4.3.1.3.11.***

Es werden keine nicht randomisierten vergleichenden Studien dargestellt.

### 4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

#### 4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind.* ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

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

Es werden keine weiteren Untersuchungen dargestellt.

#### 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.***

Es werden keine weiteren Untersuchungen dargestellt.

#### 4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

##### 4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

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

Tabelle 4-180: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

| Studie | Operationalisierung |
|---|---|
| <Studie 1> | |
| | |

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

Es werden keine weiteren Untersuchungen dargestellt.

*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.*

Es werden keine weiteren Untersuchungen dargestellt.

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 4.3.1.3.11.***

Es werden keine weiteren Untersuchungen dargestellt.

### 4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

*Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.*

Es werden keine Ergebnisse aus weiteren Unterlagen dargestellt.

## 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.*

Für die Nutzenbewertung von Sarilumab im vorliegenden Anwendungsgebiet liegt die Phase-III-Studie MONARCH vor. Es handelt sich dabei um eine randomisierte, aktiv-kontrollierte, doppelblinde, multinationale und multizentrische Studie. Die Studie MONARCH entspricht der Evidenzstufe Ib und damit der höchsten Evidenzstufe für eine klinische Studie laut Verfahrensordnung des G-BA (5. Kapitel, §5 Abs. 6).

Für die Studie MONARCH kann von einem niedrigen Verzerrungspotential ausgegangen werden. Es handelt sich um eine doppelblinde RCT. Die Randomisierung erfolgte stratifiziert nach Region, so dass von einer grundsätzlichen Gleichheit der Behandlungsarme innerhalb der Studie ausgegangen werden kann. Dies zeigte sich auch bei Auswertung der Patientencharakteristika zu Beginn der Studie. Somit war die interne Validität der Studie gewährleistet. Sowohl die Patienten als auch das Studienpersonal waren in Bezug auf die Studienmedikation verblindet. Die Erhebung und Auswertung der Endpunkte erfolgte ebenfalls verblindet. Es lag eine ergebnisunabhängige Berichterstattung vor. Für alle Endpunkte kann von einem niedrigen Verzerrungspotential ausgegangen werden. Die dargestellten Endpunkte sind patientenrelevant.

Zusammenfassend kann von einer hohen qualitativen Ergebnissicherheit der vorliegenden Nachweise ausgegangen werden.

### 4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

*Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.*

*Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.*

*Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):*

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

*Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.*

Zur Bewertung des Zusatznutzens für Sarilumab wurde im Folgenden die relevante Studienpopulation der Studie MONARCH herangezogen, da für alle Endpunkte keine Effektmodifikation durch die Anzahl der csDMARD-Vortherapien bestand. Die Ergebnisse der relevanten Studienpopulation gelten somit gleichermaßen für die Patienten in den Teilanwendungsgebieten 2 und 3.

Sofern auf den jeweiligen Endpunkt anwendbar, wurde das Ausmaß des Zusatznutzens anhand der Konfidenzintervalle der dichotomen Effektmaße bestimmt. Dabei wurde das RR betrachtet, außer es war aufgrund von null Ereignissen in einem Arm nur das Peto-OR verfügbar.

Entsprechend der IQWiG-Methodik wurde bei schwerwiegenden (bzw. schweren) Symptomen und Nebenwirkungen ein Effekt als „erheblich“, „beträchtlich“ oder „gering“ eingestuft, falls die obere Grenze des 95 % KI einen Wert von 0,75, 0,90 oder 1,00 unterschritt. Bei Vorliegen von Effekten > 1 wurden diese Schwellenwerte mittels Kehrwertbildung auf $1,\overline{3}$, $1,\overline{1}$ bzw. 1,0 festgelegt, d.h. ein Effekt wurde dann als „erheblich“, „beträchtlich“ oder „gering“ eingestuft, falls die untere Grenze des 95 % KI einen Wert von $1,\overline{3}$, $1,\overline{1}$ bzw. 1,0 überschritt.

Bei den Patienten der Studie MONARCH handelte es sich um Patienten mit einer schweren rheumatoiden Arthritis. Alle Patienten hatten zu Beginn der Studie eine hohe Krankheitsaktivität, starke Schmerzen sowie eine deutlich eingeschränkte körperliche Funktionsfähigkeit. Aus diesem Grund wurden die untersuchten Morbiditätsendpunkte als schwerwiegende Symptome der Erkrankung eingeschätzt.

Bei nicht schwerwiegenden Symptomen und Nebenwirkungen wurde ein Effekt als „beträchtlich“ oder „gering“ eingestuft, falls die obere Grenze des 95 % KI einen Wert von 0,80 oder 0,90 unterschritt. Bei Vorliegen von Effekten > 1 wurden diese Schwellenwerte mittels Kehrwertbildung auf 1,25 bzw. $1,\overline{1}$ festgelegt, d.h. ein Effekt wurde dann als

„beträchtlich“ oder „gering“ eingestuft, falls die untere Grenze des 95 % KI einen Wert von 1,25 bzw. $1,\overline{1}$ überschritt.

Die unerwünschten Ereignisse, bei denen Unterschiede zwischen den Behandlungsgruppen auftraten, wurden aufgrund ihrer fehlenden klinischen Relevanz als nicht schwerwiegende Nebenwirkungen eingestuft.

Lagen relevante, stetige Effektmaße vor, wurde bei statistisch signifikanten Mittelwertdifferenzen eine standardisierte Mittelwertdifferenz in Form von Hedges’ g zur Bewertung der Relevanz des statistisch signifikanten Gruppenunterschiedes herangezogen. Nach einem Vorschlag von Cohen (1986) indiziert eine standardisierte Mittelwertdifferenz von 0,2 einen *kleinen* Effekt, 0,5 einen *mittleren* und 0,8 einen *großen* Effekt [94]. Um diese Effektstärken auf das Ausmaß des Zusatznutzens zu übertragen, wurde in Anlehnung an einen Vorschlag des IQWiG auf Grundlage des 95% KI zu Hedges’ g folgendes Klassifikationsschema angewendet [74]:

| **Das 95%-Konfidenzintervall zu Hedges’ g...** | **Bewertung** |
|---|---|
| schließt Werte im Bereich [-0,2 ; 0,2] nicht aus | kein relevanter Effekt |
| schließt Werte im Bereich [-0,2 ; 0,2] aus,<br>nicht aber Werte im Bereich [-0,5 ; 0,5] | relevanter Effekt mit kleiner Effektstärke |
| schließt Werte im Bereich [-0,5 ; 0,5] aus,<br>nicht aber Werte im Bereich [-0,8 ; 0,8] | relevanter Effekt mit mittlerer Effektstärke |
| schließt Werte im Bereich [-0,8 ; 0,8] aus | relevanter Effekt mit großer Effektstärke |

Die nachfolgende Tabelle 4-181 zeigt in einer Übersicht das Ausmaß des Zusatznutzens für die einzelnen Endpunkte sowie die zugrunde liegenden Ergebnisse aus der Studie MONARCH. Es wurden Endpunkte der Kategorien Morbidität, gesundheitsbezogene Lebensqualität und Sicherheit untersucht. Ergebnisse zum Endpunkt Mortalität liegen nicht vor.

Tabelle 4-181: Bewertung des Zusatznutzens von Sarilumab gegenüber Adalimumab als Monotherapien für Patienten mit rheumatoider Arthritis und einer MTX-Intoleranz

| **Endpunktkategorie Endpunkt Operationalisierung** | **Effektschätzer [95 % KI] p-Wert** | **Ausmaß des Zusatznutzens[a]** |
|---|---|---|
| ***Morbidität[b]*** | | |
| Remission | | |
| DAS28-ESR < 2,6 | RR: 4,39 [1,78; 10,85] **0,0014** | erheblich |
| CDAI ≤ 2,8 | RR: 3,37 [0,73; 15,45] 0,1182 | - |
| SDAI ≤ 3,3 | RR: 3,75 [0,83; 16,88] 0,0844 | - |
| ACR/EULAR-Remission | RR: 2,20 [0,43; 11,16] 0,3423 | - |
| Niedrige Krankheitsaktivität | | |
| DAS28-ESR < 3,2 | RR: 4,22 [2,10; 8,46] **< 0,0001** | erheblich |
| CDAI ≤ 10 | RR: 2,39 [1,44; 3,97] **0,0008** | erheblich |
| SDAI ≤ 11 | RR: 2,29 [1,40; 3,74] **0,0010** | erheblich |
| Therapieansprechen | | |
| gemäß EULAR-Kriterien | RR: 1,22 [1,02; 1,46] **0,0328** | gering |
| ACR50 | RR: 1,73 [1,14; 2,63] **0,0104** | beträchtlich |
| ACR70 | RR: 2,62 [1,17; 5,85] **0,0194** | beträchtlich |
| Körperlicher Funktionsstatus | | |
| HAQ-DI (Verbesserung um ≥ 0,22) | RR: 1,37 [1,05; 1,78] **0,0213** | gering |
| HAQ-DI (Verbesserung um ≥ 0,375) | RR: 1,49 [1,11; 2,01] **0,0090** | gering |
| Gelenksymptome | | |
| 28TJC | LSMD: -1,452 [-3,126; 0,222] 0,0887 | - |
| 68TJC | LSMD: -1,281 [-4,314; 1,752] 0,4056 | - |
| 28SJC | LSMD: -1,333 [-2,336; -0,329] **0,0095**<br>H´g: -0,431 [-0,755; -0,107] | - |
| 66SJC | LSMD: -1,513 [-3,080; 0,054] 0,0584 | - |
| Schmerz | | |
| Schmerz-VAS (Verbesserung um ≥ 20,4) | RR: 1,07 [0,83; 1,39] 0,5931 | - |

| **Endpunktkategorie Endpunkt Operationalisierung** | **Effektschätzer [95 % KI] p-Wert** | **Ausmaß des Zusatznutzens[a]** |
|---|---|---|
| Morgensteifigkeit | | |
| Morgensteifigkeit-VAS | LSMD: -10,967 [-18,841; -3,094] **0,0066**<br>H´g: -0,456 [-0,783; -0,129] | - |
| Fatigue | | |
| FACIT-Fatigue (Verbesserung um ≥ 4) | RR: 1,27 [0,99; 1,64] 0,0631 | - |
| Allgemeiner Gesundheitszustand | | |
| PtGA (Verbesserung um ≥ 18,4) | RR: 1,18 [0,90; 1,55] 0,2196 | - |
| EQ-5D-VAS | LSMD: 9,244 [2,683; 15,806] **0,0061**<br>H´g: 0,461 [0,134; 0,788] | - |
| **Endpunktkategorie Morbidität** | | **erheblicher Zusatznutzen** |
| ***Lebensqualität*** | | |
| SF-36 | | |
| SF-36-PCS (Verbesserung um ≥ 2,5) | RR: 1,23 [0,95; 1,59] 0,1138 | - |
| SF-36-PCS (Verbesserung um ≥ 5,1) | RR: 1,30 [0,97; 1,75] 0,0803 | - |
| SF-36-PCS (Verbesserung um ≥ 7,2) | RR: 1,38 [0,96; 1,97] 0,0788 | - |
| SF-36-MCS | LSMD: 2,437 [-0,806; 5,681] 0,1397 | - |
| RAID | | |
| RAID (Verbesserung um ≥ 3) | RR: 1,62 [1,10; 2,37] **0,0142** | gering |
| **Endpunktkategorie Lebensqualität** | | **geringer Zusatznutzen** |
| ***Sicherheit / Verträglichkeit*** | | |
| Gesamtrate UE | RR: 0,98 [0,79; 1,20] 0,8216 | - |
| Gesamtrate SUE | RR: 0,75 [0,21; 2,72] 0,6657 | - |
| UE mit Todesfolge | Peto-OR: 6,98 [0,14; 352,19] 0,3316 | - |
| UE, die zum dauerhaften Abbruch der Behandlung führen | RR: 2,20 [0,59; 8,26] 0,2421 | - |

| **Endpunktkategorie Endpunkt Operationalisierung** | **Effektschätzer [95 % KI] p-Wert** | **Ausmaß des Zusatznutzens[a]** |
|---|---|---|
| Infektionen | RR: 1,02 [0,63; 1,66] 0,9212 | - |
| Schwerwiegende Infektionen | Peto-OR: 7,06 [0,44; 113,92] 0,1685 | - |
| Leukopenien | RR: 6,60 [1,54; 28,3] **0,0112** | geringerer Nutzen: Ausmaß beträchtlich |
| Schwerwiegende Leukopenien | Peto-OR: 6,98 [0,14; 352,19] 0,3316 | - |
| Leberwerterhöhungen | RR: 1,89 [0,49; 7,32] 0,3589 | - |
| Schwerwiegende Leberwerterhöhungen | Peto-OR: 0,13 [0,0; 6,43] 0,3030 | - |
| Lipidwerterhöhungen[c] | Peto-OR: 0,12 [0,02; 0,89] **0,0376** | gering |
| Überempfindlichkeit | RR: 0,24 [0,03; 2,08] 0,1926 | - |
| Reaktionen der Injektionsstelle | RR: 1,65 [0,50; 5,45] 0,4107 | - |
| **Endpunktkategorie Sicherheit / Verträglichkeit** | | **Zusatznutzen nicht belegt** |

DAS28-ESR: *Disease Activity Score with 28 Joint Counts* mit Erythrozytensedimentationsrate; CDAI: *Clinical Disease Activity Index*; SDAI: *Simplified Disease Activity Index*; ACR: *American College of Rheumatology*; EULAR: *European League Against Rheumatism*; TJC: *Tender Joint Count*; SJC: *Swollen Joint Count*; VAS: visuelle Analogskala; FACIT: *Functional Assessment of Chronic Illness Therapy*; HAQ-DI: *Health Assessment Questionnaire – Disability Index*; PtGA: *Patients Global Assessment*; SF-36-PCS: *Physical Component Summary Score* des SF-36; SF-36-MCS: *Mental Component Summary Score* des SF-36; RAID: *Rheumatoid Arthritis Impact of Disease*; UE: unerwünschte Ereignisse; SUE: schwerwiegende unerwünschte Ereignisse

[a] Ein nicht belegter Zusatznutzen ist durch „-“ gekennzeichnet.

[b] Alle Morbiditätsendpunkte wurden als schwerwiegende Symptome der Erkrankung bewertet, da alle Patienten der Studie MONARCH zu Studienbeginn eine hohe Krankheitsaktivität sowie starke Schmerzen und eine deutlich eingeschränkte körperliche Funktionsfähigkeit aufwiesen.

[c] Die aufgetretenen Lipidwerterhöhungen wurden als nicht-schwerwiegende UEs eingestuft.

### *Endpunktkategorie Morbidität*

Die Behandlung mit Sarilumab führte im Vergleich mit der zweckmäßigen Vergleichstherapie Adalimumab zu einer deutlichen Verminderung der Krankheitslast und Abschwächung schwerwiegender Symptome.

Bei den Patienten der Studie MONARCH handelte es sich um Patienten mit einer schweren rheumatoiden Arthritis. Für diese Patienten stellt das Erreichen einer niedrigen Krankheitsaktivität das vorrangige Therapieziel dar, da ein Großteil dieser Patienten den Zustand der Remission nicht mehr erreichen kann. Alle Patienten hatten zu Beginn der Studie eine hohe Krankheitsaktivität, starke Schmerzen sowie eine deutlich eingeschränkte körperliche Funktionsfähigkeit. Aus diesem Grund sind die untersuchten Morbiditätsendpunkte als schwerwiegende Symptome der Erkrankung einzuschätzen.

Das Erreichen der Remission stellt prinzipiell das primäre Therapieziel für Patienten mit rheumatoider Arthritis dar. Die Wahrscheinlichkeit, eine Remission gemäß DAS28-ESR zu erreichen, war unter Sarilumab gegenüber Adalimumab statistisch signifikant erhöht. Das Ausmaß dieses Effekts entspricht einem erheblichen Zusatznutzen. Bei Anwendung der strengeren CDAI- und SDAI- sowie der booleschen ACR/EULAR-Kriterien für eine Remission zeigte sich ein numerischer Vorteil für Sarilumab gegenüber Adalimumab, die Behandlungsunterschiede erreichten jedoch keine Signifikanz.

Für diejenigen Patienten, die den Zustand der Remission nicht erreichen können, stellt auch gemäß aktuellster Empfehlungen das Erreichen einer niedrigen Krankheitsaktivität das primäre Therapieziel dar, insbesondere für Patienten mit einer lang andauernden, schweren RA, welche in der hier dargestellten Studie MONARCH untersucht wurden. Die Wahrscheinlichkeit, eine niedrige Krankheitsaktivität zu erreichen, war unter Sarilumab im Vergleich zu Adalimumab ebenfalls signifikant erhöht, und zwar sowohl bei Betrachtung des DAS28-ESR als auch des CDAI und SDAI. Das Ausmaß dieses Effektes für den Endpunkt „Niedrige Krankheitsaktivität“ entspricht in allen drei Fällen einem erheblichen Zusatznutzen.

Die Behandlung mit Sarilumab führte zu einem signifikant besseren Therapieansprechen als die Behandlung mit Adalimumab. Die Wahrscheinlichkeit für ein gutes bis moderates Ansprechen gemäß EULAR-Kriterien war unter Sarilumab im Vergleich zu Adalimumab signifikant erhöht. Das Ausmaß des Effektes entspricht einem geringen Zusatznutzen. Dabei zeigte sich vor allem ein deutlicher Vorteil von Sarilumab hinsichtlich des Anteils der Patienten mit gutem Ansprechen. Die Wahrscheinlichkeit, eine 50%ige oder 70%ige Verringerung der Krankheitsaktivität gemäß ACR-Kriterien zu erreichen, war unter Sarilumab im Vergleich zu Adalimumab ebenfalls signifikant erhöht. Das Ausmaß der Effekte entspricht jeweils einem beträchtlichen Zusatznutzen.

