# Dossier zur Nutzenbewertung gemäß § 35a SGB V

Tofacitinib (XELJANZ®)

Pfizer Pharma GmbH als örtlicher Vertreter des Zulassungsinhabers Pfizer Limited

# Modul 4 A

Behandlung der mittelschweren bis schweren aktiven rheumatoiden Arthritis bei erwachsenen Patienten, die unzureichend auf eine oder mehrere csDMARD-Behandlungen ansprechen oder diese nicht vertragen

Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen

# Inhaltsverzeichnis

S	Seite
Fabellenverzeichnis	6
Abbildungsverzeichnis	16
Abkürzungsverzeichnis	20
4 Modul 4 – allgemeine Informationen	22
4.1 Zusammenfassung der Inhalte von Modul 4	23
4.2 Methodik	
4.2.1 Fragestellung	42
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	45
4.2.3 Informationsbeschaffung	47
4.2.3.1 Studien des pharmazeutischen Unternehmers	48
4.2.3.2 Bibliografische Literaturrecherche	48
4.2.3.3 Suche in Studienregistern	49
4.2.3.4 Selektion relevanter Studien	50
4.2.4 Bewertung der Aussagekraft der Nachweise	51
4.2.5 Informationssynthese und -analyse	
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen	
Studien	52
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien	53
4.2.5.3 Meta-Analysen	
4.2.5.4 Sensitivitätsanalysen	
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren	64
4.2.5.6 Indirekte Vergleiche	
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	70
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden	
Arzneimittel	70
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden	
Arzneimittel	70
4.3.1.1.1 Studien des pharmazeutischen Unternehmers	
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche	
4.3.1.1.3 Studien aus der Suche in Studienregistern	
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden	
Arzneimittel	78
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT	
mit dem zu bewertenden Arzneimittel	79
4.3.1.2.1 Studiendesign und Studienpopulationen	
4.3.1.2.2 Verzerrungspotenzial auf Studienebene	
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien	
4.3.1.3.1 DAS28-4(ESR) nach 12 Monaten – RCT	
4.3.1.3.2 DAS28-4(CRP) nach 12 Monaten – RCT	101
4.3.1.3.3 EULAR-Response nach 12 Monaten – RCT	
4.3.1.3.4 ACR/EULAR-Remission und ACR Response nach 12 Monaten –	, ,
RCT	122

4.3.1.3.5	SDAI nach 12 Monaten – RCT	132
4.3.1.3.6	CDAI nach 12 Monaten – RCT	139
4.3.1.3.7	Schmerzbeurteilung nach 12 Monaten durch den Patienten – RCT	146
4.3.1.3.8	Gesamtbeurteilung der Krankheitsaktivität nach 12 Monaten durch	
	den Patienten (PatGA) – RCT	151
4.3.1.3.9	Gesamtbeurteilung der Krankheitsaktivität nach 12 Monaten durch	
	den Arzt (PhysGA) – RCT	156
4.3.1.3.10	Anzahl empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten	
	- RCT	161
4.3.1.3.11	Anzahl geschwollener Gelenke nach 12 Monaten – RCT	
4.3.1.3.12	SF-36 nach 12 Monaten – RCT	175
	FACIT-F nach 12 Monaten – RCT	
	HAQ-DI nach 12 Monaten – RCT	
	WLQ nach 12 Monaten – RCT	
	MOS Schlafscore nach 12 Monaten – RCT	
	Unerwünschte Ereignisse bis 12 Monate – RCT	
	Schwerwiegende unerwünschte Ereignisse bis 12 Monate – RCT	
	Unerwünschte Ereignisse, die bis Monat 12 zum Therapieabbruch	
1.3.1.3.17	geführt hatten– RCT	226
431320	Unerwünschte Ereignisse von besonderem Interesse: Schwere	220
1.3.1.3.20	Infektionen bis 12 Monate – RCT	229
431321	Unerwünschte Ereignisse von besonderem Interesse: Jedwede	,
7.3.1.3.21	Infektion und parasitäre Erkrankung bis 12 Monate – RCT	232
131322	Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster	232
4.3.1.3.22	bis 12 Monate – RCT	235
131323	Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose bis	233
7.5.1.5.25	<u> </u>	238
131321	Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom	236
4.5.1.5.24	Hautkrebs bis 12 Monate – RCT	2/1
121225	Unerwünschte Ereignisse von besonderem Interesse:	2 <del>4</del> 1
4.3.1.3.23	Gastrointestinale Perforation bis 12 Monate – RCT	244
121226		244
4.3.1.3.20	Unerwünschte Ereignisse von besonderem Interesse: Malignität	247
421227	(ohne Nicht-Melanom Hautkrebs) bis 12 Monate – RCT	247
4.3.1.3.27	Unerwünschte Ereignisse von besonderem Interesse: Erhöhte	250
121220	Leberwerte bis 12 Monate – RCT	230
4.3.1.3.28	Unerwünschte Ereignisse von besonderem Interesse: Erhöhte	252
421220	Kreatininwerte im Blut bis 12 Monate – RCT	255
4.3.1.3.29	Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis	250
4 2 1 2 20	12 Monate – RCT	256
4.3.1.3.30	Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie	250
4 2 1 2 21	bis 12 Monate – RCT	259
4.3.1.3.31	Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12	2.52
10100	Monate – RCT	
	Subgruppenanalysen – RCT	265
4.3.1.3	.32.1 Subgruppenanalysen für Endpunkt DAS28-4(ESR) nach 12	2
4015	Monaten – RCT	266
4.3.1.3	.32.2 Subgruppenanalysen für Endpunkt EULAR-Response (ESR)	•
	nach 12 Monaten – RCT	269

4.3.1.3.32.3 Subgruppenanalysen für Endpunkt ACR/EULAR-Remission	
und ACR Response nach 12 Monaten – RCT	. 271
4.3.1.3.32.4 Subgruppenanalysen für Endpunkt SDAI nach 12 Monaten – RCT	. 273
4.3.1.3.32.5 Subgruppenanalysen für Endpunkt CDAI nach 12 Monaten –	
RCT	. 275
4.3.1.3.32.6 Subgruppenanalysen für Endpunkt Schmerzbeurteilung durch	
den Patienten nach 12 Monaten – RCT	. 277
4.3.1.3.32.7 Subgruppenanalysen für Endpunkt Gesamtbeurteilung der	
Krankheitsaktivität durch den Patienten (PatGA) nach 12	
Monaten – RCT	. 280
4.3.1.3.32.8 Subgruppenanalysen für Endpunkt Gesamtbeurteilung der	
Krankheitsaktivität durch den Arzt (PhysGA) nach 12 Monaten	
– RCT	. 283
4.3.1.3.32.9 Subgruppenanalysen für Endpunkt Anzahl	
empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten –	
RCT	. 286
4.3.1.3.32.10 Subgruppenanalysen für Endpunkt Anzahl geschwollener	
Gelenke nach 12 Monaten – RCT	
4.3.1.3.32.11 Subgruppenanalysen für Endpunkt SF-36 nach 12 Monaten –	
RCT	. 291
4.3.1.3.32.12 Subgruppenanalysen für Endpunkt FACIT-F nach 12	
Monaten – RCT	. 301
4.3.1.3.32.13 Subgruppenanalysen für Endpunkt HAQ-DI nach 12	
Monaten – RCT	. 303
4.3.1.3.32.14 Subgruppenanalysen für Endpunkt WLQ nach 12 Monaten –	
RCT	
4.3.1.3.32.15 Subgruppenanalysen für Endpunkt MOS Schlafscore nach 12	
Monaten – RCT	. 310
4.3.1.3.32.16 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
bis 12 Monate – RCT	. 314
4.3.1.3.32.17 Subgruppenanalysen für Endpunkt schwerwiegende	245
unerwünschte Ereignisse bis 12 Monate – RCT	.317
4.3.1.3.32.18 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse,	220
die bis Monat 12 zum Therapieabbruch geführt hatten – RCT	. 320
4.3.1.3.32.19 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Schwere Infektionen –	
RCT	. 322
4.3.1.3.32.20 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Jedwede Infektion	22.4
und parasitäre Erkrankung – RCT	. 324
4.3.1.3.32.21 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	207
von besonderem Interesse bis 12 Monate: Herpes Zoster – RCT	. 321
4.3.1.3.32.22 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	220
von besonderem Interesse bis 12 Monate: Tuberkulose – RCT	. 329
4.3.1.3.32.23 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Nicht-Melanom	220
Hautkrebs – RCT	. 329