Die Behandlung mit Sarilumab führte des Weiteren zu einer signifikanten und klinisch relevanten Verbesserung des körperlichen Funktionsstatus gemessen mittels HAQ-DI. Das Ausmaß des Effektes für den Endpunkt „Körperlicher Funktionsstatus“ entspricht einem geringen Zusatznutzen.

Für alle weiteren Morbiditätsendpunkte – Gelenksymptome, Schmerzen, Morgensteifigkeit, Fatigue und den allgemeinen Gesundheitszustand – konnte kein Zusatznutzen belegt werden. Allerdings zeigten sich über alle Endpunkte hinweg konsistent numerische Vorteile zugunsten von Sarilumab oder gar statistisch signifikante Vorteile, für die jedoch die klinische Relevanz nicht belegt werden konnte.

Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bestätigen die Ergebnisse der 24-wöchigen doppelblinden Behandlungsphase hinsichtlich der Verbesserung der Morbidität.

In der Gesamtschau ergibt sich gemäß § 5 Abs. 7 AM-NutzenV für die Endpunktkategorie „Morbidität“ aufgrund der erreichten langfristigen Freiheit von schwerwiegenden Symptomen eine große Verbesserung des therapierelevanten Nutzens und damit ein **erheblicher Zusatznutzen** von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab.

***Endpunktkategorie Gesundheitsbezogene Lebensqualität***

Die Behandlung mit Sarilumab führte im Vergleich mit der zweckmäßigen Vergleichstherapie Adalimumab zu einer deutlichen Verbesserung der gesundheitsbezogenen Lebensqualität.

Die Auswertung des generischen Messinstruments SF-36 ergab tendenziell einen Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich der körperlichen Summenskala (PCS). Die Wahrscheinlichkeit, eine Verbesserung der SF-36-PCS um mindestens 2,5, 5,1 und 7,2 Punkte zu erreichen, war unter Sarilumab numerisch erhöht. Die Behandlungsunterschiede waren jedoch statistisch nicht signifikant. Hinsichtlich der psychischen Summenskala (MCS) des SF-36 zeigte sich kein relevanter Unterschied zwischen den Behandlungsarmen. Insgesamt kann für die Verbesserung der gesundheitsbezogenen Lebensqualität – gemessen mit dem generischen Fragebogen SF-36 – kein Zusatznutzen abgeleitet werden.

Die Auswertung des spezifischen Fragebogens RAID, der speziell für Patienten mit rheumatoider Arthritis entwickelt und validiert wurde, ergab jedoch eine statistisch signifikant stärkere Verbesserung der Lebensqualität der Patienten unter Sarilumab. Das Ausmaß dieses Effektes entspricht einem geringen Zusatznutzen.

Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bestätigen die Ergebnisse der 24-wöchigen doppelblinden Behandlungsphase hinsichtlich der Verbesserung der Lebensqualität.

In der Gesamtschau ergibt sich für die Endpunktkategorie „Gesundheitsbezogene Lebensqualität“ aufgrund der festgestellten relevanten Verbesserung der Lebensqualität ein **geringer Zusatznutzen** von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab.

***Endpunktkategorie Sicherheit/Verträglichkeit***

Zusammenfassend zeigte sich in der Studie MONARCH überwiegend ein vergleichbares Sicherheitsprofil von Sarilumab und der zweckmäßigen Vergleichstherapie Adalimumab.

In Bezug auf die Gesamtraten der unerwünschten Ereignisse, der schwerwiegenden unerwünschten Ereignisse sowie der unerwünschten Ereignisse, die zum dauerhaften Abbruch der Behandlung führen, ergaben sich keine statistisch signifikanten Vor- oder Nachteile.

Unter der Behandlung mit Sarilumab traten signifikant häufiger Leukopenien auf als unter Adalimumab, wobei es sich fast ausschließlich um Neutropenien handelte. Dieser Effekt würde formal einem geringeren Nutzen vom Ausmaß beträchtlich entsprechen. Bei Neutropenien handelt es sich jedoch um Veränderungen eines Laborparameters, die nicht per se symptomatisch und für den Patienten spürbar sind. Mit einer Ausnahme handelte es sich um nicht schwerwiegende Neutropenien.

Neutropenien stellen ein bekanntes Risiko für Infektionen dar. In der Studie MONARCH zeigte sich jedoch kein Unterschied zwischen den Behandlungsarmen bezüglich der Häufigkeit von Infektionen. Schwerwiegende Infektionen traten sehr selten auf und waren

nicht mit einer Neutropenie assoziiert. Somit stellt das erhöhte Neutropenie-Risiko unter Sarilumab keinen patientenrelevanten Nachteil dar und das Auftreten von Neutropenien ist in diesem Fall als Surrogatparameter ohne klinische Relevanz zu betrachten.

In der Studie MONARCH zeigte sich des Weiteren ein signifikanter Vorteil von Sarilumab gegenüber Adalimumab hinsichtlich des Auftretens von Lipidwerterhöhungen. Dabei handelte es sich um reine Laborwertveränderungen, die nicht als schwerwiegend eingestuft wurden. Das Ausmaß des Behandlungseffekts würde formal einem geringen Zusatznutzen entsprechen.

Für alle übrigen unerwünschten Ereignisse von besonderem Interesse wurden keine signifikanten Vor- oder Nachteile für Sarilumab beobachtet.

Die Auswertungen nach 48 Wochen Behandlung mit Sarilumab bestätigen die Ergebnisse der 24-wöchigen doppelblinden Behandlungsphase hinsichtlich der Sicherheit und Verträglichkeit.

In der Gesamtschau ergibt sich für die Endpunktkategorie „Sicherheit/Verträglichkeit“ aufgrund der fehlenden klinischen Relevanz der beobachteten Unterschiede **kein Zusatznutzen und kein geringerer Nutzen** für Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab.

***Übertragbarkeit auf den deutschen Versorgungskontext***

Die Studie MONARCH wurde weltweit durchgeführt. Die Studienpopulation bestand zu etwa 90 % aus weißen (kaukasischen) Studienteilnehmern. Je nach ausgewerteter Teilpopulation stammt ein Drittel bis die Hälfte der Patienten aus westlichen Ländern (siehe Tabelle 4-15), bei denen von einem ähnlichen Standard der Gesundheitsversorgung wie in Deutschland ausgegangen werden kann.

Die für die Nutzenbewertung relevante Studienpopulation umfasst Patienten mit einer schweren aktiven rheumatoiden Arthritis, die bereits mit mindestens einem csDMARD vorbehandelt wurden, prognostisch ungünstige Faktoren aufweisen und die eine Behandlung mit MTX nicht vertragen. Diese Patienten entsprechen dem Anwendungsgebiet für eine Sarilumab-Monotherapie [1]. Die Behandlung dieser relevanten Studienpopulation in MONARCH entspricht den Empfehlungen der für Deutschland relevanten europäischen Therapie-Leitlinie [3].

Zusammenfassend ist die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext gegeben.

***Zusammenfassung der Aussagen zum Zusatznutzen***

Bezüglich der Endpunktkategorie Morbidität konnte ein erheblicher Zusatznutzen von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab abgeleitet werden. Bezüglich der Lebensqualität zeigte sich ein geringer Zusatznutzen von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab. Die Daten zu Sicherheit und

Verträglichkeit belegen in der Gesamtschau keinen Zusatznutzen aber auch keinen geringeren Nutzen von Sarilumab gegenüber der zweckmäßigen Vergleichstherapie Adalimumab.

Die qualitative Ergebnissicherheit der vorliegenden Studie ist insgesamt als hoch einzuschätzen. Über alle Endpunkte hinweg zeigten sich konsistente Ergebnisse mit gleicher Effektrichtung. Somit kann in der Gesamtschau von einer hohen Aussagesicherheit und damit einem Hinweis auf einen Zusatznutzen ausgegangen werden.

In der Gesamtbetrachtung ergibt sich somit unter Berücksichtigung der Schwere und der chronischen Natur der Erkrankung für Sarilumab ein **Hinweis auf einen erheblichen Zusatznutzen**, da eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte große Verbesserung des therapierelevanten Nutzens erreicht wird, insbesondere eine langfristige Freiheit von schwerwiegenden Symptomen sowie eine relevante Verbesserung der gesundheitsbezogenen Lebensqualität.

### 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 4-182: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

| Bezeichnung der Patientengruppen | | Ausmaß des Zusatznutzens |
|---|---|---|
| 1 | Patienten, die mit einem csDMARD vorbehandelt sind und keine ungünstigen Prognosefaktoren aufweisen[a] | Zusatznutzen nicht belegt[b] |
| 2a | Für MTX geeignete Patienten, die mit einem csDMARD vorbehandelt sind und ungünstige Prognosefaktoren aufweisen | Zusatznutzen nicht belegt[b] |
| 2b | Patienten mit MTX-Intoleranz, die mit einem csDMARD vorbehandelt sind und ungünstige Prognosefaktoren aufweisen | **Hinweis auf einen erheblichen Zusatznutzen** |
| 3a | Für MTX geeignete Patienten, die mit mehreren csDMARDs vorbehandelt sind | Zusatznutzen nicht belegt[b] |
| 3b | Patienten mit MTX-Intoleranz, die mit mehreren csDMARDs vorbehandelt sind | **Hinweis auf einen erheblichen Zusatznutzen** |
| 4 | Patienten, die mit einem oder mehreren bDMARDs vorbehandelt sind | Zusatznutzen nicht belegt[b] |

[a] Anmerkung bezüglich Teilanwendungsgebiet 1: Es handelt sich dabei um Patienten, welche zwar formal von der Zulassung für Sarilumab umfasst sind, die aber im klinischen Alltag noch nicht mit einem Biologikum behandelt werden würden. Die EULAR-Leitlinie sieht für diese Patienten eine Therapie mit einem weiteren csDMARD vor [3]. Aus diesem Grund ist davon auszugehen, dass die Patienten in Teilanwendungsgebiet 1 im Regelfall nicht mit Sarilumab behandelt werden. Das Teilanwendungsgebiet 1 wird somit für die Nutzenbewertung von Sarilumab als nicht relevant erachtet.

[b] Der Zusatznutzen konnte nicht belegt werden, weil keine geeigneten Daten zur Verfügung standen.

## 4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

### 4.5.1 Begründung für die Vorlage indirekter Vergleiche

*Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.*

Es wurden keine indirekten Vergleiche durchgeführt.

### 4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

*Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.*

Es wurden keine nicht randomisierten vergleichenden Studien oder weitere Untersuchungen vorgelegt.

### 4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

*Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.*

Nicht zutreffend, es liegen valide Daten zu patientenrelevanten Endpunkten vor.

### 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[28], Molenberghs 2010[29]). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006[30]) 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[31]) soll

[28] *Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.*
[29] *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.*
[30] *Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.*
[31] *Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.*

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 Liste der eingeschlossenen Studien

*Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).*

| Studie | Synonym | Studienberichte | Registereinträge | Publikationen |
|---|---|---|---|---|
| MONARCH | EFC14092 | [71] | [82-86] | [87] |

### 4.7 Referenzliste

*Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.*

1. Sanofi-Aventis Deutschland GmbH (2017): Kevzara® 200 mg / 150 mg Injektionslösung in einem Fertigpen / Fertigspritze; Fachinformation. Stand: Juni 2017; Fachinformation. Stand: Juni 2017 [Zugriff: 30.06.2017]. URL: http://www.fachinfo.de.
2. Gemeinsamer Bundesausschuss (2016): Niederschrift zum Beratungsgespräch gemäß § 8 Absatz 1 AM-NutzenV, Beratungsanforderung 2016-B-101.
3. Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, et al. (2017): EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Annals of the rheumatic diseases; 76:960–77.
4. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2017): Biotechnologisch hergestellte Wirkstoffe bei rheumatoider Arthritis; Berichtsplan. [Zugriff: 31.07.2017]. URL: https://www.iqwig.de/download/A16-70_Biotechnologisch-hergestellte-Wirkstoffe-bei-rheumatoider-Arthritis_Berichtsplan_V1-0.pdf.
5. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. (2014): EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Annals of the rheumatic diseases; 73(3):492-509.
6. ICH Expert Working Group (1994): ICH HARMONISED TRIPARTITE GUIDELINE: The Extend of Population Exposure to Assess Clinical Safety for Drugs Intended for Long-Term Treatment of Non-Life-Threatening Conditions; E1, current step 4 version. [Zugriff: 03.03.2017]. URL: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E1/Step4/E1_Guideline.pdf.
7. Gemeinsamer Bundesausschuss (2015): 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 - Secukinumab. [Zugriff: 03.03.2017]. URL: https://www.g-ba.de/downloads/40-268-3450/2015-11-27_AM-RL-XII_Secukinumab_2015-06-01-D167_TrG.pdf.
8. Novartis Pharma GmbH (2015): Dossier zur Nutzenbewertung gemäß § 35a SGB V; Secukinumab (Cosentyx®); Modul 4C. [Zugriff: 17.05.2017]. URL: https://www.g-ba.de/downloads/92-975-1273/2015-12-11_Modul4C_Secukinumab.pdf.
9. Novartis Pharma GmbH (2015): Dossier zur Nutzenbewertung gemäß § 35a SGB V; Secukinumab (Cosentyx®); Modul 4B. [Zugriff: 17.05.2017]. URL: https://www.g-ba.de/downloads/92-975-1272/2015-12-11_Modul4B_Secukinumab.pdf.
10. Bayer Vital GmbH (1999): Resochin® Tabletten 250 mg; Fachinformation. Stand: Juli 2016 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
11. medac (2012): metex® PEN; Fachinformation. Stand: November 2016 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
12. Novartis Pharma GmbH (1997): Sandimmun® Optoral Weichkapseln; Fachinformation. Stand: Juli 2015 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
13. Sanofi-Aventis Deutschland GmbH (1999): Arava® 10 mg/20 mg/100 mg Filmtabletten; Fachinformation. Stand: Dezember 2015 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.

14. MSD SHARP & DOHME GMBH (2009): Simponi® 100 mg Injektionslösung in einer Fertigspritze; Fachinformation. Stand: Februar 2017 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
15. Pfizer Pharma GmbH (2000): Enbrel® 50 mg Injektionslösung im Fertigpen; Fachinformation. Stand: März 2017 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
16. UCB Pharma GmbH (2009): Cimzia® 200 mg Injektionslösung; Fachinformation. Stand: Januar 2017 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
17. Roche Pharma AG (2009): RoActemra® i. v.; Fachinformation. Stand: Juli 2016 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
18. AbbVie Ltd (2003): Humira® 40 mg/0,8 ml Injektionslösung im Fertigpen; Fachinformation. Stand: Dezember 2016 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
19. Roche Pharma AG (1998): MabThera® i.v.; Fachinformation. Stand: September 2016 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
20. Bristol-Myers Squibb GmbH (2007): ORENCIA ® 125 mg Injektionslösung im Fertigpen; Fachinformation. Stand: Mai 2017 [Zugriff: 22.06.2017]. URL: http://www.fachinfo.de.
21. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. (2010): CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol; 63(8):e1-37.
22. Schulz KF, Altman DG, Moher D, Consort Group (2010): CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. PLoS Med; 7(3):e1000251.
23. Fuchs HA, Brooks RH, Callahan LF, Pincus T (1989): A simplified twenty-eight-joint quantitative articular index in rheumatoid arthritis. Arthritis Rheum; 32(5):531-7.
24. Gemeinsamer Bundesausschuss (2016): Niederschrift zum Beratungsgespräch gemäß § 8 Absatz 1 AM-NutzenV, Beratungsanforderung 2015-B-178.
25. Ward MM, Guthrie LC, Alba MI (2015): Clinically important changes in individual and composite measures of rheumatoid arthritis activity: thresholds applicable in clinical trials. Annals of the rheumatic diseases; 74(9):1691-6.
26. Maska L, Anderson J, Michaud K (2011): Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire Disability Index (HAQ), Modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire II (HAQ-II), Improved Health Assessment Questionnaire (Improved HAQ), and Rheumatoid Arthritis Quality of Life (RAQoL). Arthritis Care Res (Hoboken); 63 Suppl 11:S4-13.
27. Wolfe F, Michaud K, Gefeller O, Choi HK (2003): Predicting mortality in patients with rheumatoid arthritis. Arthritis Rheum; 48(6):1530-42.
28. Redelmeier DA, Lorig K (1993): Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. Archives of internal medicine; 153(11):1337-42.
29. Wells GA, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA (1993): Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. The Journal of rheumatology; 20(3):557-60.
30. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2013): Biotechnologisch hergestellte Arzneimittel in der Zweitlinientherapie bei der rheumatoiden Arthritis; Abschlussbericht. [Zugriff: 21.04.2017]. URL:

https://www.iqwig.de/download/A10-01_Abschlussbericht_Biologika-Zweitlinientherapie-bei-rheumatoider-Arthritis.pdf.
31. Aletaha D, Landewe R, Karonitsch T, Bathon J, Boers M, Bombardier C, et al. (2008): Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Arthritis Rheum; 59(10):1371-7.
32. van Riel PL (2014): The development of the disease activity score (DAS) and the disease activity score using 28 joint counts (DAS28). Clinical and experimental rheumatology; 32(5 Suppl 85):S-65-74.
33. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL (1995): Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum; 38(1):44-8.
34. Anderson JK, Zimmerman L, Caplan L, Michaud K (2011): Measures of rheumatoid arthritis disease activity: Patient (PtGA) and Provider (PrGA) Global Assessment of Disease Activity, Disease Activity Score (DAS) and Disease Activity Score with 28-Joint Counts (DAS28), Simplified Disease Activity Index (SDAI), Clinical Disease Activity Index (CDAI), Patient Activity Score (PAS) and Patient Activity Score-II (PASII), Routine Assessment of Patient Index Data (RAPID), Rheumatoid Arthritis Disease Activity Index (RADAI) and Rheumatoid Arthritis Disease Activity Index-5 (RADAI-5), Chronic Arthritis Systemic Index (CASI), Patient-Based Disease Activity Score With ESR (PDAS1) and Patient-Based Disease Activity Score without ESR (PDAS2), and Mean Overall Index for Rheumatoid Arthritis (MOI-RA). Arthritis Care Res (Hoboken); 63 Suppl 11:S14-36.
35. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL (1996): Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. Arthritis Rheum; 39(1):34-40.
36. van Gestel AM, Haagsma CJ, van Riel PL (1998): Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum; 41(10):1845-50.
37. Siemons L, Vonkeman HE, ten Klooster PM, van Riel PL, van de Laar MA (2014): Interchangeability of 28-joint disease activity scores using the erythrocyte sedimentation rate or the C-reactive protein as inflammatory marker. Clin Rheumatol; 33(6):783-9.
38. Aletaha D, Smolen J (2005): The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clinical and experimental rheumatology; 23(5 Suppl 39):S100-8.
39. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. (2005): Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther; 7(4):R796-806.
40. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. (2003): A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology (Oxford); 42(2):244-57.
41. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. (2011): American College of Rheumatology/European League Against Rheumatism

provisional definition of remission in rheumatoid arthritis for clinical trials. Arthritis Rheum; 63(3):573-86.
42. Fransen J, Creemers MC, Van Riel PL (2004): Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. Rheumatology (Oxford); 43(10):1252-5.
43. Fransen J, van Riel PL (2006): DAS remission cut points. Clinical and experimental rheumatology; 24(6 Suppl 43):S-29-32.
44. Aletaha D, Ward MM, Machold KP, Nell VP, Stamm T, Smolen JS (2005): Remission and active disease in rheumatoid arthritis: defining criteria for disease activity states. Arthritis Rheum; 52(9):2625-36.
45. Zhang B, Combe B, Rincheval N, Felson DT (2012): Validation of ACR/EULAR definition of remission in rheumatoid arthritis from RA practice: the ESPOIR cohort. Arthritis Res Ther; 14(3):R156.
46. Bykerk VP, Massarotti EM (2012): The new ACR/EULAR remission criteria: rationale for developing new criteria for remission. Rheumatology (Oxford); 51 Suppl 6:vi16-20.
47. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. (1995): American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum; 38(6):727-35.
48. Felson D (2007): A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. Arthritis Rheum; 57(2):193-202.
49. Ranganath VK, Khanna D, Paulus HE (2006): ACR remission criteria and response criteria. Clinical and experimental rheumatology; 24(6 Suppl 43):S-14-21.
50. European Medicines Agency (2015): Guideline on clinical investigation of medicinal products other than NSAIDs for treatment of rheumatoid arthritis, Draft. [Zugriff: 14.12.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/06/WC500187583.pdf.
51. da Silva JA, Phillips S, Buttgereit F (2011): Impact of impaired morning function on the lives and well-being of patients with rheumatoid arthritis. Scandinavian journal of rheumatology Supplement; 125:6-11.
52. Wolfe F, Hawley DJ, Wilson K (1996): The prevalence and meaning of fatigue in rheumatic disease. The Journal of rheumatology; 23(8):1407-17.
53. Hewlett S, Chalder T, Choy E, Cramp F, Davis B, Dures E, et al. (2011): Fatigue in rheumatoid arthritis: time for a conceptual model. Rheumatology (Oxford); 50(6):1004-6.
54. Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D, et al. (2005): Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. Arthritis Rheum; 53(5):697-702.
55. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J (2005): Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. The Journal of rheumatology; 32(5):811-9.
56. EuroQol Group (1990): EuroQol - a new facility for the measurement of health-related quality of life. Health policy (Amsterdam, Netherlands); 16(3):199-208.
57. Hurst NP, Jobanputra P, Hunter M, Lambert M, Lochhead A, Brown H (1994): Validity of Euroqol - a generic health status instrument - in patients with rheumatoid

arthritis. Economic and Health Outcomes Research Group. British journal of rheumatology; 33(7):655-62.
58. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A (1997): Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). British journal of rheumatology; 36(5):551-9.
59. McHorney CA, Ware JE, Jr., Raczek AE (1993): The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Medical care; 31(3):247-63.
60. Ware JE, Jr., Sherbourne CD (1992): The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical care; 30(6):473-83.
61. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE, Jr. (2000): Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum; 43(7):1478-87.
62. Ward MM, Guthrie LC, Alba MI (2014): Clinically important changes in short form 36 health survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. Arthritis Care Res (Hoboken); 66(12):1783-9.
63. Strand V, Reaney M, Chen CI, Proudfoot CW, Guillonneau S, Bauer D, et al. (2017): Sarilumab improves patient-reported outcomes in rheumatoid arthritis patients with inadequate response/intolerance to tumour necrosis factor inhibitors. RMD open; 3(1):e000416.
64. Strand V, Lee EB, Fleischmann R, Alten RE, Koncz T, Zwillich SH, et al. (2016): Tofacitinib versus methotrexate in rheumatoid arthritis: patient-reported outcomes from the randomised phase III ORAL Start trial. RMD open; 2(2):e000308.
65. Strand V, Kosinski M, Chen CI, Joseph G, Rendas-Baum R, Graham NM, et al. (2016): Sarilumab plus methotrexate improves patient-reported outcomes in patients with active rheumatoid arthritis and inadequate responses to methotrexate: results of a phase III trial. Arthritis Res Ther; 18:198.
66. Strand V, Rentz AM, Cifaldi MA, Chen N, Roy S, Revicki D (2012): Health-related quality of life outcomes of adalimumab for patients with early rheumatoid arthritis: results from a randomized multicenter study. The Journal of rheumatology; 39(1):63-72.
67. Keystone E, Burmester GR, Furie R, Loveless JE, Emery P, Kremer J, et al. (2008): Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. Arthritis Rheum; 59(6):785-93.
68. Genovese MC, Schiff M, Luggen M, Becker JC, Aranda R, Teng J, et al. (2008): Efficacy and safety of the selective co-stimulation modulator abatacept following 2 years of treatment in patients with rheumatoid arthritis and an inadequate response to anti-tumour necrosis factor therapy. Annals of the rheumatic diseases; 67(4):547-54.
69. Gossec L, Paternotte S, Aanerud GJ, Balanescu A, Boumpas DT, Carmona L, et al. (2011): Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient-derived composite measure of impact of rheumatoid arthritis: a EULAR initiative. Annals of the rheumatic diseases; 70(6):935-42.
70. Dougados M, Brault Y, Logeart I, van der Heijde D, Gossec L, Kvien T (2012): Defining cut-off values for disease activity states and improvement scores for patient-reported outcomes: the example of the Rheumatoid Arthritis Impact of Disease (RAID). Arthritis Res Ther; 14(3):R129.

71. Burmester G (2016): Clinical Study Report: A randomized, double-blind, parallel-group study assessing the efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy in patients with rheumatoid arthritis; STUDY NUMBER: EFC14092; STUDY NAME: SARIL-RA-MONARCH.
72. Sanofi-Aventis Deutschland GmbH (2017): Studie MONARCH: Post hoc Analysen 24 Wochen; Efficacy-Endpunkte.
73. Sanofi-Aventis Deutschland GmbH (2017): Studie MONARCH: Post hoc Analysen 24 Wochen; Safety-Endpunkte.
74. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2016): Allgemeine Methoden, Version 5.0 vom 10.07.2017. [Zugriff: 31.07.2017]. URL: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf.
75. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. (2010): 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Annals of the rheumatic diseases; 69(9):1580-8.
76. M. Schneider, M. Lelgemann, H.-H. Abholz, M. Blumenroth, C. Flügge, M. Gerken, et al. (2011): Management der frühen rheumatoiden Arthritis. [Zugriff: 01.06.2017]. URL: http://www.awmf.org/uploads/tx_szleitlinien/060-002l_S3_Management_fruehe_rheumatoide_Arthritis_2011-abgelaufen.pdf.
77. Minichiello E, Semerano L, Boissier MC (2016): Time trends in the incidence, prevalence, and severity of rheumatoid arthritis: A systematic literature review. Joint Bone Spine; 83(6):625-30.
78. Diffin JG, Lunt M, Marshall T, Chipping JR, Symmons DPM, Verstappen SMM (2014): Has the Severity of Rheumatoid Arthritis at Presentation Diminished Over Time? The Journal of rheumatology; 41(8):1590-9.
79. Haugeberg G, Hansen IJW, Soldal DM, Sokka T (2015): Ten years of change in clinical disease status and treatment in rheumatoid arthritis: results based on standardized monitoring of patients in an ordinary outpatient clinic in southern Norway. Arthritis Research & Therapy; 17(1):219.
80. Louie GH, Ward MM (2010): Changes in the rates of joint surgery among patients with rheumatoid arthritis in California, 1983–2007. Annals of the rheumatic diseases; 69(5):868-71.
81. Nikiphorou E, Carpenter L, Morris S, Macgregor AJ, Dixey J, Kiely P, et al. (2014): Hand and foot surgery rates in rheumatoid arthritis have declined from 1986 to 2011, but large-joint replacement rates remain unchanged: results from two UK inception cohorts. Arthritis Rheumatol; 66(5):1081-9.
82. Sanofi, Regeneron Pharmaceuticals (2015): EFC14092, 2014-002541-22, U1111-1160-6154 - Efficacy and Safety of Sarilumab and Adalimumab Monotherapy in Patients With Rheumatoid Arthritis (SARIL-RA-MONARCH) (NCT02332590). Stand des Eintrags: 07.12.2016. [Zugriff: 18.05.2017]. URL: https://ClinicalTrials.gov/show/NCT02332590
83. Sanofi-Aventis (2014): EFC14092 - A randomized, double-blind, parallel-group study assessing the efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy in patients with rheumatoid arthritis (MONARCH) (2014-002541-22). [Zugriff: 18.05.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-002541-22
84. Sanofi-Aventis (2014): EFC14092;2014-002541-22-DE - A randomized, double-blind, parallel-group study assessing the efficacy and safety of sarilumab monotherapy

versus adalimumab monotherapy in patients with rheumatoid arthritis - SARIL-RA-MONARCH (EUCTR2014-002541-22-CZ). Stand des Eintrags: 09.01.2017. [Zugriff: 18.05.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-002541-22-CZ

85. Sanofi Aventis (2015): EFC14092 - A Randomized, Double-Blind, Parallel-Group Study Assessing the Efficacy and Safety of Sarilumab Monotherapy Versus Adalimumab Monotherapy in Patients with Rheumatoid. (MONARCH) (PER-081-14). Stand des Eintrags: 19. Juni 2017. [Zugriff: 29.06.2017]. URL: http://www.ins.gob.pe/ensayosclinicos/rpec/recuperarECPBNuevoEN.asp?numec=081-14
86. Sanofi-Aventis (2014): EFC14092 - A randomized, double-blind, parallel-group study assessing the efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy in patients with rheumatoid arthritis (2014-002541-22). [Zugriff: 18.05.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html
87. Burmester GR, Lin Y, Patel R, van Adelsberg J, Mangan EK, Graham NM, et al. (2017): Efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy for the treatment of patients with active rheumatoid arthritis (MONARCH): a randomised, double-blind, parallel-group phase III trial. Annals of the rheumatic diseases; 76(5):840-7.
88. Sanofi-Aventis Deutschland GmbH (2017): Studie MONARCH: Post hoc Analysen; Patientencharakteristika.
89. European Medicines Agency (2017): Kevzara: EPAR - Public assessment report. [Zugriff: 29.06.2017]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004254/WC500230070.pdf.
90. Sanofi-Aventis Deutschland GmbH (2017): Studie MONARCH: Post hoc Analysen 48 Wochen; Efficacy-Endpunkte.
91. European Medicines Agency (2016): Guideline on the clinical development of medicinal products intended for the treatment of pain (EMA/CHMP/970057/2011). [Zugriff: 02.03.2017]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/12/WC500219131.pdf.
92. Sanofi-Aventis Deutschland GmbH (2017): Studie MONARCH: Post hoc Analysen 48 Wochen; Safety-Endpunkte.
93. Roche Pharma AG (2009): RoActemra® s.c.; Fachinformation. Stand: März 2017 [Zugriff: 23.06.2017]. URL: http://www.fachinfo.de.
94. Cohen J (1988): The t Test for Means. Chapter 2. In: Statistical power analysis for the behavioral sciences (2nd ed) Hillsdale, NJ, Lawrence Erlbaum Associates. S. 19-74.
95. Wong SS, Wilczynski NL, Haynes RB (2006): Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. Journal of the Medical Library Association : JMLA; 94(4):451-5.

## Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

*Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50") und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):*

| | | |
|---|---|---|
| **Datenbankname** | EMBASE | |
| **Suchoberfläche** | Ovid | |
| **Datum der Suche** | 08.12.2010 | |
| **Zeitsegment** | 1980 to 2010 week 50 | |
| **Suchfilter** | Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[32]] – Strategy minimizing difference between sensitivity and specificity | |
| **#** | **Suchbegriffe** | **Ergebnis** |
| 1 | Meglitinide/ | 848 |
| 2 | Nateglinide/ | 1686 |
| 3 | Repaglinide/ | 2118 |
| 4 | (glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti. | 1069 |
| 5 | (starlix or novonorm or novo norm or prandin).ab,ti. | 32 |
| 6 | (105816-04-4 or 135062-02-1).rn. | 2854 |
| 7 | or/1-6 | 3467 |
| 8 | Diabetes mellitus/ | 224164 |
| 9 | Non Insulin dependent Diabetes mellitus/ | 91081 |
| 10 | (diabet* or niddm or t2dm).ab,ti. | 379777 |
| 11 | or/8-10 | 454517 |
| 12 | (random* or double-blind*).tw. | 650136 |
| 13 | placebo*.mp. | 243550 |
| 14 | or/12-13 | 773621 |
| 15 | and/7,11,14 | 719 |

[32] 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

Tabelle 4-183: Bibliografische Literaturrecherche in EMBASE

| | | |
|---|---|---|
| **Datenbankname** | EMBASE | |
| **Suchoberfläche** | Ovid | |
| **Datum der Suche** | 30.06.2017 | |
| **Zeitsegment** | 1974 to 2017 June 29 | |
| **Suchfilter** | Filter für randomisierte kontrollierte Studien nach Wong [95] | |
| **#** | **Suchbegriffe** | **Ergebnis** |
| 1 | random*.tw. | 1.212.380 |
| 2 | placebo*.mp. | 399.947 |
| 3 | double-blind*.tw. | 180.186 |
| 4 | 1 OR 2 OR 3 | 1.452.443 |
| 5 | exp sarilumab/ | 184 |
| 6 | sarilumab*.mp. | 188 |
| 7 | SAR-153191.mp. | 1 |
| 8 | SAR153191.mp. | 12 |
| 9 | SAR 153191.mp. | 12 |
| 10 | REGN-88.mp. | 20 |
| 11 | REGN88.mp. | 4 |
| 12 | REGN 88.mp. | 20 |
| 13 | kevzara*.mp. | 1 |
| 14 | 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 | 195 |
| 15 | 4 AND 14 | 102 |

Tabelle 4-184: Bibliografische Literaturrecherche in MEDLINE

| | | |
|---|---|---|
| **Datenbankname** | MEDLINE | |
| **Suchoberfläche** | Ovid MEDLINE(R) 1946 to present; Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations; Ovid MEDLINE(R) Daily Update; Ovid MEDLINE(R) Epub Ahead of Print | |
| **Datum der Suche** | 30.06.2017 | |
| **Zeitsegment** | 1946 to 2017 June 29 | |
| **Suchfilter** | Filter für randomisierte kontrollierte Studien nach Wong [95] | |
| **#** | **Suchbegriffe** | **Ergebnis** |
| 1 | randomized controlled trial.pt. | 467.493 |
| 2 | randomized.mp. | 738.646 |
| 3 | placebo.mp. | 196.881 |
| 4 | 1 OR 2 OR 3 | 796.938 |
| 5 | sarilumab*.mp. | 33 |
| 6 | SAR153191.mp. | 0 |
| 7 | SAR-153191.mp. | 0 |
| 8 | SAR 153191.mp. | 0 |
| 9 | REGN88.mp. | 2 |
| 10 | REGN-88.mp. | 0 |
| 11 | REGN 88.mp. | 0 |
| 12 | kevzara*.mp. | 0 |
| 13 | 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 | 35 |
| 14 | 4 AND 13 | 12 |

Tabelle 4-185: Bibliografische Literaturrecherche in der Cochrane Library

| | | |
|---|---|---|
| **Datenbankname** | Cochrane Database | |
| **Suchoberfläche** | The Cochrane Library | |
| **Datum der Suche** | 30.06.2017 | |
| **Zeitsegment** | uneingeschränkt | |
| **Suchfilter** | keine Einschränkung | |
| **#** | **Suchbegriffe** | **Ergebnis** |
| 1 | sarilumab* | 50 |
| 2 | SAR153191 | 2 |
| 3 | SAR 153191 | 0 |
| 4 | SAR-153191 | 0 |
| 5 | REGN88 | 2 |
| 6 | REGN 88 | 0 |
| 7 | REGN-88 | 0 |
| 8 | Kevzara* | 0 |
| 9 | 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 | 50 |
| 10 | Einschränkung: Trials | 46 |

Die bibliografische Literaturrecherche ergab insgesamt 160 Treffer, davon waren 35 Dubletten. Die verbleibenden 125 Treffer wurden anhand von Titel und Abstract gescreent.

### Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Es wurde keine Suche nach RCT für indirekte Vergleiche durchgeführt.

### Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Es wurde keine Suche nach nicht randomisierten vergleichenden Studien durchgeführt.

### Anhang 4-A4: Suche nach weiteren Untersuchungen

Es wurde keine Suche nach weiteren Untersuchungen durchgeführt.