4.3.1.3.32.24 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Gastrointestinale	
Perforation – RCT	. 329
4.3.1.3.32.25 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Malignität (ohne	
Nicht-Melanom Hautkrebs) – RCT	. 329
4.3.1.3.32.26 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Erhöhte Leberwerte –	
RCT	
4.3.1.3.32.27 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Erhöhte	
Kreatininwerte im Blut – RCT	. 332
4.3.1.3.32.28 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Neutropenie – RCT	. 332
4.3.1.3.32.29 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Lymphopenie – RCT.	. 332
4.3.1.3.32.30 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse	
von besonderem Interesse bis 12 Monate: Anämie – RCT	. 332
4.3.1.3.33 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten	
Studien	. 333
4.3.2 Weitere Unterlagen	
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte	
Vergleiche	. 343
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche	
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen	
4.3.2.1.3.1 <endpunkt xxx=""> – indirekte Vergleiche aus RCT</endpunkt>	
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT	
4.3.2.2 Nicht randomisierte vergleichende Studien	
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte	
vergleichende Studien	. 347
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien	
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien	
4.3.2.2.3.1 <endpunkt xxx=""> – nicht randomisierte vergleichende Studien</endpunkt>	
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende	
Studien	. 349
4.3.2.3 Weitere Untersuchungen	
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen	
4.3.2.3.2 Charakteristika der weiteren Untersuchungen	
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen	
4.3.2.3.3.1 <endpunkt xxx=""> – weitere Untersuchungen</endpunkt>	
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen	
4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen	
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens	
4.4.1 Beurteilung der Aussagekraft der Nachweise	
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit	
und Ausmaß	. 354

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer	
Zusatznutzen besteht	. 365
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	
4.5.1 Begründung für die Vorlage indirekter Vergleiche	. 366
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und	
weiterer Untersuchungen	. 366
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da	
valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	. 366
4.5.4 Verwendung von Surrogatendpunkten	. 367
4.6 Liste der eingeschlossenen Studien	. 368
4.7 Referenzliste	. 369
Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche	. 373
Anhang 4-B: Suchstrategien – Suche in Studienregistern	. 380
Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente	
mit Ausschlussgrund (bibliografische Literaturrecherche)	. 382
Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in	
Studienregistern)	. 393
Anhang 4-E: Methodik der eingeschlossenen Studien – RCT	. 413
Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	. 427

# **Tabellenverzeichnis**

Sei	ite
Tabelle 4-1: Übersicht über Ein- und Ausschlusskriterien der zu berücksichtigenden Studien	27
Tabelle 4-2: Übersicht über Wahrscheinlichkeit und Ausmaß des Zusatznutzens aus RCT auf Endpunktebene – Population c)-Kombi	30
Tabelle 4-3: Übersicht über Wahrscheinlichkeit und Ausmaß des Zusatznutzens aus RCT auf Endpunktebene – Population d)-Kombi	35
Tabelle 4-4: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens	41
Tabelle 4-5: Übersicht über Ein- und Ausschlusskriterien der zu berücksichtigenden Studien - RCT	47
Tabelle 4-6: Definition der EULAR-Response	56
Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	71
Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel	74
Tabelle 4-9: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	77
Tabelle 4-10: Studienpool – RCT mit dem zu bewertenden Arzneimittel	78
Tabelle 4-11: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	80
Tabelle 4-12: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel	81
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel	84
Tabelle 4-14: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel	88
Tabelle 4-15: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	89
Tabelle 4-16: Operationalisierung von DAS28-4(ESR) nach 12 Monaten	92
Tabelle 4-17: Bewertung des Verzerrungspotenzials für DAS28-4(ESR) nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	93
Tabelle 4-18: Ergebnisse für DAS28-4(ESR) < 2,6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	94
Tabelle 4-19: Ergebnisse für DAS28-4(ESR) ≤ 3,2 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	97
Tabelle 4-20: Sensitivitätsanalysen für DAS28-4(ESR) < 2,6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	99

Tabelle 4-21: Operationalisierung von DAS28-4(CRP) nach 12 Monaten	101
Tabelle 4-22: Bewertung des Verzerrungspotenzials für DAS28-4(CRP) nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	102
Tabelle 4-23: Ergebnisse für DAS28-4(CRP) < 2,6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	103
Tabelle 4-24: Ergebnisse für DAS28-4(CRP) ≤ 3,2 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	105
Tabelle 4-25: Operationalisierung von EULAR-Response nach 12 Monaten	108
Tabelle 4-26: Bewertung des Verzerrungspotenzials für EULAR-Response nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	109
Tabelle 4-27: Ergebnisse für EULAR-Response (ESR) nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	111
Tabelle 4-28: Sensitivitätsanalysen für EULAR-Response (ESR) nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	115
Tabelle 4-29: Ergebnisse für EULAR-Response (CRP) nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	117
Tabelle 4-30: Operationalisierung von ACR Response und Remission	122
Tabelle 4-31: Bewertung des Verzerrungspotenzials für ACR/EULAR-Remission und ACR Response nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	123
Tabelle 4-32: Ergebnisse für ACR70 Response nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	124
Tabelle 4-33: Ergebnisse für ACR50 Response nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	126
Tabelle 4-34: Ergebnisse für ACR/EULAR-Remission nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	128
Tabelle 4-35: Sensitivitätsanalysen für ACR/EULAR-Remission nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	130
Tabelle 4-36: Operationalisierung von SDAI nach 12 Monaten	132
Tabelle 4-37: Bewertung des Verzerrungspotenzials für SDAI nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	133
Tabelle 4-38: Ergebnisse für SDAI ≤ 3,3 (Remission) nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	134
Tabelle 4-39: Ergebnisse für SDAI ≤ 11 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	136
Tabelle 4-40: Operationalisierung von CDAI nach 12 Monaten	
Tabelle 4-41: Bewertung des Verzerrungspotenzials für CDAI nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	
Tabelle 4-42: Ergebnisse für CDAI ≤ 2,8 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	141

Tabelle 4-43: Ergebnisse für CDAI ≤ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	143
Tabelle 4-44: Operationalisierung von Schmerzbeurteilung nach 12 Monaten durch den Patienten	146
Tabelle 4-45: Bewertung des Verzerrungspotenzials für Schmerzbeurteilung nach 12 Monaten durch den Patienten in RCT mit dem zu bewertenden Arzneimittel	146
Tabelle 4-46: Ergebnisse für Schmerzbeurteilung nach 12 Monaten durch den Patienten ≥ 10 aus RCT mit dem zu bewertenden Arzneimittel	148
Tabelle 4-47: Operationalisierung von PatGA nach 12 Monaten	151
Tabelle 4-48: Bewertung des Verzerrungspotenzials für PatGA nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	. 151
Tabelle 4-49: Ergebnisse für PatGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	. 153
Tabelle 4-50: Operationalisierung von PhysGA	156
Tabelle 4-51: Bewertung des Verzerrungspotenzials für PhysGA nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	. 156
Tabelle 4-52: Ergebnisse für PhysGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	. 158
Tabelle 4-53: Operationalisierung von Anzahl empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten	
Tabelle 4-54: Bewertung des Verzerrungspotenzials für Anzahl empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	162
Tabelle 4-55: Ergebnisse für Anzahl empfindlicher/druckschmerzhafter Gelenke (28) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	. 163
Tabelle 4-56: Ergebnisse für Anzahl empfindlicher/druckschmerzhafter Gelenke (68) $\leq$ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	. 165
Tabelle 4-57: Operationalisierung von Anzahl geschwollener Gelenke nach 12 Monaten	168
Tabelle 4-58: Bewertung des Verzerrungspotenzials für Anzahl geschwollener Gelenke nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	. 169
Tabelle 4-59: Ergebnisse für Anzahl geschwollener Gelenke (28) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	. 170
Tabelle 4-60: Ergebnisse für Anzahl geschwollener Gelenke (66) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	. 172
Tabelle 4-61: Operationalisierung von SF-36 nach 12 Monaten	175
Tabelle 4-62: Bewertung des Verzerrungspotenzials für SF-36 nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	. 176
Tabelle 4-63: Ergebnisse für SF-36 nach 12 Monaten – Responderanalysen der Komponentenscores aus RCT mit dem zu bewertenden Arzneimittel	. 177
Tabelle 4-64: Sensitivitätsanalysen für SF-36 nach 12 Monaten - Komponentenscores aus RCT mit dem zu bewertenden Arzneimittel	