## Anhang 4-B: Suchstrategien – Suche in Studienregistern

*Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. clinicaltrials.gov), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. http://www.clinicaltrials.gov), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:*

| | |
|---|---|
| **Studienregister** | clinicaltrials.gov |
| **Internetadresse** | http://www.clinicaltrials.gov |
| **Datum der Suche** | 08.12.2010 |
| **Suchstrategie** | (Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE] |
| **Treffer** | 23 |

## Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-186: Suchstrategie der Recherche in Studienregistern

| | |
|---|---|
| **Studienregister** | clinicaltrials.gov |
| **Internetadresse** | http://www.clinicaltrials.gov |
| **Datum der Suche** | 26.06.2017 |
| **Suchstrategie** | sarilumab OR SAR153191 OR REGN88 OR kevzara [a] |
| **Treffer** | **24** |
| | |
| **Studienregister** | EU Clinical Trials Register |
| **Internetadresse** | https://www.clinicaltrialsregister.eu/ctr-search/search |
| **Datum der Suche** | 30.06.2017 |
| **Suchstrategie** | sarilumab OR SAR153191 OR REGN88 OR kevzara [a] |
| **Treffer** | **14** |
| | |
| **Studienregister** | International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal) |
| **Internetadresse** | http://apps.who.int/trialsearch/ |
| **Datum der Suche** | 30.06.2017 |
| **Suchstrategie** | sarilumab* OR SAR153191 OR REGN88 OR kevzara* [a] |
| **Treffer** | **32** |
| | |
| **Studienregister** | Klinische Prüfungen PharmNet.Bund |
| **Internetadresse** | http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html |
| **Datum der Suche** | 30.06.2017 |
| **Suchstrategie** | sarilumab* [active substance] OR SAR153191 [active substance] OR REGN88 [active substance] OR sarilumab* [Textfelder] OR SAR153191 [Textfelder] OR REGN88 [Textfelder] OR kevzara* [active substance] OR kevzara* [Textfelder] [a] |
| **Treffer** | **9** |
| | |
| [a] Die Suche nach Variationen von SAR153191 (wie SAR 153191 etc.) oder REGN88 führte nicht zu weiteren Treffern. | |

Die Suche in clinicaltrials.gov ergab 24 Treffer. Darunter waren 19 Studien im Anwendungsgebiet der rheumatoiden Arthritis, wobei es sich bei allen Studien um Studien des pharmazeutischen Unternehmers handelt. 18 Studien wurden ausgeschlossen (Ausschlussgründe siehe Tabelle 4-187). Als einzig relevante Studie wurde die Studie MONARCH eingeschlossen.

Die Suche in EU-CTR ergab 14 Treffer. Es handelt sich dabei ausnahmslos um Studien, die bereits durch die Suche in clinicaltrials.gov identifiziert wurden. Davon war 1 Treffer die relevante Studie MONARCH, alle übrigen Treffer wurden ausgeschlossen (Ausschlussgründe siehe Tabelle 4-188).

Die Suche in ICTRP ergab 32 Treffer. Es handelt sich dabei ausnahmslos um Studien, die bereits durch die Suche in clinicaltrials.gov identifiziert wurden. Die höhere Trefferzahl im Vergleich zu clinicaltrials.gov ergibt sich durch Studien, die mit mehreren Einträgen vertreten sind. Zwei Treffer bezogen sich auf die relevante Studie MONARCH, alle übrigen Treffer wurden ausgeschlossen (Ausschlussgründe siehe Tabelle 4-189).

Die Suche in PharmNet.Bund ergab 9 Treffer. Es handelt sich dabei ausnahmslos um Studien, die bereits durch die Suche in clinicaltrials.gov identifiziert wurden. Davon war 1 Treffer die relevante Studie MONARCH, alle übrigen Treffer wurden ausgeschlossen (Ausschlussgründe siehe Tabelle 4-190).

**Anhang 4-B2: Suche nach RCT für indirekte Vergleiche**

Es wurde keine Suche nach RCT für indirekte Vergleiche durchgeführt.

**Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien**

Es wurde keine Suche nach nicht randomisierten vergleichenden Studien durchgeführt.

**Anhang 4-B4: Suche nach weiteren Untersuchungen**

Es wurde keine Suche nach weiteren Untersuchungen durchgeführt.

## 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

Es wurden keine im Volltext gesichteten Dokumente ausgeschlossen.

### Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Es wurde keine Suche nach RCT für indirekte Vergleiche durchgeführt.

### Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Es wurde keine Suche nach nicht randomisierten vergleichenden Studien durchgeführt.

### Anhang 4-C4: Suche nach weiteren Untersuchungen

Es wurde keine Suche nach weiteren Untersuchungen durchgeführt.

## Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

*Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.*

### Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-187: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – ClinicalTrials.gov

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| 01 | LTS11298 / NCT01118728<br>Sanofi, Regeneron Pharmaceuticals (2010): LTS11298 / 2010-019263-11 - Extension Study for Long Term Evaluation of SAR153191 (REGN88) in Patients With Ankylosing Spondylitis (NCT01118728). Stand des Eintrags: Mai 2017. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01118728 | A1: falsche Studienpopulation |
| 02 | MOBILITY / EFC11072 / NCT01061736<br>Sanofi, Regeneron Pharmaceuticals (2010): EFC11072 / 2009-016266-90 - Evaluation of Sarilumab (SAR153191/REGN88) on Top of Methotrexate in Rheumatoid Arthritis Patients (NCT01061736). Stand des Eintrags: November 2014. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01061736 | A3: Komparator entspricht nicht der zVT |
| 03 | EASY / MSC12665 / NCT02057250<br>Sanofi, Regeneron Pharmaceuticals (2014): MSC12665 / 2012-004339-21 / U1111-1130-9931 - To Evaluate Sarilumab - SAR153191 (REGN88) - Auto-injector Device In Patients With Rheumatoid Arthritis (NCT02057250). Stand des Eintrags: Mai 2017. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT02057250 | A6: Studiendauer < 24 Wochen |
| 04 | INT12684 / NCT02017639<br>Sanofi, Regeneron Pharmaceuticals (2014): INT12684 / U1111-1140-5082 - Sarilumab Effect on the Pharmacokinetics of Simvastatin (NCT02017639). Stand des Eintrags: April 2016. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT02017639 | A5: keine RCT |
| 05 | ONE / EFC13752 / NCT02121210<br>Sanofi, Regeneron Pharmaceuticals (2014): EFC13752 / 2013-002790-22 / U1111-1143-4344 - To Evaluate the Immunogenicity and Safety of Sarilumab Administered as Monotherapy in Patients With Rheumatoid Arthritis (RA) (NCT02121210). Stand des Eintrags: Mai 2017. [Zugriff: 26.06.2017]. URL: | A3: Komparator entspricht nicht der zVT |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | https://ClinicalTrials.gov/show/NCT02121210 | |
| 06 | ACT11575 / NCT01217814<br>Sanofi, Regeneron Pharmaceuticals (2010): ACT11575 / 2010-021020-94 / U1111-1115-3763 - Effect of SAR153191 (REGN88) With Methotrexate in Patients With Active Rheumatoid Arthritis Who Failed TNF-α Blockers (NCT01217814). Stand des Eintrags: November 2012. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01217814 | A6: Studiendauer < 24 Wochen |
| 07 | DRI13925 / NCT02776735<br>Sanofi, Regeneron Pharmaceuticals (2016): DRI13925 / 2015-003999-79 / U1111-1177-3487 - An Open-label, Ascending, Repeated Dose-finding Study of Sarilumab in Children and Adolescents With Polyarticular-course Juvenile Idiopathic Arthritis (pcJIA) (NCT02776735). Stand des Eintrags: Juni 2017. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT02776735 | A1: falsche Studienpopulation |
| 08 | DRI13926 / NCT02991469<br>Sanofi, Regeneron Pharmaceuticals (2016): DRI13926 / 2015-004000-35 / U1111-1177-3584 - A Repeated Dose-finding Study of Sarilumab in Children and Adolescents With Systemic Juvenile Idiopathic Arthritis (sJIA) (NCT02991469). Stand des Eintrags: Juni 2017. [Zugriff: 26.06.2017]. URL: https://clinicalTrials.gov/show/NCT02991469 | A1: falsche Studienpopulation |
| 09 | 6R88-RA-1309 / NCT02097524<br>Pharmaceuticals R, Sanofi (2014): 6R88-RA-1309 - Single-dose Study to Describe the Pharmacodynamics (PD) and Safety of Sarilumab (REGN88/SAR153191) and Tocilizumab in Adults With Rheumatoid Arthritis (RA) (NCT02097524). Stand des Eintrags: Mai 2015. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT02097524 | A6: Studiendauer < 24 Wochen |
| 10 | PKM12058 / NCT01328522<br>Sanofi, Regeneron Pharmaceuticals (2011): PKM12058 / U1111-1119-2883 - Comparison of the Safety and Pharmacokinetics of Two SAR153191 (REGN88) Drug Products in Rheumatoid Arthritis Patients (NCT01328522). Stand des Eintrags: Oktober 2011. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01328522 | A6: Studiendauer < 24 Wochen |
| 11 | DRI11073 / NCT01061723<br>Sanofi, Regeneron Pharmaceuticals (2010): DRI11073 / 2009-016068-35 - Dose Ranging Study to Evaluate the Efficacy and Safety of SAR153191 (REGN88) in Patients With Ankylosing Spondylitis (NCT01061723). Stand des Eintrags: August 2012. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01061723 | A1: falsche Studienpopulation |
| 12 | TDU10809 / 6R88-RA-0801 / NCT01055899<br>Pharmaceuticals R, Sanofi (2008): 6R88-RA-0801 - Safety Study | A6: Studiendauer < 24 Wochen |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | of a Single Dose of REGN88(SAR153191)in Patients With Rheumatoid Arthritis (NCT01055899). Stand des Eintrags: September 2013. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01055899 | |
| 13 | TARGET / EFC10832 / NCT01709578<br>Sanofi (2012): EFC10832 / U1111-1115-8466 - To Evaluate The Effect Of SAR153191 (REGN88) Added To Other RA Drugs In Patients With RA Who Are Not Responding To Or Intolerant Of Anti-TNF Therapy (SARIL-RA-TARGET) (NCT01709578). Stand des Eintrags: November 2015. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01709578 | A3: Komparator entspricht nicht der zVT |
| 14 | PDY14191 / NCT02404558<br>Sanofi, Regeneron Pharmaceuticals (2015): PDY14191 / U1111-1163-1359 - Single-dose Study to Describe the Safety of Sarilumab and Tocilizumab in Patients With Rheumatoid Arthritis (NCT02404558). Stand des Eintrags: März 2016. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT02404558 | A6: Studiendauer < 24 Wochen |
| 15 | EXTEND / LTS11210 / NCT01146652<br>Sanofi, Regeneron Pharmaceuticals (2010): LTS11210\|2010-019262-86 - Long Term Evaluation of Sarilumab in Rheumatoid Arthritis Patients (SARIL-RA-EXTEND) (NCT01146652). Stand des Eintrags: Oktober 2015. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01146652 | A5: keine RCT |
| 16 | TDR10805 / 6R88-RA-0802 / NCT01011959<br>Pharmaceuticals R, Sanofi (2008): 6R88-RA-0802 - A Multiple-dose Study of the Safety and Tolerability of REGN88(SAR153191) in Rheumatoid Arthritis Subjects (NCT01011959). Stand des Eintrags: September 2013. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01011959 | A6: Studiendauer < 24 Wochen |
| 17 | ACT13480 / NCT01900431<br>Sanofi (2013): ACT13480 / 2012-004845-34 / U1111-1130-6500 - Phase II Study to Analyze Sarilumab in Non-Infectious Uveitis (NCT01900431). Stand des Eintrags: Mai 2017. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01900431 | A1: falsche Studienpopulation |
| 18 | TDU13402 / NCT01850680<br>Sanofi, Regeneron Pharmaceuticals (2013): TDU13402 / U1111-1134-0048 - Single Ascending Dose Study of Safety and Tolerability of Sarilumab and Methotrexate in Japanese Patients With Rheumatoid Arthritis (NCT01850680). Stand des Eintrags: Januar 2014. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01850680 | A6: Studiendauer < 24 Wochen |
| 19 | ACT10804 / 6R88-RA-0803 / NCT01026519<br>Pharmaceuticals R, Sanofi (2008): 6R88-RA-0803 - A Safety and Tolerability Study of REGN88(SAR153191) in Subjects With Rheumatoid Arthritis (NCT01026519). Stand des Eintrags: September 2013. [Zugriff: 26.06.2017]. URL: | A6: Studiendauer < 24 Wochen |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | https://ClinicalTrials.gov/show/NCT01026519 | |
| 20 | KAKEHASI / EFC14059 / NCT02293902<br>Sanofi, Regeneron Pharmaceuticals (2014): EFC14059 / U1111-1155-7401 - A Study Assessing the Efficacy and Safety of Sarilumab Added to MTX in Japanese Patients With Moderately to Severely Active Rheumatoid Arthritis (SARIL-RA-KAKEHASI) (NCT02293902). Stand des Eintrags: Dezember 2016. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT02293902 | A3: Komparator entspricht nicht der zVT |
| 21 | ASCERTAIN / SFY13370 / NCT01768572<br>Sanofi, Regeneron Pharmaceuticals (2013): SFY13370 / 2012-003536-23 / U1111-1133-7839 - To Evaluate The Safety of SAR153191 (REGN88) and Tocilizumab Added to Other RA Drugs in Patients With RA Who Are Not Responding to or Intolerant of Anti-TNF Therapy (SARIL-RA-ASCERTAIN) (NCT01768572). Stand des Eintrags: November 2015. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01768572 | A3: Dosierung des Komparators entspricht nicht der in Deutschland zugelassenen Dosierung [a] |
| 22 | COMPARE / EFC11574 / NCT01764997<br>Sanofi, Regeneron Pharmaceuticals (2013): EFC11574 / 2012-001984-66 / U1111-1131-6653 - An Evaluation of Sarilumab Plus Methotrexate Compared to Etanercept Plus Methotrexate in RA Patients Not Responding to Adalimumab Plus Methotrexate (NCT01764997). Stand des Eintrags: November 2015. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT01764997 | A7: Studie wurde vorzeitig beendet, es liegt lediglich ein verkürzter Studienbericht vor [b] |
| 23 | HARUKA / LTS13618 / NCT02373202<br>Sanofi, Regeneron Pharmaceuticals (2015): LTS13618 / U1111-1160-6525 - A Study Assessing the Safety and Efficacy of Sarilumab Added to Non-MTX DMARDs or as Monotherapy in Japanese Patients With Active Rheumatoid Arthritis (SARIL-RA-HARUKA) (NCT02373202). Stand des Eintrags: Dezember 2016. [Zugriff: 26.06.2017]. URL: https://ClinicalTrials.gov/show/NCT02373202 | A3: Komparator entspricht nicht der zVT |

[a] In der Studie ASCERTAIN wurde die Dosierung des Studienkomparators Tocilizumab der US-amerikanischen Zulassung entsprechend gewählt. Diese sieht eine initiale Dosis von 4 mg/kg Tocilizumab i.v. alle vier Wochen vor, welche bei unzureichendem Ansprechen auf 8 mg/kg alle vier Wochen erhöht werden kann. In Deutschland ist lediglich die Dosierung 8 mg/kg Tocilizumab i.v. alle vier Wochen zugelassen.

[b] Die Studie COMPARE wurde vorzeitig abgebrochen, da der Einschluss von Patienten in die randomisierte doppelblinde Behandlungsphase nicht im vorgesehenen Zeitplan realisiert werden konnte. Die Beendigung der Studie geschah nicht aus Sicherheitsgründen. Insgesamt wurden 43 Patienten in die doppelblinde Behandlungsphase eingeschlossen und auf drei Behandlungsarme randomisiert. Von diesen 43 Patienten erhielten lediglich 11 Patienten die vorgesehene 24-wöchige Behandlung mit Sarilumab oder Etanercept. Aus diesem Grund wurde keine Auswertung zur Wirksamkeit der Behandlung durchgeführt. Die Auswertungen zu den geplanten Sicherheitsendpunkten sind im verkürzten Studienbericht dargestellt. Aufgrund der kleinen Patientenzahl und überwiegend zu kurzen Behandlungsdauer sind diese Auswertungen für die Nutzenbewertung von Sarilumab nicht relevant.