Tabelle 4-65: Ergebnisse für SF-36 nach 12 Monaten – Responderanalysen der Einzeldimensionen aus RCT mit dem zu bewertenden Arzneimittel	183
Tabelle 4-66: Ergebnisse für SF-36 nach 12 Monaten – mittlere Änderungen aus RCT mit dem zu bewertenden Arzneimittel	185
Tabelle 4-67: Operationalisierung von FACIT-F nach 12 Monaten	189
Tabelle 4-68: Bewertung des Verzerrungspotenzials für FACIT-F nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	189
Tabelle 4-69: Ergebnisse für FACIT-F ≥ 4 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	191
Tabelle 4-70: Sensitivitätsanalysen für FACIT-F ≥ 4 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	193
Tabelle 4-71: Operationalisierung von HAQ-DI nach 12 Monaten	195
Tabelle 4-72: Bewertung des Verzerrungspotenzials für HAQ-DI nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	196
Tabelle 4-73: Ergebnisse für HAQ-DI ≥ 0,22 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	197
Tabelle 4-74: Ergebnisse für HAQ-DI ≥ 0,3 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	199
Tabelle 4-75: Ergebnisse für HAQ-DI ≥ 0,5 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	201
Tabelle 4-76: Sensitivitätsanalysen für HAQ-DI ≥ 0,22 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	203
Tabelle 4-77: Sensitivitätsanalysen für HAQ-DI ≥ 0,3 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	204
Tabelle 4-78: Sensitivitätsanalysen für HAQ-DI ≥ 0,5 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	205
Tabelle 4-79: Operationalisierung von WLQ nach 12 Monaten	207
Tabelle 4-80: Bewertung des Verzerrungspotenzials für WLQ nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel	208
Tabelle 4-81: Ergebnisse für WLQ nach 12 Monaten – Responderanalysen aus RCT mit dem zu bewertenden Arzneimittel	209
Tabelle 4-82: Ergebnisse für WLQ nach 12 Monaten – mittlere Änderungen aus RCT mit dem zu bewertenden Arzneimittel	211
Tabelle 4-83: Operationalisierung von MOS Schlafscore nach 12 Monaten	214
Tabelle 4-84: Bewertung des Verzerrungspotenzials für MOS Schlafscore nach 12  Monaten in RCT mit dem zu bewertenden Arzneimittel	
Tabelle 4-85: Ergebnisse für MOS Schlafscore ≥ 6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	216
Tabelle 4-86: Sensitivitätsanalysen für MOS Schlafscore ≥ 6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel	

20
21
23
24
25
26
26
27
29
29
30
32
32
33
35
35
36
38
2 2 2 2 2 3 3 3 3

Tabelle 4-106: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	. 238
Tabelle 4-107: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	. 239
Tabelle 4-108: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs bis 12 Monate"	. 241
Tabelle 4-109: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	. 241
Tabelle 4-110: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	. 242
Tabelle 4-111: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation bis 12 Monate"	. 244
Tabelle 4-112: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	. 244
Tabelle 4-113: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	. 245
Tabelle 4-114: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs) bis 12 Monate"	. 247
Tabelle 4-115: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs) bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	. 247
Tabelle 4-116: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs) bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	. 248
Tabelle 4-117: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte bis 12 Monate"	. 250
Tabelle 4-118: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	. 250
Tabelle 4-119: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	. 251
Tabelle 4-120: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut bis 12 Monate"	. 253
Tabelle 4-121: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	. 253

Tabelle 4-122: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	254
Tabelle 4-123: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis 12 Monate"	256
Tabelle 4-124: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	257
Tabelle 4-125: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	258
Tabelle 4-126: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie bis 12 Monate"	259
Tabelle 4-127: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	260
Tabelle 4-128: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	261
Tabelle 4-129: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12 Monate"	262
Tabelle 4-130: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel	262
Tabelle 4-131: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel	263
Tabelle 4-132: p-Werte der Interaktionstests für DAS28-4(ESR) nach 12 Monaten	266
Tabelle 4-133: Subgruppenergebnisse für DAS28-4(ESR) < 2,6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad	268
Tabelle 4-134: p-Werte der Interaktionstests für EULAR-Response (ESR) nach 12 Monaten	269
Tabelle 4-135: p-Werte der Interaktionstests für ACR/EULAR-Remission und ACR Response nach 12 Monaten	271
Tabelle 4-136: Subgruppenergebnisse für ACR70 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region	272
Tabelle 4-137: p-Werte der Interaktionstests für SDAI nach 12 Monaten	273
Tabelle 4-138: p-Werte der Interaktionstests für CDAI nach 12 Monaten	275
Tabelle 4-139: Subgruppenergebnisse für CDAI ≤ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region	276
Tabelle 4-140: p-Werte der Interaktionstests für Schmerzbeurteilung durch den Patienten ≥ 10 nach 12 Monaten	277
Tabelle 4-141: Subgruppenergebnisse für Schmerzbeurteilung durch den Patienten ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter	278

Tabelle 4-142: Subgruppenergebnisse für Schmerzbeurteilung durch den Patienten ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region	279
Tabelle 4-143: p-Werte der Interaktionstests für PatGA ≥ 10 nach 12 Monaten2	280
Tabelle 4-144: Subgruppenergebnisse für PatGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht	281
Tabelle 4-145: Subgruppenergebnisse für PatGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region	282
Tabelle 4-146: p-Werte der Interaktionstests für PhysGA ≥ 10 nach 12 Monaten2	283
Tabelle 4-147: Subgruppenergebnisse für PhysGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht	284
Tabelle 4-148: Subgruppenergebnisse für PhysGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region	285
Tabelle 4-149: p-Werte der Interaktionstests für Anzahl empfindlicher/druckschmerzhafter Gelenke ≤ 1 nach 12 Monaten	286
Tabelle 4-150: p-Werte der Interaktionstests für Anzahl geschwollener Gelenke ≤ 1 nach 12 Monaten	288
Tabelle 4-151: Subgruppenergebnisse für Anzahl geschwollener Gelenke (66) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter	289
Tabelle 4-152: Subgruppenergebnisse für Anzahl geschwollener Gelenke (28) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region	290
Tabelle 4-153: p-Werte der Interaktionstests für SF-36 Responderanalysen der Komponentenscores nach 12 Monaten (Verbesserung jeweils ≥ 2,5)	291
Tabelle 4-154: Subgruppenergebnisse für SF-36 Körperlicher Komponentenscore Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel:  Alter	292
Tabelle 4-155: p-Werte der Interaktionstests für SF-36 Responderanalysen der Einzeldimensionen nach 12 Monaten	294
Tabelle 4-156: Subgruppenergebnisse für SF-36 Einzeldimensionen Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht	296
Tabelle 4-157: Subgruppenergebnisse für SF-36 Einzeldimensionen Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter	297
Tabelle 4-158: Subgruppenergebnisse für SF-36 Einzeldimensionen Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region	298
Tabelle 4-159: Subgruppenergebnisse für SF-36 Einzeldimensionen Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad	300
Tabelle 4-160: p-Werte der Interaktionstests für FACIT-F Responderanalyse nach 12 Monaten (Verbesserung $\geq$ 4)	301
Tabelle 4-161: p-Werte der Interaktionstests für HAQ-DI nach 12 Monaten 3	303
Tabelle 4-162: Subgruppenergebnisse für HAQ-DI ≥ 0,3 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter	304

Tabelle 4-163: Subgruppenergebnisse für HAQ-DI ≥ 0,22 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad	305
Tabelle 4-164: p-Werte der Interaktionstests für WLQ nach 12 Monaten	307
Tabelle 4-165: Subgruppenergebnisse für WLQ Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad	309
Tabelle 4-166: p-Werte der Interaktionstests für MOS Schlafscore Responderanalyse (Verbesserung $\geq$ 6)	310
Tabelle 4-167: Subgruppenergebnisse für MOS Schlafscore Responderanalyse nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter	311
Tabelle 4-168: Subgruppenergebnisse für MOS Schlafscore Responderanalyse nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region	312
Tabelle 4-169: Subgruppenergebnisse für MOS Schlafscore Responderanalyse nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad	313
Tabelle 4-170: p-Werte der Interaktionstests für unerwünschte Ereignisse bis 12 Monate	314
Tabelle 4-171: Subgruppenergebnisse für unerwünschte Ereignisse bis 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht	315
Tabelle 4-172: p-Werte der Interaktionstests für schwerwiegende unerwünschte Ereignisse bis 12 Monate	317
Tabelle 4-173: Subgruppenergebnisse für schwerwiegende unerwünschte Ereignisse bis 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter	318
Tabelle 4-174: p-Werte der Interaktionstests für unerwünschte Ereignisse, die bis Monat 12 zum Therapieabbruch geführt hatten	320
Tabelle 4-175: p-Werte der Interaktionstests für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Schwere Infektionen	322
Tabelle 4-176: p-Werte der Interaktionstests für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Jedwede Infektion und parasitäre Erkrankung	324
Tabelle 4-177: Subgruppenergebnisse für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Jedwede Infektion und parasitäre Erkrankung aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht	325
Tabelle 4-178: p-Werte der Interaktionstests für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Herpes Zoster	
Tabelle 4-179: p-Werte der Interaktionstests für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Erhöhte Leberwerte	
Tabelle 4-180: Übersicht über die Ergebnisse aus RCT auf Endpunktebene nach 12 Monaten – Population c)-Kombi	334
Tabelle 4-181: Übersicht über die Ergebnisse aus RCT auf Endpunktebene nach 12 Monaten – Population d)-Kombi	338
Tabelle 4-182: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	344
Tabelle 4-183: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden	344