Tabelle 4-188: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – EU-CTR

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| 24 | EASY / MSC12665<br>Sanofi-Aventis (2014): MSC12665 - A Multicenter, Randomized, Open-Label, Parallel-Group Usability Study Of The Sarilumab Auto-Injector Device And A Prefilled Syringe In Patients With Moderate To Severe Active Rheumatoid Arthritis Who Are Candidates For Anti-IL6R Therapy (2012-004339-21). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-004339-21 | A6: Studiendauer < 24 Wochen |
| 25 | ONE / EFC13752<br>Sanofi-Aventis (2014): EFC13752 - An open-label, randomized, parallel group study assessing the immunogenicity and safety of sarilumab administered as monotherapy in patients with active rheumatoid arthritis (2013-002790-22). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-002790-22 | A3: Komparator entspricht nicht der zVT |
| 26 | COMPARE / EFC11574<br>Sanofi-Aventis (2013): EFC11574 - A Randomized, Controlled Study of Sarilumab and Methotrexate (MTX) Versus Etanercept and MTX in Patients with Rheumatoid Arthritis (RA) and an Inadequate Response to 4 Months of Treatment with Adalimumab and MTX (2012-001984-66). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-001984-66 | A7: Studie wurde vorzeitig beendet, es liegt lediglich ein verkürzter Studienbericht vor [a] |
| 27 | ASCERTAIN / SFY13370<br>Sanofi-Aventis (2013): SFY13370 - A Randomized, Double-Blind, Double-Dummy Study Assessing The Safety and Tolerability of Sarilumab and Tocilizumab In Patients With Rheumatoid Arthritis Who Are Inadequate Responders to or Intolerant of TNF Antagonists (2012-003536-23). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-003536-23 | A3: Dosierung des Komparators entspricht nicht der in Deutschland zugelassenen Dosierung [b] |
| 28 | EXTEND / LTS11210<br>Sanofi-Aventis (2010): LTS11210 - A multi-center, uncontrolled extension study evaluating the efficacy and safety of sarilumab in patients with active Rheumatoid Arthritis (RA) (2010-019262-86). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-019262-86 | A5: keine RCT |
| 29 | MOBILITY / EFC11072<br>Sanofi-Aventis (2010): EFC11072 - A randomized, double-blind, placebo-controlled, multicenter, two-part, dose ranging and confirmatory study with an operationally seamless design, evaluating efficacy and safety of SAR153191 on top of methotrexate (MTX) in patients with active rheumatoid arthritis who are inadequate responders to MTX therapy (2009-016266-90). | A3: Komparator entspricht nicht der zVT |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-016266-90 | |
| 30 | ACT13480<br>Sanofi-Aventis (2013): ACT13480 - A Randomized, Double-Masked and Placebo-Controlled Study to Evaluate the Efficacy and Safety of Sarilumab Administered Subcutaneously Every 2 Weeks in Patients with Non-Infectious, Intermediate, Posterior or Pan-Uveitis (NIU) (2012-004845-34). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-004845-34 | A1: falsche Studienpopulation |
| 31 | TARGET / EFC10832<br>Sanofi-Aventis (2012): EFC10832 - A Randomized, Double-blind, Parallel, Placebo-controlled Study Assessing The Efficacy and Safety of Sarilumab Added To Non-biologic DMARD Therapy In Patients With Rheumatoid Arthritis Who Are Inadequate Responders To Or Intolerant Of TNF-α Antagonists (2011-003538-16). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-003538-16 | A3: Komparator entspricht nicht der zVT |
| 32 | DRI13925<br>Sanofi-Aventis (2016): DRI13925 - An Open-label, Sequential, Ascending, Repeated Dose-finding Study of Sarilumab, Administered with Subcutaneous (SC) Injection, in Children and Adolescents, Aged 2 to 17 Years, with Polyarticular-course Juvenile Idiopathic Arthritis (pcJIA) Followed by an Extension Phase (2015-003999-79). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003999-79 | A1: falsche Studienpopulation |
| 33 | DRI13926<br>Sanofi-Aventis (2016): DRI13926 - An Open-label, Sequential, Ascending, Repeated Dose-finding Study of Sarilumab, Administered with Subcutaneous (SC) Injection, in Children and Adolescents, Aged 1 to 17 Years, with Systemic Juvenile Idiopathic Arthritis (sJIA), Followed by an Extension Phase (2015-004000-35). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-004000-35 | A1: falsche Studienpopulation |
| 34 | DRI11073<br>Sanofi-Aventis (2009): DRI11073 - A randomized double blind-placebo controlled dose ranging study to evaluate the efficacy and safety of SAR153191 in patients with Ankylosing Spondylitis (AS) (2009-016068-35). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-016068-35 | A1: falsche Studienpopulation |
| 35 | ACT11575<br>Sanofi-Aventis (2010): ACT11575 - A randomized, double-blind, | A6: Studiendauer < 24 Wochen |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | parallel-group, placebo- and active calibrator-controlled study assessing the clinical benefit of SAR153191 subcutaneous (SC) on top of methotrexate (MTX) in patients with active rheumatoid arthritis (RA) who have failed previous tumor necrosis factor-alpha (TNF-α) antagonists. (2010-021020-94). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-021020-94 | |
| 36 | LTS11298<br>Sanofi-Aventis (2010): LTS11298 - A multicenter uncontrolled extension study evaluating the long term safety and efficacy of SAR153191 in patients with Ankylosing Spondylitis (AS) (2010-019263-11). [Zugriff: 30.06.2017]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-019263-11 | A1: falsche Studienpopulation |

[a] Die Studie COMPARE wurde vorzeitig abgebrochen, da der Einschluss von Patienten in die randomisierte doppelblinde Behandlungsphase nicht im vorgesehenen Zeitplan realisiert werden konnte. Die Beendigung der Studie geschah nicht aus Sicherheitsgründen. Insgesamt wurden 43 Patienten in die doppelblinde Behandlungsphase eingeschlossen und auf drei Behandlungsarme randomisiert. Von diesen 43 Patienten erhielten lediglich 11 Patienten die vorgesehene 24-wöchige Behandlung mit Sarilumab oder Etanercept. Aus diesem Grund wurde keine Auswertung zur Wirksamkeit der Behandlung durchgeführt. Die Auswertungen zu den geplanten Sicherheitsendpunkten sind im verkürzten Studienbericht dargestellt. Aufgrund der kleinen Patientenzahl und überwiegend zu kurzen Behandlungsdauer sind diese Auswertungen für die Nutzenbewertung von Sarilumab nicht relevant.

[b] In der Studie ASCERTAIN wurde die Dosierung des Studienkomparators Tocilizumab der US-amerikanischen Zulassung entsprechend gewählt. Diese sieht eine initiale Dosis von 4 mg/kg Tocilizumab i.v. alle vier Wochen vor, welche bei unzureichendem Ansprechen auf 8 mg/kg alle vier Wochen erhöht werden kann. In Deutschland ist lediglich die Dosierung 8 mg/kg Tocilizumab i.v. alle vier Wochen zugelassen.

Tabelle 4-189: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – ICTRP

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| 37 | DRI13926<br>Sanofi-Aventis (2016): DRI13926 - An Open-label, Sequential, Ascending, Repeated Dose-finding Study of Sarilumab, Administered with Subcutaneous (SC) Injection, in Children and Adolescents, Aged 1 to 17 Years, with Systemic Juvenile Idiopathic Arthritis (sJIA), Followed by an Extension Phase (EUCTR2015-004000-35-GB). Stand des Eintrags: 03.04.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-004000-35-GB | A1: falsche Studienpopulation |
| 38 | DRI13925<br>Sanofi-Aventis (2016): DRI13925 - An Open-label, Sequential, Ascending, Repeated Dose-finding Study of Sarilumab, Administered with Subcutaneous (SC) Injection, in Children and Adolescents, Aged 2 to 17 Years, with Polyarticular-course Juvenile Idiopathic Arthritis (pcJIA) Followed by an Extension Phase (EUCTR2015-003999-79-CZ). Stand des Eintrags: | A1: falsche Studienpopulation |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | 29.05.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003999-79-CZ | |
| 39 | PDY14191 / NCT02404558<br>Sanofi-Aventis (2015): U1111-1163-1359;PDY14191 - An Open-label, Randomized, Parallel Group, Single-dose Study to Describe the Safety of IL-6 Receptor Blockade With Sarilumab or Tocilizumab Monotherapy in Japanese Patients With Rheumatoid Arthritis (NCT02404558). Stand des Eintrags: 11.04.2016. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02404558 | A6: Studiendauer < 24 Wochen |
| 40 | HARUKA / LTS13618 / NCT02373202<br>Sanofi-Aventis (2015): U1111-1160-6525;LTS13618 - A Randomized, Double-blind, Multicenter Study Evaluating the Safety and Efficacy of Sarilumab Added to Non-MTX DMARDs or as Monotherapy in Japanese Patients With Active Rheumatoid Arthritis (NCT02373202). Stand des Eintrags: 09.01.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02373202 | A3: Komparator entspricht nicht der zVT |
| 41 | KAKEHASI / EFC14059 / NCT02293902<br>Sanofi-Aventis (2014): U1111-1155-7401;EFC14059 - A Randomized, Double-blind, Multicenter Study With a Placebo-controlled Period Assessing the Efficacy and Safety of Sarilumab Added to Methotrexate (MTX) in Japanese Patients With Moderately to Severely Active Rheumatoid Arthritis Who Are Inadequate Responders to MTX Therapy (NCT02293902). Stand des Eintrags: 09.01.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02293902 | A3: Komparator entspricht nicht der zVT |
| 42 | ONE / EFC13752<br>Sanofi-Aventis (2014): EFC13752 - An open-label, randomized, parallel group study assessing the immunogenicity and safety of sarilumab administered as monotherapy in patients with active rheumatoid arthritis - SARIL-RA-ONE (EUCTR2013-002790-22-HU). Stand des Eintrags: 10.07.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-002790-22-HU | A3: Komparator entspricht nicht der zVT |
| 43 | EXTEND / LTS11210<br>Sanofi-Aventis (2014): LTS11210 - A Multi-Center, Uncontrolled Extension Study Evaluating Efficacy and Safety of Sar153191 on Top of Dmards in Patients with Active Rheumatoid Arthritis (Ra) (PER-022-13). Stand des Eintrags: 19.06.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=PER-022-13 | A5: keine RCT |
| 44 | EXTEND / LTS11210<br>Sanofi-Aventis (2012): LTS11210 - A Multi-Center, Uncontrolled Extension Study Evaluating Safety and Efficacy of Sar153191 on Top of Dmards in Patients with Active Rheumatoid Arthritis (Ra) | A5: keine RCT |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | (PER-077-11). Stand des Eintrags: 19.06.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=PER-077-11 | |
| 45 | EXTEND / LTS11210<br>Sanofi-Aventis (2010): LTS11210 - A multi-center, uncontrolled extension study evaluating efficacy and safety of SAR153191 on top of DMARDs in patients with active Rheumatoid Arthritis (RA) (EUCTR2010-019262-86-ES). Stand des Eintrags: 02.10.2012. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-019262-86-ES | A5: keine RCT |
| 46 | 6R88-RA-1309 / NCT02097524<br>Regeneron Pharmaceuticals (2014): 6R88-RA-1309 - Single-dose Study to Describe the Pharmacodynamics (PD) and Safety of Sarilumab (REGN88/SAR153191) and Tocilizumab in Adults With Rheumatoid Arthritis (RA) (NCT02097524). Stand des Eintrags: 25.05.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02097524 | A6: Studiendauer < 24 Wochen |
| 47 | TARGET / EFC10832<br>Sanofi-Aventis (2014): EFC10832 - A RANDOMIZED, DOUBLE-BLIND, PARALLEL, PLACEBO-CONTROLLED STUDY ASSESSING THE EFFICACY AND SAFETY OF SARILUMAB ADDED TO NON-BIOLOGIC DMARD THERAPY IN PATIENTS WITH RHEUMATOID ARTHRITIS WHO ARE INADEQUATE RESPONDERS TO OR INTOLERANT OF TNF-ANTAGONISTS (PER-033-13). Stand des Eintrags: 19.06.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=PER-033-13 | A3: Komparator entspricht nicht der zVT |
| 48 | TARGET / EFC10832<br>Sanofi-Aventis (2012): EFC10832 - A Randomized, Double-blind, Parallel, Placebo-controlled Study Assessing The Efficacy and Safety of Sarilumab Added To non-biologic DMARD Therapy In Patients With Rheumatoid Arthritis Who Are Inadequate Responders To Or Intolerant Of TNF-a Antagonists - SARIL-RA-TARGET (EUCTR2011-003538-16-LT). Stand des Eintrags: 10.07.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-003538-16-LT | A3: Komparator entspricht nicht der zVT |
| 49 | TARGET / EFC10832 / NCT01709578<br>Sanofi-Aventis (2012): U1111-1115-8466;EFC10832 - A Randomized, Double-blind, Parallel, Placebo-controlled Study Assessing the Efficacy and Safety of Sarilumab Added to Non-biologic DMARD Therapy in Patients With Rheumatoid Arthritis Who Are Inadequate Responders to or Intolerant of TNF-a Antagonists (NCT01709578). Stand des Eintrags: 11.05.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01709578 | A3: Komparator entspricht nicht der zVT |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| 50 | COMPARE / EFC11574<br>Sanofi-Aventis (2014): EFC11574 - A Randomized, Controlled Study of Sarilumab and Methotrexate (Mtx) Versus Etanercept and Mtx in Patients with Rheumatoid Arthritis (Ra) and an Inadequate Response to 4 Months of Treatment with Adalimumab and Mtx (PER-042-13). Stand des Eintrags: 19.06.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=PER-042-13 | A7: Studie wurde vorzeitig beendet, es liegt lediglich ein verkürzter Studienbericht vor [a] |
| 51 | COMPARE / EFC11574 / NCT01764997<br>Sanofi-Aventis (2013): 2012-001984-66;U1111-1131-6653;EFC11574 - A Randomized, Controlled Study of Sarilumab and Methotrexate (MTX) Versus Etanercept and MTX in Patients With Rheumatoid Arthritis (RA) and an Inadequate Response to 4 Months of Treatment With Adalimumab and MTX (NCT01764997). Stand des Eintrags: 02.03.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01764997 | A7: Studie wurde vorzeitig beendet, es liegt lediglich ein verkürzter Studienbericht vor [a] |
| 52 | EASY / MSC12665 / NCT02057250<br>Sanofi-Aventis (2014): 2012-004339-21;U1111-1130-9931;MSC12665 - A Multicenter, Randomized, Open-Label, Parallel-Group Usability Study Of The Sarilumab Auto-Injector Device And A Prefilled Syringe In Patients With Moderate To Severe Active Rheumatoid Arthritis Who Are Candidates For Anti-IL6R Therapy (NCT02057250). Stand des Eintrags: 17.11.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02057250 | A6: Studiendauer < 24 Wochen |
| 53 | INT12684 / NCT02017639<br>Sanofi-Aventis (2013): U1111-1140-5082;INT12684 - A Multi-center, Open-label, 2-treatment, Single-sequence Clinical Study to Evaluate the Effects of a Single 200 mg Subcutaneous Injection of Sarilumab on the Pharmacokinetics of a Single 40 mg Oral Dose of Simvastatin, With Optional 1-year Extension of Open Label Treatment of Sarilumab, in Patients With Rheumatoid Arthritis (NCT02017639). Stand des Eintrags: 18.04.2016. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02017639 | A5: keine RCT |
| 54 | ACT13480<br>Sanofi-Aventis (2013): ACT13480 - A Randomized, Double-Masked and Placebo-Controlled Study to Evaluate the Efficacy and Safety of Sarilumab Administered Subcutaneously Every 2 Weeks in Patients with Non-Infectious, Intermediate, Posterior or Pan-Uveitis (NIU) - SARILNIUSATURN (EUCTR2012-004845-34-CZ). Stand des Eintrags: 08.08.2016. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-004845-34-CZ | A1: falsche Studienpopulation |
| 55 | TDU13402 / NCT01850680<br>Sanofi-Aventis (2013): U1111-1134-0048;TDU13402 - A Randomized, Double-Blind, Placebo-Controlled, Single Ascending | A6: Studiendauer < 24 Wochen |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | Dose Study of the Safety and Tolerability of Subcutaneously Administered Sarilumab in Japanese Patients With Rheumatoid Arthritis Receiving Concomitant Methotrexate (NCT01850680). Stand des Eintrags: 19.02.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01850680 | |
| 56 | ASCERTAIN / SFY13370<br>Sanofi-Aventis (2013): SFY13370 - A Randomized, Double-Blind, Double-Dummy Study Assessing The Safety and Tolerability of Sarilumab and Tocilizumab In Patients With Rheumatoid Arthritis Who Are Inadequate Responders to or Intolerant of TNF Antagonists - ASCERTAIN (EUCTR2012-003536-23-CZ). Stand des Eintrags: 04.08.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-003536-23-CZ | A3: Dosierung des Komparators entspricht nicht der in Deutschland zugelassenen Dosierung [b] |
| 57 | MOBILITY / EFC11072<br>Sanofi-Aventis (2011): EFC11072 - A Randomized, Double-Blind, Placebo-Controlled, Multicenter, Two-Part, Dose Ranging and Confirmatory Study with an Operationally Seamless Design, Evaluating Efficacy and Safety of Sar153191 in Addition to Methotrexate (Mtx) in Patients with Active Rheumatoid Arthritis Who Are Inadequate Responders to Mtx Therapy (PER-072-11). Stand des Eintrags: 19.06.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=PER-072-11 | A3: Komparator entspricht nicht der zVT |
| 58 | MOBILITY / EFC11072<br>Sanofi-Aventis (2009): EFC11072;EFC11072 - A randomized, double-blind, placebo-controlled, multicenter, two-part, dose ranging and confirmatory study with an operationally seamless design, evaluating efficacy and safety of SAR153191 on top of methotrexate (MTX) in patients with active rheumatoid arthritis who are inadequate responders to MTX therapy - MOBILITY (EUCTR2009-016266-90-ES). Stand des Eintrags: 11.09.2012. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-016266-90-ES | A3: Komparator entspricht nicht der zVT |
| 59 | MOBILITY / EFC11072<br>Sanofi-Aventis (2010): EFC11072 - A Ramdomized, Double-Blind, Placebo-Controled, Multicenter, Two-Part, Dose Ranging and Confirmatory Study with an Operationally Seamless Design, Evaluating Efficacy and Safety of Sar153191 on Top Pf Methotrexate (Mtx) in Patients with Active Rheumatoid Arthritis Who Are Inadequate Responers to Mtx Therapy (PER-049-10). Stand des Eintrags: 19.06.2017. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=PER-049-10 | A3: Komparator entspricht nicht der zVT |
| 60 | PKM12058 / NCT01328522<br>Sanofi-Aventis (2011): U1111-1119-2883;PKM12058 - Comparison of the Safety and Pharmacokinetics of Two SAR153191 (REGN88) Drug Products in Rheumatoid Arthritis Patients (NCT01328522). Stand des Eintrags: 19.02.2015. [Zugriff: | A6: Studiendauer < 24 Wochen |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01328522 | |
| 61 | ACT11575<br>Sanofi-Aventis (2010): ACT11575 - A randomized, double-blind, parallel-group, placebo- and active calibrator-controlled study assessing the clinical benefit of SAR153191 subcutaneous (SC) on top of methotrexate (MTX) in patients with active rheumatoid arthritis (RA) who have failed previous tumor necrosis factor-alpha (TNF-a) antagonists. - Effect of SAR153191with Methotrexate in patients with active rheumatoid arthritis who failed TNF-a (EUCTR2010-021020-94-DE). Stand des Eintrags: 19.03.2012. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-021020-94-DE | A6: Studiendauer < 24 Wochen |
| 62 | LTS11298<br>Sanofi-Aventis (2010): LTS11298;2010-019263-11-BE - A multicenter uncontrolled extension study evaluating the long term safety and efficacy of SAR153191 in patients with Ankylosing Spondylitis (AS) - SUSTAIN (EUCTR2010-019263-11-AT). Stand des Eintrags: 26.06.2012. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-019263-11-AT | A1: falsche Studienpopulation |
| 63 | TDU10809 / 6R88-0801 / NCT01055899<br>Regeneron Pharmaceuticals (2010): 6R88-RA-0801 - A Randomized, Double-Blind, Placebo-Controlled, Single Ascending Dose Study of the Safety and Tolerability of Subcutaneously Administered REGN88 in Subjects With Rheumatoid Arthritis Receiving Concomitant Methotrexate (NCT01055899). Stand des Eintrags: 19.02.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01055899 | A6: Studiendauer < 24 Wochen |
| 64 | ACT10804 / 6R88-RA-0803 / NCT01026519<br>Regeneron Pharmaceuticals (2009): 6R88-RA-0803 - A Single-dose, Double-Blind, Placebo-Controlled, Parallel Group Safety, Tolerability and Pharmacodynamic Study of Subcutaneous REGN88 in Subjects With Rheumatoid Arthritis Receiving Concomitant Methotrexate (NCT01026519). Stand des Eintrags: 19.02.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01026519 | A6: Studiendauer < 24 Wochen |
| 65 | DRI11073<br>Sanofi-Aventis (2009): DRI11073 - A randomized double blind-placebo controlled dose ranging study to evaluate the efficacy and safety of SAR153191 in patients with Ankylosing Spondylitis (AS) - ALIGN (EUCTR2009-016068-35-BE). Stand des Eintrags: 18.08.2014. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-016068-35-BE | A1: falsche Studienpopulation |
| 66 | TDR10805 / 6R88-RA-0802 / NCT01011959 | A6: Studiendauer < 24 Wochen |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | Regeneron Pharmaceuticals (2009): 6R88-RA-0802 - A Multi-Center, Randomized, Double-Blind, Placebo-Controlled, Ascending Parallel-group Study of the Safety and Tolerability of REGN88 in Subjects With Rheumatoid Arthritis Receiving Concomitant Methotrexate (NCT01011959). Stand des Eintrags: 19.02.2015. [Zugriff: 30.06.2017]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01011959 | |