Tabelle 4-184: Operationalisierung von <endpunkt xxx=""></endpunkt>	345
Tabelle 4-185: Bewertung des Verzerrungspotenzials für <endpunkt xxx=""> in RCT für indirekte Vergleiche</endpunkt>	345
Tabelle 4-186: Ergebnisse für <endpunkt xxx=""> aus RCT für indirekte Vergleiche</endpunkt>	346
Tabelle 4-187: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	348
Tabelle 4-188: Operationalisierung von <endpunkt xxx=""></endpunkt>	348
Tabelle 4-189: Verzerrungsaspekte für < Endpunkt xxx> – nicht randomisierte vergleichende Studien	349
Tabelle 4-190: Operationalisierung von <endpunkt xxx=""> – weitere Untersuchungen</endpunkt>	351
Tabelle 4-191: Übersicht über Wahrscheinlichkeit und Ausmaß des Zusatznutzens aus RCT auf Endpunktebene – Population c)-Kombi	355
Tabelle 4-192: Übersicht über Wahrscheinlichkeit und Ausmaß des Zusatznutzens aus RCT auf Endpunktebene – Population d)-Kombi	360
Tabelle 4-193: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens	365
Tabelle 4-194 (Anhang): Studiendesign und -methodik für Studie A3921064 (ORAL STANDARD)	414
Tabelle 4-195 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie A3921064 (ORAL STANDARD)	428

# Abbildungsverzeichnis

Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel
Abbildung 2: Kaplan-Meier Kurven für den Endpunkt DAS28-4(ESR) < 2,6 nach 12 Monaten (Population c)-Kombi)
Abbildung 3: Kaplan-Meier Kurven für den Endpunkt DAS28-4(ESR) < 2,6 nach 12 Monaten (Population d)-Kombi)
Abbildung 4: Kaplan-Meier Kurven für den Endpunkt DAS28-4(ESR) $\leq$ 3,2 nach 12 Monaten (Population c)-Kombi)
Abbildung 5: Kaplan-Meier Kurven für den Endpunkt DAS28-4(ESR) ≤ 3,2 nach 12 Monaten (Population d)-Kombi)
Abbildung 6: Kaplan-Meier Kurven für den Endpunkt DAS28-4(CRP) < 2,6 nach 12 Monaten (Population c)-Kombi)
Abbildung 7: Kaplan-Meier Kurven für den Endpunkt DAS28-4(CRP) < 2,6 nach 12 Monaten (Population d)-Kombi)
Abbildung 8: Kaplan-Meier Kurven für den Endpunkt DAS28-4(CRP) $\leq$ 3,2 nach 12 Monaten (Population c)-Kombi)
Abbildung 9: Kaplan-Meier Kurven für den Endpunkt DAS28-4(CRP) $\leq$ 3,2 nach 12 Monaten (Population d)-Kombi)
Abbildung 10: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (ESR) nach 12 Monaten ("gute Verbesserung", Population c)-Kombi)
Abbildung 11: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (ESR) nach 12 Monaten ("gute Verbesserung", Population d)-Kombi)
Abbildung 12: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (ESR) nach 12 Monaten ("gute oder moderate Verbesserung", Population c)-Kombi)
Abbildung 13: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (ESR) nach 12 Monaten ("gute oder moderate Verbesserung", Population d)-Kombi)
Abbildung 14: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (CRP) nach 12 Monaten ("gute Verbesserung", Population c)-Kombi)
Abbildung 15: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (CRP) nach 12 Monaten ("gute Verbesserung", Population d)-Kombi)
Abbildung 16: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (CRP) nach 12 Monaten ("gute oder moderate Verbesserung", Population c)-Kombi)
Abbildung 17: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (CRP) nach 12 Monaten ("gute oder moderate Verbesserung", Population d)-Kombi)
Abbildung 18: Kaplan-Meier Kurven für den Endpunkt ACR70 Response nach 12 Monaten (Population c)-Kombi)
Abbildung 19: Kaplan-Meier Kurven für den Endpunkt ACR70 Response nach 12 Monaten (Population d)-Kombi)

Abbildung 20: Kaplan-Meier Kurven für den Endpunkt ACR50 Response nach 12 Monaten (Population c)-Kombi)	127
Abbildung 21: Kaplan-Meier Kurven für den Endpunkt ACR50 Response nach 12 Monaten (Population d)-Kombi)	127
Abbildung 22: Kaplan-Meier Kurven für den Endpunkt ACR/EULAR-Remission nach 12 Monaten (Population c)-Kombi)	129
Abbildung 23: Kaplan-Meier Kurven für den Endpunkt ACR/EULAR-Remission nach 12 Monaten (Population d)-Kombi)	129
Abbildung 24: Kaplan-Meier Kurven für den Endpunkt SDAI ≤ 3,3 nach 12 Monaten (Population c)-Kombi)	135
Abbildung 25: Kaplan-Meier Kurven für den Endpunkt SDAI ≤ 3,3 nach 12 Monaten (Population d)-Kombi)	135
Abbildung 26: Kaplan-Meier Kurven für den Endpunkt SDAI ≤ 11 nach 12 Monaten (Population c)-Kombi)	137
Abbildung 27: Kaplan-Meier Kurven für den Endpunkt SDAI ≤ 11 nach 12 Monaten (Population d)-Kombi)	137
Abbildung 28: Kaplan-Meier Kurven für den Endpunkt CDAI ≤ 2,8 nach 12 Monaten (Population c)-Kombi)	142
Abbildung 29: Kaplan-Meier Kurven für den Endpunkt CDAI ≤ 2,8 nach 12 Monaten (Population d)-Kombi)	142
Abbildung 30: Kaplan-Meier Kurven für den Endpunkt CDAI ≤ 10 nach 12 Monaten (Population c)-Kombi)	144
Abbildung 31: Kaplan-Meier Kurven für den Endpunkt CDAI ≤ 10 nach 12 Monaten (Population d)-Kombi)	144
Abbildung 32: Kaplan-Meier Kurven für den Endpunkt Schmerzbeurteilung nach 12 Monaten durch den Patienten ≥ 10 (Population c)-Kombi)	149
Abbildung 33: Kaplan-Meier Kurven für den Endpunkt Schmerzbeurteilung nach 12 Monaten durch den Patienten ≥ 10 (Population d)-Kombi)	149
Abbildung 34: Kaplan-Meier Kurven für den Endpunkt PatGA ≥ 10 nach 12 Monaten (Population c)-Kombi)	154
Abbildung 35: Kaplan-Meier Kurven für den Endpunkt PatGA ≥ 10 nach 12 Monaten (Population d)-Kombi)	154
Abbildung 36: Kaplan-Meier Kurven für den Endpunkt PhysGA ≥ 10 nach 12 Monaten (Population c)-Kombi)	159
Abbildung 37: Kaplan-Meier Kurven für den Endpunkt PhysGA ≥ 10 nach 12 Monaten (Population d)-Kombi)	159
Abbildung 38: Kaplan-Meier Kurven für den Endpunkt empfindlicher/druckschmerzhafter Gelenke (28) ≤ 1 nach 12 Monaten (Population c)-Kombi)	164