[a] Die Studie COMPARE wurde vorzeitig abgebrochen, da der Einschluss von Patienten in die randomisierte doppelblinde Behandlungsphase nicht im vorgesehenen Zeitplan realisiert werden konnte. Die Beendigung der Studie geschah nicht aus Sicherheitsgründen. Insgesamt wurden 43 Patienten in die doppelblinde Behandlungsphase eingeschlossen und auf drei Behandlungsarme randomisiert. Von diesen 43 Patienten erhielten lediglich 11 Patienten die vorgesehene 24-wöchige Behandlung mit Sarilumab oder Etanercept. Aus diesem Grund wurde keine Auswertung zur Wirksamkeit der Behandlung durchgeführt. Die Auswertungen zu den geplanten Sicherheitsendpunkten sind im verkürzten Studienbericht dargestellt. Aufgrund der kleinen Patientenzahl und überwiegend zu kurzen Behandlungsdauer sind diese Auswertungen für die Nutzenbewertung von Sarilumab nicht relevant.

[b] In der Studie ASCERTAIN wurde die Dosierung des Studienkomparators Tocilizumab der US-amerikanischen Zulassung entsprechend gewählt. Diese sieht eine initiale Dosis von 4 mg/kg Tocilizumab i.v. alle vier Wochen vor, welche bei unzureichendem Ansprechen auf 8 mg/kg alle vier Wochen erhöht werden kann. In Deutschland ist lediglich die Dosierung 8 mg/kg Tocilizumab i.v. alle vier Wochen zugelassen.

Tabelle 4-190: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – PharmNet.Bund

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| 67 | DRI13925<br>Sanofi-Aventis (2015): DRI13925; U1111-1177-3487 - An Open-label, Sequential, Ascending, Repeated Dose-finding Study of Sarilumab, Administered with Subcutaneous (SC) Injection, in Children and Adolescents, Aged 2 to 17 Years, with Polyarticular-course Juvenile Idiopathic Arthritis (pcJIA) Followed by an Extension Phase (2015-003999-79). [Zugriff: 30.06.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html | A1: falsche Studienpopulation |
| 68 | ACT13480<br>Sanofi-Aventis (2012): ACT13480 - A Randomized, Double-Masked and Placebo-Controlled Study to Evaluate the Efficacy and Safety of Sarilumab Administered Subcutaneously Every 2 Weeks in Patients with Non-Infectious, Intermediate, Posterior or Pan-Uveitis (NIU) (2012-004845-34). [Zugriff: 30.06.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html | A1: falsche Studienpopulation |
| 69 | COMPARE / EFC11574<br>Sanofi-Aventis (2012): EFC11574 - A Randomized, Controlled Study of Sarilumab and Methotrexate (MTX) Versus Etanercept and MTX in Patients with Rheumatoid Arthritis (RA) and an Inadequate Response to 4 Months of Treatment with Adalimumab and MTX (2012-001984-66). [Zugriff: 30.06.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische- | A7: Studie wurde vorzeitig beendet, es liegt lediglich ein verkürzter Studienbericht vor [a] |

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| | pruefungen/index.html | |
| 70 | TARGET / EFC10832<br>Sanofi-Aventis (2011): EFC10832 - A Randomized, Double-blind, Parallel, Placebo-controlled Study Assessing The Efficacy and Safety of Sarilumab Added To Non-biologic DMARD Therapy In Patients With Rheumatoid Arthritis Who Are Inadequate Responders To Or Intolerant Of TNF-alpha Antagonists (2011-003538-16). [Zugriff: 30.06.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html | A3: Komparator entspricht nicht der zVT |
| 71 | ACT11575<br>Sanofi-Aventis (2010): ACT11575 - A randomized, double-blind, parallel-group, placebo- and active calibrator-controlledstudy assessing the clinical benefit of SAR153191 subcutaneous (SC) on top of methotrexate (MTX) in patients with active rheumatoid arthritis (RA) who have failed previous tumor necrosis factor-alpha (TNF-alpha) antagonists (2010-021020-94). [Zugriff: 30.06.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html | A6: Studiendauer < 24 Wochen |
| 72 | EXTEND / LTS11210<br>Sanofi-Aventis (2010): LTS11210 - A multi-center, uncontrolled extension study evaluating the efficacy and safety of sarilumab in patients with active Rheumatoid Arthritis (RA) (2010-019262-86). [Zugriff: 30.06.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html | A5: keine RCT |
| 73 | MOBILITY / EFC11072<br>Sanofi-Aventis (2009): EFC11072 - A randomized, double-blind, placebo-controlled, multicenter, two-part, dose ranging and confirmatory study with an operationally seamless design, evaluating efficacy and safety of SAR153191 on top of methotrexate (MTX) in patients with active rheumatoid arthritis who are inadequate responders to MTX therapy (2009-016266-90). [Zugriff: 30.06.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html | A3: Komparator entspricht nicht der zVT |
| 74 | DRI11073<br>Sanofi-Aventis (2009): DRI11073 - A randomized double blind-placebo controlled dose ranging study to evaluate the efficacy and safety of SAR153191 in patients with Ankylosing Spondylitis (AS) (2009-016068-35). [Zugriff: 30.06.2017]. URL: https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html | A1: falsche Studienpopulation |

[a] Die Studie COMPARE wurde vorzeitig abgebrochen, da der Einschluss von Patienten in die randomisierte doppelblinde Behandlungsphase nicht im vorgesehenen Zeitplan realisiert werden konnte. Die Beendigung der Studie geschah nicht aus Sicherheitsgründen. Insgesamt wurden 43 Patienten in die doppelblinde Behandlungsphase eingeschlossen und auf drei Behandlungsarme randomisiert. Von diesen 43 Patienten erhielten lediglich 11 Patienten die vorgesehene 24-wöchige Behandlung mit Sarilumab oder Etanercept. Aus diesem Grund wurde keine Auswertung zur Wirksamkeit der Behandlung durchgeführt. Die Auswertungen zu den geplanten Sicherheitsendpunkten sind im verkürzten Studienbericht dargestellt. Aufgrund der kleinen Patientenzahl und überwiegend zu kurzen Behandlungsdauer sind diese Auswertungen für die

| Nr. | Studie | Ausschlussgrund |
|---|---|---|
| Nutzenbewertung von Sarilumab nicht relevant. | | |

**Anhang 4-D2: Suche nach RCT für indirekte Vergleiche**

Es wurde keine Suche nach RCT für indirekte Vergleiche durchgeführt.

**Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien**

Es wurde keine Suche nach nicht randomisierten vergleichenden Studien durchgeführt.

**Anhang 4-D4: Suche nach weiteren Untersuchungen**

Es wurde keine Suche nach weiteren Untersuchungen durchgeführt.

## Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

*Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-191 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 4-191 inklusive eines Flow-Charts für den Patientenfluss.*