Abbildung 39: Kaplan-Meier Kurven für den Endpunkt empfindlicher/druckschmerzhafter Gelenke (28) ≤ 1 nach 12 Monaten (Population d)-Kombi)	164
Abbildung 40: Kaplan-Meier Kurven für den Endpunkt empfindlicher/druckschmerzhafter Gelenke (68) ≤ 1 nach 12 Monaten (Population c)-Kombi)	166
Abbildung 41: Kaplan-Meier Kurven für den Endpunkt empfindlicher/druckschmerzhafter Gelenke (68) ≤ 1 nach 12 Monaten (Population d)-Kombi)	166
Abbildung 42: Kaplan-Meier Kurven für den Endpunkt geschwollener Gelenke (28) ≤ 1 nach 12 Monaten (Population c)-Kombi)	171
Abbildung 43: Kaplan-Meier Kurven für den Endpunkt geschwollener Gelenke (28) ≤ 1 nach 12 Monaten (Population d)-Kombi)	171
Abbildung 44: Kaplan-Meier Kurven für den Endpunkt geschwollener Gelenke (66) ≤ 1 nach 12 Monaten (Population c)-Kombi)	173
Abbildung 45: Kaplan-Meier Kurven für den Endpunkt geschwollener Gelenke (66) ≤ 1 nach 12 Monaten (Population d)-Kombi)	173
Abbildung 46: Kaplan-Meier Kurven für den Endpunkt SF-36 Körperlicher Komponentenscore ≥ 2,5 nach 12 Monaten (Population c)-Kombi)	179
Abbildung 47: Kaplan-Meier Kurven für den Endpunkt SF-36 Körperlicher Komponentenscore ≥ 2,5 nach 12 Monaten (Population d)-Kombi)	179
Abbildung 48: Kaplan-Meier Kurven für den Endpunkt SF-36 Psychischer Komponentenscore ≥ 2,5 nach 12 Monaten (Population c)-Kombi)	180
Abbildung 49: Kaplan-Meier Kurven für den Endpunkt SF-36 Psychischer Komponentenscore ≥ 2,5 nach 12 Monaten (Population d)-Kombi)	180
Abbildung 50: Kaplan-Meier Kurven für den Endpunkt FACIT-F ≥ 4 nach 12 Monaten (Population c)-Kombi)	192
Abbildung 51: Kaplan-Meier Kurven für den Endpunkt FACIT-F ≥ 4 nach 12 Monaten (Population d)-Kombi)	192
Abbildung 52: Kaplan-Meier Kurven für den Endpunkt HAQ-DI ≥ 0,22 nach 12 Monaten (Population c)-Kombi)	198
Abbildung 53: Kaplan-Meier Kurven für den Endpunkt HAQ-DI ≥ 0,22 nach 12 Monaten (Population d)-Kombi)	198
Abbildung 54: Kaplan-Meier Kurven für den Endpunkt HAQ-DI ≥ 0,3 nach 12 Monaten (Population c)-Kombi)	200
Abbildung 55: Kaplan-Meier Kurven für den Endpunkt HAQ-DI ≥ 0,3 nach 12 Monaten (Population d)-Kombi)	
Abbildung 56: Kaplan-Meier Kurven für den Endpunkt HAQ-DI ≥ 0,5 nach 12 Monaten (Population c)-Kombi)	
Abbildung 57: Kaplan-Meier Kurven für den Endpunkt HAQ-DI ≥ 0,5 nach 12 Monaten (Population d)-Kombi)	

Abbildung 58: Kaplan-Meier Kurven für den Endpunkt MOS Schlafscore ≥ 6 nach 12 Monaten (Population c)-Kombi)	217
Abbildung 59: Kaplan-Meier Kurven für den Endpunkt MOS Schlafscore ≥ 6 nach 12 Monaten (Population d)-Kombi)	217
Abbildung 60: Flussdiagramm der Studie A3921064 (ORAL STANDARD)	426

# Abkürzungsverzeichnis

Abkürzung	Bedeutung		
ACR	American College of Rheumatology		
ALT	Alanin-Aminotransferase		
AST	Aspartat-Aminotransferase		
bDMARD	biological DMARD		
BID	zweimal täglich (bis in die)		
csDMARD	conventional synthetic DMARD		
CDAI	Clinical Disease Activity Index		
CONSORT	Consolidated Standards of Reporting Trials		
CP-690,550	Bezeichnung für Tofacitinib vor der Vergabe dieses Freinamens (INN) für den Wirkstoff		
CRP	C-reaktives Protein		
DAS	Krankheitsaktivitätsscore (Disease Activity Score)		
DMARD	disease-modifying anti-rheumatic drugs		
DIMDI	Deutsches Institut für Medizinische Dokumentation		
EG	Europäische Gemeinschaft		
EKG	Elektrokardiogramm		
EQ-5D	Lebensqualitätsinstrument entwickelt von der European Quality of Life (EuroQoL) Group		
ESR	Erythrozytensedimentationsrate (Erythrocyte Sedimentation Rate)		
FACIT	Functional Assessment of Chronic Illness Therapy		
FACIT-F	Functional Assessment of Chronic Illness Therapy Fatigue Scale		
FAS	Full Analysis Set		
GCP	Good Clinical Practice		
HAQ-DI	Health Assessment Questionnaire-Disability Index		
INN	Internationaler Freiname (international nonproprietary name)		
ITT	Intention to treat		
IVRS	Interactive Voice Response System		
LOCF	Last Observation Carried Forward		
MedDRA	Medical Dictionary for Regulatory Activities		
MOS	Medical Outcomes Study		
MTC	Mixed Treatment Comparison		

MTX	Methotrexat
NRI	Nonresponder Imputation
PatGA	Patient Global Assessment
PhysGA	Physician Global Assessment
RA	Rheumatoide Arthritis
RCT	Randomized Controlled Trial
SAP	Statistischer Analyseplan
SC	subkutan (subcutaneous)
SDAI	Simplified Disease Activity Index
SF-36	Short Form-36
SGB	Sozialgesetzbuch
SOC	System Organ Class
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes Unerwünschtes Ereignis
TNFi	Tumornekrosefaktor-Inhibitor (tumor necrosis factor inhibitor)
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschtes Ereignis
WHO	World Health Organization
WLQ	Work Limitations Questionnaire
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**

Die Fragestellung für das vorliegende Nutzendossier zu Tofacitinib in der Indikation für erwachsene Patienten mit moderater bis schwerer rheumatoiden Arthritis (RA) betrifft die Darlegung der verfügbaren aktuellen Daten- und Erkenntnislage zur Nutzenbewertung von Tofacitinib. Für das Anwendungsgebiet A wird die Fragestellung im hier vorliegenden Modul 4A für eine spezifische Teilpopulation im Rahmen der Indikation von Tofacitinib behandelt:

Welcher Zusatznutzen besteht für die Behandlung mit Tofacitinib in Kombination mit MTX (ggf. auch als Monotherapie bei MTX-Unverträglichkeit) im Vergleich zur zweckmäßigen Vergleichstherapie bei erwachsenen Patienten mit moderater bis schwerer rheumatoider Arthritis, die unzureichend auf eine oder mehrere csDMARD-Behandlungen ansprechen oder diese nicht vertragen?

#### **Patientenpopulation**

Das hier beschriebene Anwendungsgebiet A umfasst die Patienten im Rahmen der Zulassung von Tofacitinib, die unzureichend auf eine vorhergehende Therapie mit mindestens einem klassischen synthetischen krankheitsmodifizierenden antirheumatischen Arzneimittel (csDMARD) angesprochen oder diese nicht vertragen haben. Im Anwendungsgebiet A dürfen die Patienten im Indikationsgebiet noch keine Therapie mit (mindestens) einem biologisch hergestellten krankheitsmodifizierenden antirheumatischen Arzneimittel (bDMARD) erhalten haben. Entsprechend nationaler und internationaler Leitlinien kann das Anwendungsgebiet A auch durch die Bezeichnung "Zweitlinientherapie" der RA beschrieben werden. Das Anwendungsgebiet A ist eine Teilpopulation innerhalb der Indikation für Tofacitinib.

Eine zweite Teilpopulation der Indikation von Tofacitinib (Anwendungsgebiet B: Patienten, die bereits auf mindestens einem bDMARD unzureichend angesprochen haben) wird entsprechend in Modul 4B dargestellt.

Die Bezeichnung der relevanten Patienten-Teilpopulationen in diesem Modul 4A übernimmt als vereinfachende Kurzform die Kennzeichnung des G-BA aus der Niederschrift zur Beratung Tofacitinib (1). Demzufolge werden in diesem Modul 4A die folgenden drei Teilpopulationen adressiert:

- **Teilpopulation b**) Patienten ohne ungünstige Prognosefaktoren nach unzureichendem Ansprechen auf ein csDMARD
- **Teilpopulation c**) Patienten mit ungünstigen Prognosefaktoren nach unzureichendem Ansprechen auf ein csDMARD

• **Teilpopulation d**) - Patienten nach unzureichendem Ansprechen auf mehr als ein csDMARD.

Als Patienten "mit ungünstigen Prognosefaktoren" werden diejenigen Patienten eingestuft, bei denen entweder (zumindest) Gelenkerosionen vorliegen, die (zumindest) bezüglich der Rheumafaktoren und/oder Antikörper gegen citrullinierte Peptide positiv getestet wurden, oder die (zumindest) eine hohe Krankheitsaktivität (d.h. DAS28 > 5,1) aufweisen.

Die Teilpopulation a) (nach der Bennenung in der Niederschrift des G-BA) umfasst Patienten, die bislang keine krankheitsmodifizierende Therapie erhalten haben. Da diese Patienten nicht durch die Indikation von Tofacitinib abgedeckt sind, wird die Teilpopulation a) nicht in diesem Nutzendossier adressiert.

#### Intervention

Die Intervention besteht aus der Gabe von 5mg BID Tofacitinib (in Kombination mit MTX oder ggfs. als Monotherapie) entsprechend der Fachinformation (2).

#### Vergleichstherapie

Für Patienten ohne ungünstige Prognosefaktoren, die unzureichend auf eine vorangegangene Behandlung mit einem csDMARD ansprachen (Kurzbezeichnung: Population b)), werden csDMARDs als Mono- oder Kombinationstherapie als zweckmäßige Vergleichstherapie betrachtet (Population "b)-Mono" bzw. "b)-Kombi").

Für Patienten mit ungünstigen Prognosefaktoren, die unzureichend auf ein csDMARD ansprachen (Population c)) bzw. für Patienten, die unzureichend auf mehrere csDMARDs ansprachen (Population d)), wurden bDMARDs in Kombination mit MTX als Vergleichstherapie berücksichtigt (Populationen "c)-Kombi" bzw. "d)-Kombi"). Bei MTX-Unverträglichkeit kann die Vergleichstherapie für die Populationen c) und d) ggf. auch eine bDMARD-Monotherapie sein (Populationen "c)-Mono" bzw. "d)-Mono").

#### **Endpunkte**

- Morbidität:
  - o DAS28-4
  - o EULAR-Response
  - o ACR/EULAR-Remission und ACR Response
  - o SDAI (Simplified Disease Activity Index)
  - o CDAI (Clinical Disease Activity Index)
  - o Schmerzbeurteilung durch den Patienten
  - o Gesamtbeurteilung der Krankheitsaktivität durch den Patienten
  - o Gesamtbeurteilung der Krankheitsaktivität durch den Arzt

- o Anzahl empfindlicher/druckschmerzhafter bzw. geschwollener Gelenke
- o FACIT-F
- o HAQ-DI
- o WLQ
- o MOS Schlafscore
- Gesundheitsbezogene Lebensqualität:
  - o SF-36
- Sicherheit/Verträglichkeit:
  - Unerwünschte Ereignisse
  - Schwerwiegende unerwünschte Ereignisse
  - Therapieabbruch wegen unerwünschter Ereignisse
  - Unerwünschte Ereignisse von besonderem Interesse:
    - o Schwere Infektionen
    - o Jedwede Infektion und parasitäre Erkrankung
    - Herpes Zoster
    - o Tuberkulose
    - Nicht-Melanom Hautkrebs
    - o Gastrointestinale Perforation
    - o Malignität (ohne Nicht-Melanom Hautkrebs)
    - o Erhöhte Leberwerte
    - Erhöhte Kreatininwerte im Blut
    - o Neutropenie
    - o Lymphopenie
    - Anämie

Die Auswahl der o.g. unerwünschten Ereignisse von besonderem Interesse erfolgte anhand der Angaben in der Fachinformation und des Risk Management Plans des zu bewertenden Arzneimittels Tofacitinib bzw. den jeweiligen Fachinformationen der zweckmäßigen Vergleichstherapie.

#### **Studientyp**

Randomisierte kontrollierte Studie (RCT).

#### **Datenquellen**

Zur Nutzenbewertung wurden Studien herangezogen, zu denen Ergebnisse in Form einer Originalpublikation, eines ausführlichen Ergebnisberichts aus einem Studienregistereintrag oder eines Studienberichts vorlagen. Ausgeschlossen wurden Sekundärpublikationen, die keine neuen Daten (über die ohnehin in den Studienberichten/Originalpublikationen dargestellten) enthalten.

Aus der systematischen Literatur- und Studienregisterrecherche sowie den Studien des pharmazeutischen Unternehmers ergibt sich eine für die Nutzenbewertung relevante randomisierte kontrollierte Studie, die die Einschlusskriterien der Studienselektion erfüllt: Studie A3921064 (ORAL STANDARD).

### Ein-/Ausschlusskriterien für Studien

Tabelle 4-1: Übersicht über Ein- und Ausschlusskriterien der zu berücksichtigenden Studien

	Einschlusskriterien	E	Ausschlusskriterien	A
Patientenpopulation	Erwachsene Patienten mit moderater bis schwerer rheumatoider Arthritis nach unzureichendem Ansprechen auf mindestens ein csDMARD oder falls dieses nicht vertragen wurde	1	Erwachsene Patienten ohne moderate bis schwere rheumatoide Arthritis oder ohne bisheriges csDMARD-Versagen.	
Intervention	Tofacitinib 5mg BID (als Monotherapie oder in Kombination mit MTX)	2	Andere Interventionen	
Vergleichstherapie	Entspricht der zweckmäßigen Vergleichstherapie (ZVT) nach Festlegung des G-BA (1) für mindestens eine der Teilpopulationen b), c) oder d):  • b): csDMARDs als Mono- oder Kombinationstherapie • c) und d): bDMARDs in Kombination mit MTX bzw. als Monotherapie bei MTX-Unverträglichkeit	3	Entspricht für keine der Teilpopulation b), c) oder d) der Festlegung der ZVT	
Endpunkte	Ergebnisse zu mind. einem patientenrelevanten Endpunkt bzgl.  • Mortalität  • Morbidität (Symptomatik)  • Lebensqualität  • Sicherheit/Verträglichkeit sind berichtet.	4	Ergebnisse zu keinem Endpunkt bzgl.  • Mortalität • Morbidität (Symptomatik) • Lebensqualität • Sicherheit/Verträglichkeit sind berichtet.	
Studiendauer	≥ 24 Wochen (bzw. 6 Monate)	5	< 24 Wochen (bzw. 6 Monate)	
Studientyp	Randomisierte kontrollierte Studie	6	Keine randomisierte kontrollierte Studie	
Publikationstyp	Studienbericht, Primärpublikation oder Sekundärpublikation mit neuen Daten	7	Sekundärpublikation ohne neue Daten	
Studienstatus	Studie abgeschlossen	8	Studie läuft noch	8

# <u>Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von</u> Ergebnissen

#### Studienqualität:

Für die Darstellung des Zusatznutzens wurden die Ergebnisse von Subpopulationen einer randomisierten kontrollierten Studie (RCT) von hoher Qualität herangezogen. Die Studien ermöglicht einen direkten Vergleich der zu bewertenden Arzneimittels Tofacitinib mit der zweckmäßigen Vergleichstherapie. Bei der eingeschlossenen Studie handelt es sich um eine doppelblinde Studie mit niedrigem Verzerrungspotenzial.

#### Validität der herangezogenen Endpunkte:

Alle Endpunkte, die für die Darstellung der Wirksamkeit und Sicherheit im Rahmen der Nutzenbewertung von Tofacitinib herangezogen wurden, sind patientenrelevant. Bezüglich der Validität der einzelnen Endpunkte siehe Abschnitt 4.2.5.2 in diesem Modul 4A.

#### Evidenzstufe:

Der Nachweis für den Zusatznutzen beruht auf den Ergebnissen einer randomisierten kontrollierten Studie und entspricht somit dem Evidenzgrad Ib.

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

# Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

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

Tabelle 4-2: Übersicht über Wahrscheinlichkeit und Ausmaß des Zusatznutzens aus RCT auf Endpunktebene – Population c)-Kombi

Stand: 21.04.2017

Studie: A3921064 (ORAL STANDARD)	RR [95% KI], p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
DAS28-4(ESR) < 2,6	0,8 [0,3-1,7], p=0,484	Kein Zusatznutzen belegt
$DAS28-4(ESR) \le 3.2$	0,8 [0,4-1,4], p=0,370	Kein Zusatznutzen belegt
DAS28-4(CRP) < 2,6	1,1 [0,7-1,8], p=0,622	Kein Zusatznutzen belegt
$DAS28-4(CRP) \le 3,2$	1,1 [0,8-1,6], p=0,512	Kein Zusatznutzen belegt
EULAR-Response (ESR): gute Verbesserung	0,8 [0,4-1,4], p=0,370	Kein Zusatznutzen belegt
EULAR-Response (ESR): gute oder moderate Verbesserung	1,1 [0,9-1,4], p=0,366	Kein Zusatznutzen belegt
EULAR-Response (CRP): gute Verbesserung	1,1 [0,8-1,6], p=0,506	Kein Zusatznutzen belegt
EULAR-Response (CRP): gute oder moderate Verbesserung	1,0 [0,8-1,2], p=0,765	Kein Zusatznutzen belegt
ACR70	1,7 [0,9-3,3], p=0,084	Kein Zusatznutzen belegt
ACR50	1,4 [0,9-2,0], p=0,130	Kein Zusatznutzen belegt
ACR/EULAR-Remission	1,7 [0,5-5,5], p=0,402	Kein Zusatznutzen belegt
<b>SDAI ≤ 3,3</b>	1,9 [0,8-4,8], p=0,176	Kein Zusatznutzen belegt
SDAI ≤ 11	1,2 [0,8-1,6], p=0,407	Kein Zusatznutzen belegt
CDAI ≤ 2,8	1,6 [0,7-3,9], p=0,276	Kein Zusatznutzen belegt
CDAI ≤ 10	1,3 [0,9-1,8], p=0,190	Kein Zusatznutzen belegt
Schmerzbeurteilung durch den Patienten (Verbesserung ≥ 10)	0,9 [0,7-1,2], p=0,573	Kein Zusatznutzen belegt
Gesamtbeurteilung der Krankheitsaktivität durch den Patienten [PatGA] (Verbesserung ≥ 10)	0,9 [0,7-1,2], p=0,463	Kein Zusatznutzen belegt
Gesamtbeurteilung der Krankheitsaktivität durch den Arzt [PhysGA] (Verbesserung≥10)	1,0 [0,8-1,2], p=0,667	Kein Zusatznutzen belegt
Anzahl empfindlicher/druckschmerzhafter Gelenke ≤ 1 – 28 untersuchte Gelenke	1,0 [0,6-1,7], p=0,887	Kein Zusatznutzen belegt
Anzahl empfindlicher/druckschmerzhafter Gelenke ≤ 1 – 68 untersuchte Gelenke	0,9 [0,5-1,7], p=0,717	Kein Zusatznutzen belegt

Tofacitinib (XELJANZ®)

Seite 30 von 442

Anzahl geschwollener Gelenke ≤ 1 – 28 untersuchte Gelenke	1,0 [0,7-1,4], p=0,977	Kein Zusatznutzen belegt
Anzahl geschwollener Gelenke ≤ 1 – 66 untersuchte Gelenke	1,1 [0,8-1,7], p=0,574	Kein Zusatznutzen belegt
SF-36		
- Körperlicher Komponentenscore (Verbesserung ≥ 2,5)	1,0 [0,7-1,2], p=0,686	Kein Zusatznutzen belegt
- Psychischer Komponentenscore (Verbesserung ≥ 2,5)	0,8 [0,6-1,2], p=0,350	Kein Zusatznutzen belegt
- Vitalität (Verbesserung ≥ 5)	0,9 [0,7-1,3], p=0,527	Kein Zusatznutzen belegt
- Körperliche Funktion (Verbesserung $\geq 5$ )	1,0 [0,7-1,4], p=0,997	Kein Zusatznutzen belegt
- Körperliche Schmerzen (Verbesserung $\geq 5$ )	0,9 [0,7-1,2], p=0,522	Kein Zusatznutzen belegt
- Allgemeine Gesundheit (Verbesserung ≥ 5)	1,1 [0,7-1,7], p=0,670	Kein Zusatznutzen belegt
- Körperliche Rollenfunktion (Verbesserung $\geq 5$ )	1,1 [0,7-1,6], p=0,687	Kein Zusatznutzen belegt
- Emotionale Rollenfunktion (Verbesserung $\geq 5$ )	1,0 [0,7-1,6], p=0,920	Kein Zusatznutzen belegt
- Psychisches Wohlbefinden (Verbesserung ≥ 5)	0,9 [0,6-1,4], p=0,675	Kein Zusatznutzen belegt
- Soziale Funktion (Verbesserung $\geq 5$ )	0,8 [0,6-1,1], p=0,253	Kein Zusatznutzen belegt
FACIT-F (Verbesserung $\geq 4$ )	0,9 [0,6-1,2], p=0,332	Kein Zusatznutzen belegt
$HAQ-DI \ge 0,22$	1,0 [0,8-1,2], p=0,670	Kein Zusatznutzen belegt
$HAQ-DI \ge 0.3$	0,8 [0,6-1,1], p=0,183	Kein Zusatznutzen belegt
$HAQ-DI \ge 0.5$	0,8 [0,6-1,1], p=0,206	Kein Zusatznutzen belegt
WLQ		
- Indexscore (Verbesserung ≥ 5)	1,1 [0,5-2,7], p=0,798	Kein Zusatznutzen belegt
- Zeitmanagement (Verbesserung ≥ 13)	0,5 [0,2-0,9], p=0,030	Kein Zusatznutzen belegt
- Körperliche Anforderungen (Verbesserung $\geq 13$ )	0,7 [0,3-1,8], p=0,446	Kein Zusatznutzen belegt
- Mentale/Interpersonelle Anforderungen (Verbesserung $\geq 13$ )	1,2 [0,5-2,5], p=0,732	Kein Zusatznutzen belegt
- Ergebnisorientierte Anforderungen (Verbesserung ≥ 13)	0,7 [0,3-1,4], p=0,289	Kein Zusatznutzen belegt
MOS Schlafscore (Verbesserung ≥ 6)	0,8 [0,6-1,2], p=0,354	Kein Zusatznutzen belegt
Unerwünschte Ereignisse		

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

männlich	0,5 [0,2-0,96], p=0,039	Hinweis auf geringen Zusatznutzen
weiblich	1,2 [0,99-1,5], p=0,057	Kein Zusatznutzen belegt
Schwerwiegende unerwünschte Ereignisse		Kein Zusatznutzen belegt
Alter ≤ 65 Jahre	6,2 [1,4-26,6], p=0,014	Hinweis auf geringeren Nutzen
Alter > 65 Jahre	0,3 [0,0-3,1], p=0,324	Kein Zusatznutzen belegt
Unerwünschte Ereignisse, die zum Therapieabbruch geführt hatten	1,5 [0,5-4,4], p=0,457	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen	4,7 [0,2-96,4], p=0,315	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung	1,3 [0,9-2,0], p=0,189	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster	6,6 [0,3-125,4], p=0,210	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs	0,3 [0,0-7,6], p=0,475	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs)	2,8 [0,1-68,2], p=0,523	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte	3,8 [0,4-32,9], p=0,232	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Neutropenie	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Anämie	2,8 [1,2-68,2], p=0,523	Kein Zusatznutzen belegt

Stand: 21.04.2017

Tofacitinib (XELJANZ®)

Seite 32 von 442

Begründung für die Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens von Tofacitinib im Anwendungsgebiet A – erwachsene Patienten mit mittelschwerer bis schwerer RA, die bisher unzureichend nur auf MTX angesprochen haben und für die keine MTX-Unverträglichkeit vorliegt – Teilpopulation "c)-Kombi":

Insgesamt sind bei den dargestellten patientenrelevanten Endpunkten zur Morbidität von Tofacitinib nur sehr wenige statistisch signifikante Ergebnisse aufgetreten. Dies könnte angesichts einiger positiver Trends zugunsten von Tofacitinib (+ MTX) im Vergleich zur ZVT (Adalimumab + MTX) dadurch begründet sein, dass nur eine Subgruppe der Studie ORAL STANDARD für die Teilpopulation "c)-Kombi" zur Bewertung zur Verfügung steht. Hinsichtlich der Sicherheit/Verträglichkeit zeigen sich wenige positive Ergebnisse zugunsten von Tofacitinib (+ MTX) im Vergleich zur ZVT (Adalimumab + MTX), jedoch auch mehrere Trends zuungunsten von Tofacitinib. Auch hinsichtlich der Sicherheit/Verträglichkeit traten insgesamt nur sehr wenige statistisch signifikante Ergebnisse bei den dargestellten Endpunkten auf.

Da es sich bei der Evidenzbasis um eine Teilpopulation der doppelblinden Studie ORAL STANDARD handelt (39% der Patienten), die auf Studienebene und auf der Ebene der meisten dargestellten Endpunkte niedriges Verzerrungspotenzial besitzt, ist für alle nachfolgend diskutierten statistisch signifikanten Ergebnisse die Aussagesicherheit eines "Hinweises" gerechtfertigt.

Im Detail konnte kein Zusatznutzen zur Wirksamkeit von Tofacitinib belegt werden, da es zu keinem der dargestellten Wirksamkeits-Endpunkte statistisch signifikante Ergebnisse zu verzeichnen gab. Damit ist hinsichtlich der Wirksamkeit kein Zusatznutzen für Tofacitinib belegt. Ein geringerer Nutzen von Tofacitinib im Vergleich zur ZVT liegt ebenfalls nicht vor.