Tabelle 4-191 (Anhang): Studiendesign und -methodik für Studie MONARCH

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| **Studienziel** | | |
| **2 b** | Genaue Ziele, Fragestellung und Hypothesen | Ziel der Studie war die Untersuchung der Wirksamkeit und Sicherheit von Sarilumab (Monotherapie) versus Adalimumab (Monotherapie) bei Patienten mit aktiver rheumatoider Arthritis (RA) die entweder intolerant gegenüber Methotrexat (MTX) oder ungeeignet für eine Weiterbehandlung mit MTX sind, oder die nach mindestens 12 Wochen Behandlung mit MTX unzureichend angesprochen haben.<br><br>**Primäres Ziel:**<br>Die Überlegenheit von Sarilumab gegenüber Adalimumab nach 24-wöchiger Behandlung sollte bezüglich klinischer Anzeichen und Symptome mittels DAS28-ESR gezeigt werden.<br>**Sekundäre Ziele:**<br>Die Überlegenheit von Sarilumab gegenüber Adalimumab nach 24-wöchiger Behandlung sollte bezüglich<br>• einer Reduktion von klinischen Anzeichen und Symptomen der RA nach 24-wöchiger Behandlungsphase (DAS28-ESR-Remission, ACR20/50/70 usw.) und<br>• einer Verbesserung der Lebensqualität nach 24-wöchiger Behandlungsphase gezeigt werden.<br>Die Sicherheit und Verträglichkeit von Sarilumab inklusive der Immunogenität sollte untersucht werden.<br>**Explorative Ziele:**<br>• Untersuchung der Erhaltung der Wirksamkeit im Zeitverlauf<br>• Untersuchung der Pharmakokinetik von Sarilumab als Monotherapie |
| **Methoden** | | |
| **3** | Studiendesign | |
| **3a** | Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis | Parallelgruppen-Design, multizentrische, randomisierte, doppelblinde, doppel-Dummy Phase III-Studie.<br>Randomisierung im Verhältnis 1:1 in die parallelen Behandlungsgruppen Sarilumab und Adalimumab. |
| **3b** | Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung | **Clinical Trial Protocol Amendment No. 3 (29-Apr-2015):**<br>Nur gültig in Deutschland - Anpassungen des Protokolls im Hinblick auf die nationale EC Guidance:<br>• Ausschluss von inhaftierten Personen aus der Studie<br>• Ausschluss von Angestellten des Sponsors oder der Studienzentren<br>**Clinical Trial Protocol Amendment No. 4 (17-Jun-2015):**<br>• Konkretisierung bei unerwünschten Ereignissen von besonderem Interesse: Untersuchung von potentiellen opportunistischen Infektionen und Abbruch der Behandlung<br>• Korrektur von Fehlern oder Inkonsistenzen im Protokoll bezüglich der geplanten Studienereignisse und Fußnoten<br>• Korrektur von Inkonsistenzen bezüglich der Kriterien für eine Dosierungserhöhung |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | | • Konkretisierung der Anforderung an einen unabhängigen Begutachter der Gelenke<br>**Clinical Trial Protocol Amendment No. 5 (20-Nov-2015):**<br>• Modifikation der Studiendauer, um eine offene Langzeitbehandlung mit Sarilumab 200 mg q2w nach Woche 48 zu ermöglichen, bis zur kommerziellen Verfügbarkeit von Sarilumab oder bis 2020<br>• Ergänzung einer möglichen Dosisreduktion von Sarilumab auf 150 mg q2w in der Extensionsphase |
| **4** | Probanden / Patienten | |
| **4a** | Ein-/Ausschlusskriterien der Probanden / Patienten | **Einschlusskriterien:**<br>• Bestehende Diagnose der rheumatoiden Arthritis (RA) nach den ACR/EULAR-Klassifikationskriterien von 2010 mit mindestens 3-monatiger Krankheitsdauer<br>• Funktionelle Klasse I-III nach den überarbeiteten ACR-Kriterien von 1991<br>• Aktive RA definiert durch:<br>o Mindestens 6 von 66 geschwollene und 8 von 68 schmerzende Gelenke und<br>o Hochsensitives C-reaktives Protein (hs-CRP) ≥ 8 mg/L oder ESR ≥ 28mm/h, gemessen zwischen Screening und Randomisierung, und<br>o DAS28-ESR > 5,1 (Bestimmung der ESR in der Phase zwischen Screening und Randomisierung)<br>• Patienten, die nach Einschätzung des behandelnden Arztes entweder intolerant gegenüber MTX oder nicht geeignet für eine Weiterbehandlung mit MTX waren, oder die nach mindestens 12 Wochen Behandlung mit MTX bei geeigneter Dosierung (10 bis 25 mg/Woche, bzw. 6 bis 25 mg/Woche für Patienten aus der Asien-Pazifik Region) unzureichend angesprochen haben.<br>• Unterschriebene Einverständniserklärung<br>**Ausschlusskriterien:**<br>• Unter 18 Jahren bzw. Minderjährigkeit im jeweiligen Land<br>• Behandlung mit jeglichem Biologikum gegen RA; Behandlung mit einem Januskinase (JAK)-Inhibitor (z.B. Tofacitinib)<br>• Aktuelle Behandlung mit DMARDs/Immunsuppressiva wie MTX, Cyclosporin, Mycophenolat, Tacrolimus, Gold, Penicillamin, Sulfasalazin oder Hydroxychloroquin innerhalb von 2 Wochen vor der Randomisierungs-Visite oder Azathioprin, Cyclophosphamid innerhalb von 12 Wochen vor der Randomisierungs-Visite oder Leflunomid innerhalb von 8 Wochen vor der Randomisierungs-Visite, oder 4 Wochen nach Cholestyramin-Standard-Auswaschphase: Cholestyramin bei einer Dosierung von 8 g dreimal täglich für mindestens 24 Stunden oder Aktivkohle bei einer Dosierung von 50 g viermal täglich für mindestens 24 Stunden<br>• Die Verwendung von parenteralen Kortikosteroiden oder intraartikulären Kortikosteroiden innerhalb von 4 Wochen vor der Screening-Visite<br>• Die Verwendung von oralen Kortikosteroiden mit einer Dosierung |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | | höher als 10 mg Prednison (oder äquivalent) pro Tag, oder eine Änderung der Dosierung innerhalb von 4 Wochen vor der Randomisierung<br>• Patienten mit akuter oder vorausgegangener, autoimmuner oder entzündlicher systemischer Erkrankung oder lokaler Gelenkerkrankung, die keine RA ist<br>• Vorausgegangene juvenile idiopathische Arthritis oder Beginn der Arthritis vor einem Alter von 16 Jahren<br>• Schwere systemische RA, einschließlich, aber nicht beschränkt auf, Vaskulitis, Lungenfibrose, und/oder Felty-Syndrom<br>• Neue Behandlung mit einem NSAID oder COX-2-Hemmer oder Dosis-Anpassung innerhalb von 4 Wochen vor der Randomisierung, mit Ausnahme der Verwendung von niedrig dosierter Acetylsalicylsäure bei kardiovaskulären Erkrankungen<br>• Teilnahme an einer klinischen Studie innerhalb von 5 Halbwertszeiten oder 60 Tagen vor der Screening-Visite, je nachdem, was länger ist<br>• Patienten die innerhalb von 5 Jahren vor der Screening-Visite alkohol- oder drogenabhängig waren<br>• Patienten, die die Zustimmung während des Screening-Zeitraumes zurückziehen (nach Unterzeichnung des Einwilligungserklärung)<br>• Erreichen der maximalen Anzahl an Teilnehmern<br>• Bedingungen / Situationen wie:<br>o Patienten mit kurzer Lebenserwartung<br>o Bedingungen / Begleiterkrankungen, die es nicht möglich machen, die Wirksamkeitsendpunkte zu bewerten, z.B. Patienten mit chronischen Schmerzen, verursacht durch andere Krankheiten als RA (z.B. Fibromyalgie)<br>o Der Patient ist Prüfarzt, wissenschaftlicher Mitarbeiter, Apotheker, Studienkoordinator, anderer Mitarbeiter oder im Zusammenhang stehend, direkt an der Durchführung der Studie beteiligt, oder in Deutschland Mitarbeiter des Studienzentrums oder Sponsors<br>o Unkooperatives Verhalten oder jede Bedingung, die die Einhaltung der Studienprozeduren durch den Patienten gefährdet, in Deutschland Häftlinge oder aufgrund behördlicher Anordnung untergebrachte Personen<br>• Schwangere oder stillende Frauen<br>• Frauen im gebärfähigen Alter, die nicht durch eine hochwirksame Methode der Empfängnisverhütung geschützt sind oder keinen Schwangerschaftstest machen wollen oder können<br>• Ausschluss bezüglich Tuberkulose (TB):<br>o Akute TB oder vorausgegangene unvollständig behandelte TB<br>o Quantiferon -positive Patienten ohne aktive Erkrankung werden aus der Studie ausgeschlossen, es sei denn die folgenden Bedingungen sind erfüllt:<br>▪ Patienten mit vorausgegangener dokumentierter abgeschlossener Chemoprophylaxe für latente TB oder Behandlung einer aktiven TB mit ärztlicher Untersuchung durch einen Spezialisten, um eine aktive TB auszuschließen oder zu behandeln<br>▪ Patienten ohne vorangegangene Chemoprophylaxe für eine latente TB oder Behandlung einer aktiven |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | | TB, die einen Spezialisten konsultiert haben, um eine geeignete Chemoprophylaxe zu beginnen, die der Patienten seit ≥ 1 Monat einnimmt und toleriert<br>▪ Beratung und Zustimmung des Sponsors ist in jedem der genannten Fälle erforderlich<br>o Klinisch signifikante Anomalie einer Thorax-Röntgenaufnahme (mindestens posterior-anteriore Ansicht) innerhalb von 12 Wochen vor der Screening-Visite oder im Screening-Zeitraum, passend zu einer vorangegangenen oder aktiven TB<br>o Mutmaßliche extrapulmonale TB-Infektion<br>o Patienten mit einem hohen Infektionsrisiko (wie z.B. Personen mit engem Kontakt mit Personen mit aktiver oder latenter TB)<br>o Patienten die eine Bazillus Calmette Guerin (BCG) -Impfung innerhalb von 12 Monaten vor dem Screening erhalten haben<br>• Patienten mit akuter oder vorangegangener interstitieller Lungenerkrankung, diagnostiziert durch eine hochauflösende Computertomographie und/oder Lungenbiopsie mit konsistenten Erkenntnissen über Lungenfunktionstests und bestätigenden klinischen Befunden<br>• Neue Behandlung oder eine Dosisanpassung einer laufenden Medikation für Dyslipidämie, z.B. Statine, innerhalb von 6 Wochen vor der Randomisierung<br>• Patienten mit vorangegangenen invasiven opportunistischen Infektionen einschließlich, aber nicht beschränkt auf, Histoplasmose, Listeriose, Kokzidioidomykose, Candidiasis, Pneumocystis jirovecii, Aspergillose trotz Heilung oder John Cunningham (JC) Virus (progressive multifokale Leukoenzephalopathie [PML])<br>• Patienten mit Fieber (> 38 °C) oder chronischen, hartnäckigen, oder wiederkehrenden Infektionen, die eine aktive Behandlung mit Antibiotika, Virostatika oder Antimykotika innerhalb von 4 Wochen vor der Screening-Visite erfordern, oder andere häufige wiederkehrende Infektionen die vom Behandler als inakzeptabel eingestuft wurden<br>• Patienten mit unkontrolliertem Diabetes mellitus ($HbA_{1C} \geq 9\%$ bei der Screening-Visite)<br>• Patienten mit nicht verheilten Hautgeschwüren<br>• Patienten die abgeschwächten Lebendimpfstoff (wie Varizella Zoster, orale Polio oder Röteln-Impfstoffen) innerhalb von 3 Monaten vor der Randomisierungs-Visite erhalten haben<br>• Patienten, die einen positiven Test auf Hepatitis-B-Oberflächenantigen (HBs-Ag) oder Hepatitis-B-Core (HBc)-Antikörper hatten<br>• Patienten, die bei der Screening-Visite positiv auf Hepatitis-C Antikörper getestet wurden<br>• Patienten, die bei der Screening-Visite oder zuvor eine positiven HIV-Test hatten, oder die im Verdacht stehen, HIV-Positiv zu sein<br>• Patienten mit akutem oder vorangegangenem wiederkehrendem Herpes Zoster<br>• Patienten mit vorangegangener Gelenksinfektionen oder prothetischer Gelenksinfektion<br>• Krebserkrankung innerhalb von 5 Jahren vor Studienbeginn, mit Ausnahme von adäquat behandeltem in situ Zervixkarzinom sowie |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | | Basalzellkarzinom oder Plattenepithelkarzinom der Haut<br>• Akute oder vorangegangene Begleiterkrankung, die einen negativen Einfluss auf die Beteiligung an der Studie hat, wie kardiovaskuläre Erkrankungen (einschließlich Stadium III oder IV nach der New York Heart Assoziation Klassifikation), Nierenerkrankungen, neurologische Erkrankungen (einschließlich demyelinisierender Erkrankungen), aktive infektiöse Erkrankungen, endokrinologische, gastrointestinale, hepatobiliäre, nicht bösartige lymphoproliferative oder andere lymphatische Erkrankungen, Stoffwechselerkrankungen, Lungenerkrankungen<br>• Operation 4 Wochen vor der Screening-Visite oder geplante Operation während des Studienzeitraums<br>• Patienten mit vorangegangener systemischer Überempfindlichkeitsreaktion auf ein Biologikum, mit Ausnahme einer lokalen Reaktion an der Injektionsstelle<br>• Patienten mit einer der folgenden Anomalien der Laborwerte bei der Screening- oder Randomisierungs-Visite:<br>o Hämoglobin < 8,5g/dl<br>o Weiße Blutkörperchen < 3.000/$mm^3$<br>o Neutrophile < 2.000/$mm^3$<br>o Thrombozytenzahl < 150.000 Zellen/$mm^3$<br>o Aspartat-Aminotransferase (AST), Alanin-Aminotransferase (ALT) > 1,5 x obere Grenze des Normalbereichs (ULN)<br>o Bilirubin (total) > ULN, ausgenommen Morbus-Gilbert-Meulengracht liegt vor (durch genetische Tests diagnostiziert)<br>o Akute schwere unkontrollierte Hypercholesterinämie (> 350 mg/dL, 9,1 mmol/L) oder Hypertriglyzeridämie (> 500 mg/dL, 5,6 mmol/L)<br>o Patienten mit berechneter Kreatinin-Clearance < 30 ml/min (unter Verwendung der Cockroft-Gault-Formel)<br>• Patienten mit vorangegangener entzündlicher Darmerkrankung, schwerer Divertikulitis oder vorheriger Magen-Darm-Perforation |
| **4b** | Studienorganisation und Ort der Studiendurchführung | 86 Zentren in 14 Ländern aufgeteilt in 3 Regionen:<br>• Region 1 (Westliche Länder): Tschechien, Deutschland, Ungarn, Israel, Spanien, USA<br>• Region 2 (Südamerika): Chile, Peru<br>• Region 3 (andere): Südkorea, Polen, Südafrika, Rumänien, Russland, Ukraine<br>Covance Central Laboratory Services (Indiana, USA; Schweiz; Singapur) führte zentral die klinischen Laboruntersuchungen durch, eResearch Technology (Philadelphia, USA) führte die zentralen EKG-Messungen durch, Regeneron Pharmaceuticals Inc. (New York, USA) führte die pharmakokinetischen und immunologischen Analysen durch.<br>Ein unabhängiges Daten-Monitoring-Komitee (DMC) war für die Patientensicherheit in allen Sarilumab-Studien zuständig. Die Mitglieder des DMC hatten Zugriff auf alle Sicherheitsdaten der Studie und sprachen Empfehlungen zur Durchführung der Studie mit Blick auf die Patientensicherheit aus.<br>Ein unabhängiger Ausschuss zur Beurteilung kardiovaskulärer Ereignisse (cardiovascular adjudication committee, CAC) untersuchte alle kardiovaskulären Ereignisse nach einheitlichen Kriterien. Die Mitglieder des CAC waren bezüglich der Behandlung der Patienten |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | | verblindet. |
| **5** | Interventionen<br>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc. | Sarilumab und Adalimumab wurden durch subkutane (s.c.) Injektion in den Bauch oder den vorderen Oberschenkel verabreicht. Die Injektion erfolgte mittels vorgefüllter Glasspritzen.<br>**1.) Doppelblinde Behandlungsphase: 24 Wochen**<br>Sarilumab-Gruppe:<br>• Sarilumab 200 mg s.c. und Placebo für Adalimumab s.c. q2w.<br>Adalimumab-Gruppe:<br>• Adalimumab 40 mg s.c. und Placebo für Sarilumab q2w.<br>Dosiseskalation:<br>• Eine Dosierungserhöhung auf wöchentlich 40 mg Adalimumab ab Woche 16 bis einschließlich Woche 23 war erlaubt, falls ein Patient in zwei aufeinanderfolgenden Visiten eine Verbesserung um < 20% gegenüber dem Ausgangswert bei der Anzahl der geschwollenen (SJC) und schmerzhaften Gelenke (TJC) hatte. Alternativ konnten solche Patienten die Studie abbrechen und erielten dann eine individuelle Therapie.<br>• Um die Verblindung aufrecht zu erhalten, erhielten Patienten der Sarilumab-Gruppe, die bis Woche 16 in zwei aufeinanderfolgenden Visiten eine Verbesserung um < 20% gegenüber dem Ausgangswert bei der Anzahl der geschwollenen (SJC) und schmerzhaften Gelenke (TJC) erreichte, eine Placebo-Injektion s.c. q2w alternierend mit der Sarilumab-Injektion.<br>**2.) Open Label Extension: max. 276 Wochen**<br>Alle Patienten, die von der doppelblinden in die offene Behandlungsphase übergingen, erhielten s.c. Sarilumab 200 mg q2w. Die Dosis konnte auf 150 mg Sarilumab q2w reduziert werden, wenn präspezifizierte Werte für die Neutrophilenzahl, Thrombozytenzahl oder die Leberwerte unter- bzw. überschritten wurden. |
| **6** | Zielkriterien | |
| **6a** | Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten | **Primärer Endpunkt:**<br>• Änderung des DAS28-ESR-Wertes nach 24 Wochen im Vergleich zum Ausgangswert. Der DAS28-ESR wurde basierend auf 4 Variablen berechnet:<br>o Anzahl der schmerzhaften Gelenke (TJC) (ausgehend von 28 Gelenken)<br>o Anzahl der geschwollenen Gelenke (SJC) (ausgehend von 28 Gelenken)<br>o Generelle Einschätzung des Gesundheitszustandes durch den Patienten auf einer 100 mm VAS<br>o ESR in mm/h<br>**Sekundäre Endpunkte:**<br>• Anteil der Patienten mit DAS28-ESR-Remission in Woche 24 (DAS28-ESR < 2,6)<br>• Anteil der Patienten mit niedriger Krankheitsaktivität in Woche 24 (DAS28-ESR < 3,2)<br>• Änderung des DAS28-CRP-Wertes nach 24 Wochen im Vergleich zum Ausgangswert<br>• Anteil der Patienten mit DAS28-CRP-Remission in Woche 24 |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | | (DAS28-CRP < 2,6)<br>• Anteil der Patienten mit ACR20/50/70-Response nach 24 Wochen. Die ACR-Response wurde basierend auf 7 Variablen berechnet:<br>o Anzahl der schmerzhaften Gelenke (TJC; basierend auf 68 Gelenken)<br>o Anzahl der geschwollenen Gelenke (SJC; basierend auf 66 Gelenken)<br>o CRP-Wert<br>o Einschätzung des Patienten bezüglich Schmerz<br>o Einschätzung der Krankheitsaktivität durch den Patienten<br>o Einschätzung der Krankheitsaktivität durch den Arzt<br>o HAQ-DI<br>• Änderung der einzelnen Komponenten des ACR-Wertes im Vergleich zum Ausgangswert nach 24 Wochen<br>• Änderung des CDAI (clinical disease activity index) im Vergleich zum Ausgangswert nach 24 Wochen<br>• Anteil der Patienten mit CDAI-Remission (CDAI ≤ 2,8) in Woche 24<br>**Sicherheitsendpunkte:**<br>• Unerwünschte Ereignisse<br>• Schwerwiegende unerwünschte Ereignisse<br>• Unerwünschte Ereignisse, die zum dauerhaften Abbruch der Behandlung führen<br>• Unerwünschte Ereignisse von besonderem Interesse: Leukopenie, Thrombozytopenie, Infektionen, Leberwerterhöhungen, Divertikulitits, gastrointestinale Ulzerationen, Erhöhung der Lipidwerte, Überempfindlichkeit, Anaphylaxie, Reaktionen der Injektionsstelle, Krebserkrankungen, Lupus-ähnliches Syndrom, demyelinisierende Erkrankung, Schwangerschaft, Überdosierung<br>• kardiovaskuläre Ereignisse<br>• Todesfälle<br>• Rheumafaktor, antinukleäre Antikörper (ANA), anti-CCP-Antikörper<br>• anti-Sarilumab-Antikörper<br>• Röntgen-Thorax, Vitalfunktionen, Körpertemperatur, Blutdruck, EKG<br>**Lebensqualität/Patientenberichtete Endpunkte:**<br>• Short-Form-36 (Version 2.0)<br>• FACIT-Fatique (functional assessment of chronic illness therapy fatique scale)<br>• Morgensteifigkeit<br>• EQ-5D-3L (EuroQol)<br>• WPS-RA (rheumatoid arthritis-work productivity survey)<br>• RAID (rheumatoid arthritis impact of disease) |
| **6b** | Änderungen der Zielkriterien nach Studienbeginn, mit | Keine |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | Begründung | |
| **7** | Fallzahl | |
| **7a** | Wie wurden die Fallzahlen bestimmt? | Die Fallzahlberechnung beruhte auf dem primären Endpunkt (Änderung des DAS28-ESR-Wertes nach 24 Wochen im Vergleich zum Ausgangswert). Ein Unterschied von 0,6 im DAS28-ESR-Wert zwischen zwei aktiven Vergleichstherapien wurde als klinisch relevant gewertet.<br>Die Fallzahl wurde unter den folgenden Annahmen berechnet:<br>• Einer gemeinsamen Standardabweichung von 1,7 (basierend auf früheren Ergebnissen)<br>• Ein Unterschied von 0,6 im DAS28-ESR-Wert zwischen zwei aktiven Vergleichstherapien<br>• Einem zweiseitigen t-Test mit einem Signifikanzniveau von 5% und einer Power von 90%<br>Unter diesen Annahmen wurde eine Fallzahl von 170 Patienten pro Studienarm benötigt. |
| **7b** | Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch | Es waren keine Interimsanalysen geplant.<br>Der Sponsor konnte die Studie vorzeitig beenden. Kriterien für einen vorzeiten Studienabbruch wären unter anderen Bedenken bezüglich des Nutzen-Risiko-Verhältnisses von Sarilumab oder eine unzureichende Patientenrekrutierung gewesen. |
| **8** | Randomisierung, Erzeugung der Behandlungsfolge | |
| **8a** | Methode zur Generierung der zufälligen Zuteilung | Die Randomisierung erfolgte zentral durch ein interaktives Sprachdialogsystem (IVRS, Interactive Voice Response System). |
| **8b** | Einzelheiten (z. B. Blockrandomisierung, Stratifizierung) | Randomisierung im Verhältnis 1:1 in die parallelen Behandlungsgruppen Sarilumab und Adalimumab mit Stratifizierung nach Region. |
| **9** | Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment)<br>Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war | Zentrale Zuteilung per Telefon/Internet (IVRS). |
| **10** | Randomisierung, Durchführung<br>Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte | Zentrale Erstellung der Randomisierungsliste und Zuteilung der Patienten: PAREXEL International, New Jersey, USA |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | die Probanden/Patienten den Gruppen zu? | |
| **11** | Verblindung | |
| **11a** | Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen? | a) Ja<br>b) Ja<br>c) Ja<br><br>Die Verblindung erfolgte bei der Randomisierung mittels IVRS. |
| **11b** | Falls relevant, Beschreibung der Ähnlichkeit von Interventionen | Sarilumab, Adalimumab und Placebo werden in identisch aussehenden, vorgefüllten Glasspritzen bereitgestellt und identisch verblindet verpackt (Kits mit jeweils einer Spritze Adalimumab oder Sarilumab sowie einer Placebo-Spritze). |
| **12** | Statistische Methoden | |
| **12a** | Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien | **Population zur Auswertung der Wirksamkeitsendpunkte:**<br>ITT (*intent-to-treat*) - Population: Alle randomisierten Patienten. Die Patienten wurden ausgewertet wie randomisiert.<br>**Population zur Auswertung der Sicherheitsendpunkte:**<br>Safety-Population: Alle randomisierten Patienten, die mindestens eine Dosis oder Teildosis der Studienintervention (Sarilumab oder Adalimumab) erhalten haben. Die Patienten wurden ausgewertet wie behandelt.<br><br>**Statistische Auswertung der Wirksamkeitsendpunkte:**<br>Primäre und sekundäre Endpunkte wurden in der ITT-Population ausgewertet.<br>**Primärer Endpunkt:**<br>Der primäre Endpunkt, DAS28-ESR, wurde mit Hilfe eines gemischten Modelles mit Messwertwiederholungen (MMRM) analysiert. In das Modell wurden als feste Effekte Behandlung, Visite (Woche 12 und 24), Interaktion zwischen Behandlung und Visite und die Region aufgenommen. Der Baseline-Wert des DAS28-ESR wurde als stetige Kovariate aufgenommen. Das Modell lieferte den Unterschied der Least-Square-Mittelwerte (LSM) zusammen mit dem korrespondierenden 95%-Konfidenzintervall und p-Wert.<br>Daten die nach Erhöhung der Adalimumab Dosierung (auf wöchentlichen Abstand) erhoben wurden, wurden mit in die primäre Analyse aufgenommen. Daten die nach Studienabbruch gesammelt wurden, wurden in der primären Analyse nicht berücksichtigt.<br>**Sekundäre Endpunkte:**<br>• Stetige Einflussvariablen: Nach Woche 12 und 24 wurden diese mit denselben Methoden wie der primäre Endpunkt untersucht.<br>• Kategorielle Einflussvariablen: Die binären Endpunkte wurden mit Hilfe eines CMH-Tests (Cochran-Mantel-Haenszel) stratifiziert nach der Region analysiert. Der Mantel-Haenszel-Schätzer für das |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | | Odds Ratio und das korrespondierende 95%-Konfidenzintervall wurden berechnet. Patienten, die die Studie abgebrochen haben, wurden als Non-Responder analysiert.<br>• Die sekundären Endpunkte wurden in einer a priori im SAP festgelegten Hierarchie ausgewertet, um das Signifikanzniveau von $\alpha = 0,05$ zu kontrollieren.<br>**Statistische Auswertung der weiteren Endpunkte:**<br>• Sicherheitsendpunkte: Die Sicherheitsendpunkte wurden deskriptiv ausgewertet und es wurden keine statischen Test a priori geplant. Ausnahmen waren Laborparameter wie Neutrophilenzahl, Leberwerte und Lipidwerte, für welche der Gruppenunterschied und das jeweilige 95% Konfidenzintervall berechnet wurden.<br>• Lebensqualitätsendpunkte: Die Änderung in Woche 24 im Vergleich zum Ausgangswert wurde für SF-36 (PCS, MCS und Einzeldomänen), EQ-5D-3L (VAS und utility score), FACIT-Fatigue, Morgensteifigkeit-VAS und RAID mit einem MMRM-Ansatz analog zur Auswertung des primären Endpunkts berechnet.<br>**Statistische Analysen in der Open Label Extension-Phase:**<br>• Alle primären und sekundären Wirksamkeitsendpunkte sowie die Lebensqualitätsendpunkte werden im Zeitverlauf für die Gruppen Sarilumab 200 mg q2w/ Sarilumab 200 mg q2w und Adalimumab 40 mg q2w/ Sarilumab 200 mg q2w dargestellt, basierend auf den Daten der ITT-Population der Extensionsphase. |
| **12b** | Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen | **Sensitivitätsanalysen:**<br>Zwei Sensitivitätsanalysen zur Änderung des DAS28-ESR wurden nach Woche 24 durchgeführt:<br>• In der ersten Analyse, wurden die Daten der Patienten, die die Studie abgebrochen haben, inkludiert.<br>• In der zweiten Analyse wurden alle Daten nach Studienabbruch oder der Dosiserhöhung in der Adalimumab Gruppe als fehlend gesetzt und mittels Multipler Imputation ersetzt.<br>**Subgruppenanalysen:**<br>Für den primären Endpunkt „Änderung des DAS28-ESR nach 24 Wochen“ sowie für den sekundären Endpunkt „Anteil der Patienten mit DAS28-ESR-Remission in Woche 24“ wurden Subgruppenanalysen für die folgenden Faktoren durchgeführt:<br>• Geschlecht (Männer / Frauen)<br>• Rasse (kaukasisch bzw. weiß / andere)<br>• Region (Region 1: Westliche Staaten / Region 2: Süd-Amerika/ Region 3: andere)<br>• Alter (<65 Jahre / ≥65 Jahre)<br>• Gewicht (<60 kg / ≥60 kg und <100 kg / ≥100 kg)<br>• BMI (<25 / ≥25 und <30 / ≥30 kg/m$^2$)<br>• Rheumafaktor (positiv / negativ)<br>• Anti-CCP-Antikörper (positiv/ negativ)<br>• Baseline-CRP (≤15 mg/L / >15 mg/L)<br>• Baseline-ESR (≤ Median / > Median)<br>• Dauer der RA (≤ Median / > Median in Jahren; ≤ 3 Jahre / > 3 Jahre) |