Hinsichtlich der Morbidität trat ein statistisch signifikanter Unterschied zuungunsten von Tofacitinib bezüglich der Responder-Rate der **Dimension "Zeitmanagement" des WLQ** (work limitations questionnaire) auf (relatives Risiko für Response (Verbesserung): 0,5, 95%-Konfidenzintervall [0,2-0,9], p=0,03). Dieser vom Patienten ausgefüllte Fragebogen (WLQ) konnte naturgemäß nur von den Studienpatienten ausgefüllt werden, die im Studienverlauf berufstätig waren. Dadurch ist die Zahl der vorhandenen Datensätze zum WLQ deutlich geringer, als es dem FAS (full analysis set) entsprochen hätte. Da jedoch die Randomisierung in der Studie nicht hinsichtlich des Faktors "Berufstätigkeit" stratifiziert worden war, ist für diesen Endpunkt insgesamt ein hohes Verzerrungspotenzial zu erwarten. Auch der hohe Anteil von fehlenden Werten (ca. 40-60% in den jeweiligen Studienarmen) führt dazu, dass eine sinnvolle Interpretation der Ergebnisse des WLQ nicht möglich ist. Deshalb ist in der Gesamteinschätzung dieser Endpunkt für die Ableitung eines Zusatznutzens nicht interpretierbar.

Zu keinem Endpunkt der gesundheitsbezogenen Lebensqualität lagen statistisch signifikante Ergebnisse vor.

Zur Domäne der Sicherheit/Verträglichkeit gab es hinsichtlich der **Häufigkeit** unerwünschter Ereignisse bei einem für die Teilpopulation "c)-Kombi" statistisch nicht signifikanten Ergebnis einen Beleg auf Effektmodifikation für den Faktor Geschlecht (Interaktions-p-Wert = 0,012). Nur für die männlichen Patienten zeigte sich ein statistisch signifikanter Effekt zugunsten von Tofacitinib in einer Reduktion der Häufigkeit unerwünschter Ereignisse versus Adalimumab (relatives Risiko 0,5 [0,2-0,96], p=0,039). Für die weiblichen Patienten ist ein Trend zuungunsten von Tofacitinib zu erkennen, der jedoch nicht statistisch signifikant ist. Insofern sind die Ergebnisse der beiden Subgruppen nicht als gleichgereichtet anzusehen, und daher wird für die **Subgruppe der männlichen Patienten** das Vorliegen eines **Hinweises auf einen geringen Zusatznutzen** von Tofacitinib gesehen.

Bei der **Häufigkeit schwerwiegender unerwünschter Ereignisse** zeigte sich ein statistisch signifikanter Effekt zuungunsten von Tofacitinib in der Subgruppe der jüngeren Patienten (≤ 65 Jahre). Da sich kein plausibles Muster bei den schwerwiegenden unerwünschten Ereignissen erkennen lässt, welches ein relevantes Schadenspotenzial von Tofacitinib im Vergleich zu bDMARD befürchten ließe und ein positives Nutzen-Risiko-Verhältnis durch die EMA bestätigt wurde, ist die Ableitung eines generell geringeren Nutzens nicht gerechtfertigt. **Ein Zusatznutzen ist somit nicht belegt.** 

Unter Tofacitinib traten in dieser Teilpopulation c) insgesamt 13 schwerwiegende unerwünschte Ereignisse auf. Diese verteilten sich auf insgesamt zehn unterschiedliche System Organ Classes (SOC) entsprechend der MedDRA Terminologie. Auf keine einzelne SOC, wie beispielsweise "Infektionen und parasitäre Erkrankungen" (*Infections and Infestations*), entfielen mehr als zwei schwerwiegende unerwünschte Ereignisse. Des Weiteren fällt bei genauerer Betrachtung der schwerwiegenden unerwünschten Ereignisse auf, dass bei nur drei schwerwiegenden unerwünschten Ereignissen der Studienarzt einen kausalen Zusammenhang zwischen der Studienmedikation und dem Ereignis für wahrscheinlich hielt. Somit lässt sich keine plausibles Muster bei den schwerwiegenden unerwünschten Ereignissen erkennen, welches ein relevantes Schadenspotenzial von Tofacitinib im Vergleich zu bDMARD befürchten ließe.

Zu keinem weiteren Endpunkt der Kategorie Sicherheit/Verträglichkeit lagen statistisch signifikante Ergebnisse vor.

In der Gesamtschau besteht die Aufgabe, einen positiven Effekt in einer Subgruppe (Männer) gegen einen negativen Effekt in einer anderen Subgruppe (jüngere Patienten) bewertend in Relation zu setzen. Da der singuläre Effekt zuungunsten von Tofacitinib nur bei berufstätigen Patienten auftrat und das Verzerrungspotenzial dieses Endpunktes als hoch eingestuft wurde, kann dieses Ergebnis ebenfalls nicht ohne weiteres auf die gesamte Teilpopulation "c)-Kombi" übertragen werden. In der Gesamtschau scheint deshalb das Fehlen eines Nachweises für einen Zusatznutzen bei gleichzeitig ebenfalls fehlendem Nachweis für einen generell geringeren Nutzen von Tofacitinib eine gerechtfertigte Bewertung der vorliegenden Ergebnisse für die Teilpopulation "c)-Kombi". Für die Teilpopulation "c)-Kombi" lässt sich somit kein Zusatznutzen ableiten.

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

Tabelle 4-3: Übersicht über Wahrscheinlichkeit und Ausmaß des Zusatznutzens aus RCT auf Endpunktebene – Population d)-Kombi

Stand: 21.04.2017

Studie: A3921064 (ORAL STANDARD)	RR [95% KI], p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
DAS28-4(ESR) < 2,6		Kein Zusatznutzen belegt
$DAS28-4(ESR) \le 5,1$	1,5 [0,7-3,5], p=0,344	Kein Zusatznutzen belegt
DAS28-4(ESR) $> 5,1$	0,4 [0,2-0,9], p=0,035	Kein Zusatznutzen belegt
$DAS28-4(ESR) \le 3,2$	0,7 [0,4-1,3], p=0,250	Kein Zusatznutzen belegt
DAS28-4(CRP) < 2,6	1,0 [0,6-1,5], p=0,849	Kein Zusatznutzen belegt
$DAS28-4(CRP) \le 3,2$	1,0 [0,8-1,4], p=0,854	Kein Zusatznutzen belegt
EULAR-Response (ESR): gute Verbesserung	0,7 [0,4-1,3], p=0,250	Kein Zusatznutzen belegt
EULAR-Response (ESR): gute oder moderate Verbesserung	1,0 [0,8-1,2], p=0,828	Kein Zusatznutzen belegt
EULAR-Response (CRP): gute Verbesserung	1,1 [0,8-1,5], p=0,705	Kein Zusatznutzen belegt
EULAR-Response (CRP): gute oder moderate Verbesserung	0,9 [0,8-1,1], p=0,336	Kein Zusatznutzen belegt
ACR70	1,2 [0,7-2,1], p=0,430	Kein Zusatznutzen belegt
ACR50	1,0 [0,7-1,4], p=0,971	Kein Zusatznutzen belegt
ACR/EULAR-Remission	1,1 [0,5-2,6], p=0,763	Kein Zusatznutzen belegt
SDAI ≤ 3,3	0,9 [0,4-1,6], p=0,621	Kein Zusatznutzen belegt
SDAI ≤ 11	1,3 [0,9-1,8], p=0,190	Kein Zusatznutzen belegt
CDAI ≤ 2,8	1,0 [0,5-2,1], p=0,933	Kein Zusatznutzen belegt
CDAI ≤ 10	1,2 [0,8-1,7], p=0,376	Kein Zusatznutzen belegt
Schmerzbeurteilung durch den Patienten (Verbesserung $\geq 10$ )	0,1 [0,8-1,2], p=0,871	Kein Zusatznutzen belegt
Gesamtbeurteilung der Krankheitsaktivität durch den Patienten [PatGA] (Verbesserung ≥ 10)	1,1 [0,9-1,3], p=0,485	Kein Zusatznutzen belegt

Tofacitinib (XELJANZ®)

Seite 35 von 442

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

	1.0.50.0.1.11 0.626	W ' 7
Gesamtbeurteilung der Krankheitsaktivität durch den Arzt [PhysGA] (Verbesserung ≥ 10)	1,0 [0,8-1,1], p=0,626	Kein Zusatznutzen belegt
männlich	1,4 [0,9-2,0], p=0,102	Kein Zusatznutzen belegt
weiblich	0,9 [0,7-1,1], p=0,180	Kein