| Item[a] | Charakteristikum | Studieninformation |
|---|---|---|
| | | • Anzahl der vorangegangenen DMARDs (≤1 / 2 / ≥3)<br>• MTX-Status (intolerant / ungeeignet / Non-Responder)<br>• Raucherstatus (Nicht-Raucher / Raucher / Ex-Raucher)<br>Für die Subgruppenanalysen des primären Endpunkts wurde ein MMRM-Modell angewendet. Das Modell enthielt als feste Effekte Baselinewert, Behandlung, Region, Subgruppe, Visite (Woche 12 und 24) sowie die Interaktionsterme Behandlung*Visite, Subgruppe*Behandlung und Subgruppe*Behandlung*Visite.<br>Für die Subgruppenanalysen des sekundären Endpunkts wurde ein CMH-Test (Cochran-Mantel-Haenszel) durchgeführt. Der Mantel-Haenszel-Schätzer für das Odds Ratio und das korrespondierende 95%-Konfidenzintervall wurden für jede Subgruppe berechnet. Eine logistische Regression mit den Termen Behandlung, Subgruppe, Interaktion Subgruppe*Behandlung und Region wurde durchgeführt. |
| **Resultate** | | |
| **13** | Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle) | |
| **13a** | Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die<br>a) randomisiert wurden,<br>b) tatsächlich die geplante Behandlung/Intervention erhalten haben,<br>c) in der Analyse des primären Zielkriteriums berücksichtigt wurden | a)<br>Total: n = 369<br>Sarilumab: n = 184; Adalimumab: n = 185<br>b)<br>Total: n = 368<br>Sarilumab: n = 184; Adalimumab: n = 184<br>c)<br>Total: n = 369 (ITT-Population)<br>Sarilumab: n = 184; Adalimumab: n = 185 |
| **13b** | Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen | Es wurden 540 Patienten gescreent. Davon erfüllten 171 Patienten die Einschlusskriterien nicht und 369 Patienten wurden randomisiert.<br>Ein Patient in der Adalimumab-Gruppe erhielt keine Behandlung, da die Einschluss-/ Ausschlusskriterien verletzt wurden (Randomisierungsfehler).<br>Studienabbrüche: siehe Flow-Chart. |
| **14** | Aufnahme / Rekrutierung | |
| **14a** | Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung | Der erste Patient wurde im Februar 2015 in die Studie aufgenommen. Die doppelblinde Behandlungsphase war im Januar 2016 abgeschlossen (letzte Visite zu Woche 24).<br>Die Open Label Extension der Studie ist noch nicht abgeschlossen. |
| **14b** | Informationen, warum die Studie endete oder beendet wurde | Die Studie ist noch nicht abgeschlossen. Die doppelblinde Behandlungsphase endete planmäßig. |
| **a: nach CONSORT 2010.** | | |

*Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.*

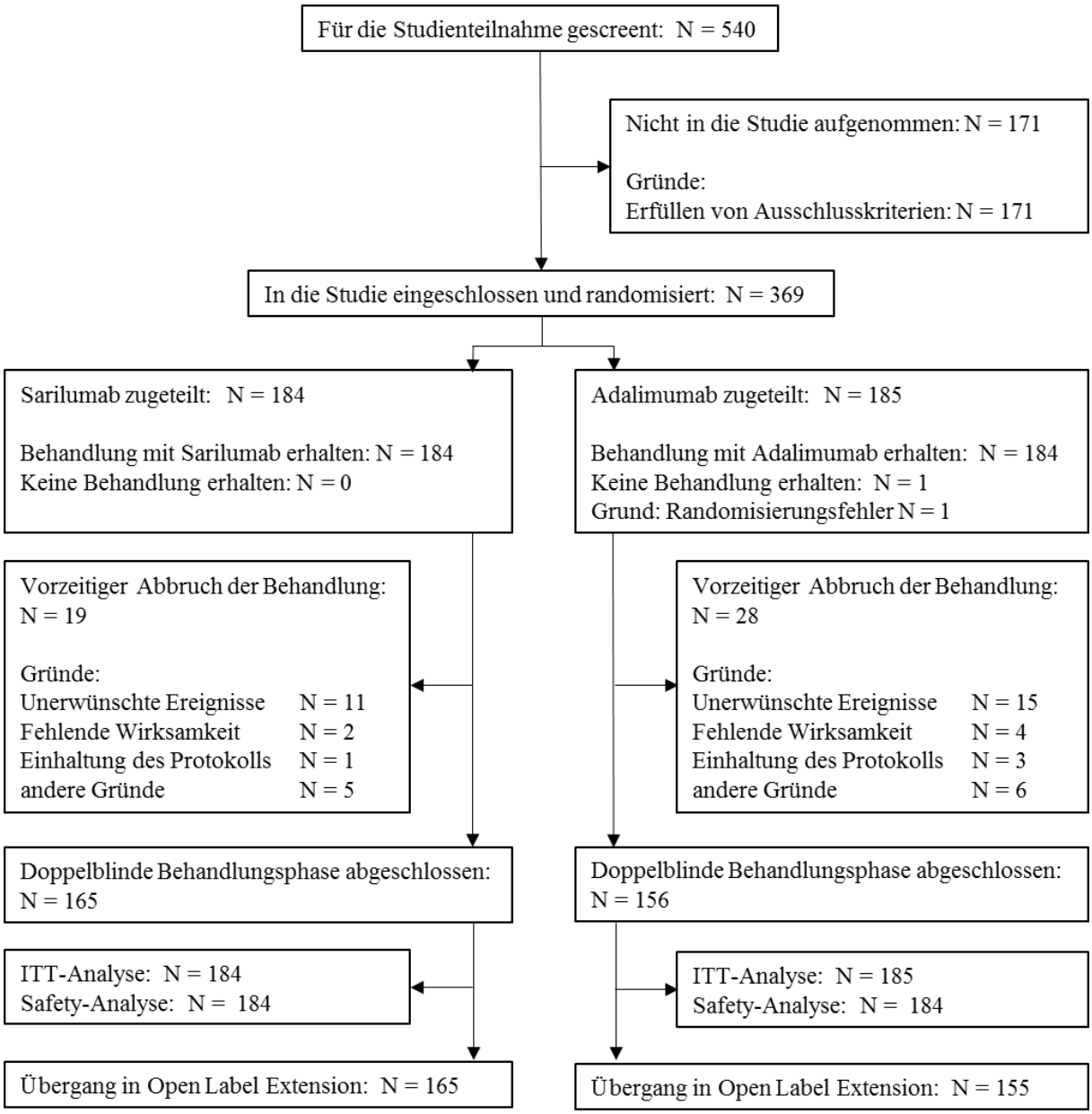



Abbildung 4-3: Patientenfluss der Studie MONARCH

**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 4-192 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MONARCH

**Studie: _____MONARCH_____**

**Tabelle: Liste der für die Bewertung herangezogenen Quellen**

| Genaue Benennung der Quelle | Kürzel |
|---|---|
| Burmester, G. (2016). Clinical Study Report: A randomized, double-blind, parallel-group study assessing the efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy in patients with rheumatoid arthritis; STUDY NUMBER: EFC14092; STUDY NAME: SARIL-RA-MONARCH. | |
| Sanofi Group (2015): Amended Clinical Trial Protocol 02: A randomized, double-blind, parallel-group study assessing the efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy in patients with rheuamtoid arthritis. Study numer: EFC14092. | |

**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:

Die Patienten wurden im Verhältnis 1:1 in die parallelen Behandlungsgruppen Sarilumab und Adalimumab stratifiziert nach Region randomisiert. Die Randomisierung erfolgte zentral durch ein interaktives Sprachdialogsystem (IVRS, Interactive Voice Response System).

**1.**

**<u>für randomisierte Studien:</u> Adäquate Erzeugung der Randomisierungssequenz**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

**<u>für nicht randomisierte Studien:</u> Zeitliche Parallelität der Gruppen**

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

**2.**

**<u>für randomisierte Studien:</u> Verdeckung der Gruppenzuteilung („allocation concealment“)**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

**<u>für nicht randomisierte Studien:</u> Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <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:

Die Patienten der Studie waren verblindet. Die Studienmedikation wurde in identisch aussehenden, vorgefüllten Glasspritzen bereitgestellt und identisch verblindet verpackt. Die Patienten und behandelnden Personen hatten bis zum Ende des randomisierten Behandlungszeitraums keinen Zugang zur Randomisierungsliste, außer ein medizinischer Notfall machte das Aufheben der Randomisierung notwendig.

**behandelnde bzw. weiterbehandelnde Personen:**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:

Die behandelnden Personen der Studie waren verblindet. Die Studienmedikation wurde in identisch aussehenden, vorgefüllten Glasspritzen bereitgestellt und identisch verblindet verpackt. Die Patienten und behandelnden Personen hatten bis zum Ende des randomisierten Behandlungszeitraums keinen Zugang zur Randomisierungsliste, außer ein medizinischer Notfall machte das Aufheben der Randomisierung notwendig.

**4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

**5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Bei der Studie MONARCH handelt es sich um eine randomisierte, aktiv kontrollierte, doppelblinde Phase-III-Studie. Sowohl die Erzeugung der Randomisierungssequenz als auch Verdeckung der Gruppenzuteilung wurde mittels IVRS realisiert. Patienten und behandelnde Personen waren verblindet. Die Erhebung und Auswertung der Endpunkte erfolgte verblindet. Es handelt sich um eine ergebnisunabhängige Berichterstattung. Änderungen des Studienprotokolls, die nach Studienbeginn vorgenommen wurden, beinhalten keine relevanten Parameter für das Verzerrungspotential.
Abschließend wird das Verzerrungspotential der Studie MONARCH als niedrig eingestuft.

**B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:**

**Endpunkt: Änderung des DAS28-ESR**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Änderung des CDAI**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Änderung des SDAI**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: DAS28-Remission**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: CDAI-Remission**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: SDAI-Remission**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: ACR/EULAR-Remission**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Niedrige Krankheitsaktivität gemäß DAS28**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Niedrige Krankheitsaktivität gemäß CDAI**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Niedrige Krankheitsaktivität gemäß SDAI**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Ansprechen gemäß EULAR-Kriterien**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Ansprechen gemäß ACR-Kriterien (ACR20/50/70)**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Gelenksymptome: Druckschmerzhafte Gelenke**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Gelenksymptome: Geschwollene Gelenke**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Schmerzen**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Globale Einschätzung des Arztes (PhGA)**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Patienten mit fehlenden Werten wurden nicht berücksichtigt.

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie und es erfolgte keine ergebnisgesteuerte Berichterstattung. Bei der Auswertung des Endpunkts wurden jedoch Patienten mit fehlenden Werten nicht berücksichtigt, da es sich nicht um einen patientenrelevanten Endpunkt handelt. Entsprechend wurde der Endpunkt nur rein deskriptiv dargestellt.

**Endpunkt: Globale Einschätzung des Patienten**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Funktionsstatus - HAQ-DI**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Laborparameter CRP + ESR**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Patienten mit fehlenden Werten wurden nicht berücksichtigt.

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Es handelt sich um eine doppelblinde Studie und es erfolgte keine ergebnisgesteuerte Berichterstattung. Bei der Auswertung des Endpunkts wurden jedoch Patienten mit fehlenden Werten nicht berücksichtigt, da es sich um Surrogatparameter und damit nicht um einen patientenrelevanten Endpunkt handelt. Entsprechend wurde der Endpunkt nur rein deskriptiv dargestellt.

**Endpunkt: Funktionsstatus - Morgensteifigkeit**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Fatigue**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Änderung des allgemeinen Gesundheitszustands - EQ-5D**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Änderung der Lebensqualität - SF-36**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Änderung der Lebensqualität - RAID**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

**Endpunkt: Sicherheit/Verträglichkeit**

**1. Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Sowohl Patienten als auch Personen, die mit der Behandlung des Patienten und der Erhebung der Zielkriterien betraut waren, waren verblindet.

**2. Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Zur Auswertung des Endpunkts Sicherheit/Verträglichkeit wurde die Safety-Population herangezogen. Diese umfasste alle randomisierten Patienten, die mindestens eine Dosis oder Teildosis der Studienintervention erhalten haben. Die Safety-Population ist mit der ITT-Population identisch.

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Für die Analyse des Endpunktes ergeben sich keine Hinweise für ein mögliches Verzerrungspotential. Es handelt sich um eine doppelblinde Studie, es erfolgte keine ergebnisgesteuerte Berichterstattung und das ITT-Prinzip wurde adäquat umgesetzt.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

**Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)**

*Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).*

**A Verzerrungsaspekte auf Studienebene:**

**Einstufung als randomisierte Studie**

☐ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein**: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.
→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

---

---

**1.**

**für randomisierte Studien:**

**Adäquate Erzeugung der Randomisierungssequenz**

☐ **ja**: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

☐ **unklar**: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

☐ **nein**: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**für nicht randomisierte Studien:**

**Zeitliche Parallelität der Gruppen**

☐ **ja**: Die Gruppen wurden zeitlich parallel verfolgt.

☐ **unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

☐ **nein**: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

**2.**

**für randomisierte Studien:**

**Verdeckung der Gruppenzuteilung („allocation concealment“)**

☐ **ja**: Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

☐ **unklar**: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

☐ **nein**: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**für nicht randomisierte Studien:**

**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**

☐ **ja**: Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

☐ **unklar**: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

☐ **nein**: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**3. Verblindung von Patienten und behandelnden Personen**

**Patient:**

☐ **ja:** Die Patienten waren verblindet.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

---

**behandelnde bzw. weiterbehandelnde Personen:**

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

---

**4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte**

*Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).*

*Beispiele zu a und b:*
- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

*Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter*
- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

*Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:*
- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
  *Zulässige Gründe sind:*
  - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
  - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
  - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche" Endpunkte nicht berichtet sind.*

*Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.*
*Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.*

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können**

*z. B.*
- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
  - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
  - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
  - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
  - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

☐ **ja**

☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

---

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):**

*Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.*

☐ **niedrig**: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

☐ **hoch**: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

---

---

**B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:**

*Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).*

**Endpunkt: _____**

**1. Verblindung der Endpunkterheber**

*Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.*
*In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.*

☐ **ja:** Der Endpunkt wurde verblindet erhoben.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

---

**2. Adäquate Umsetzung des ITT-Prinzips**

*Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set"). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.*

☐ **ja:** Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

☐ **unklar**: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

☐ **nein**: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

*Beachte die Hinweise zu Punkt A.4!*

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

*z. B.*
- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

☐ **ja**

☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

*Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.*

☐ **niedrig**: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

☐ **hoch**: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung: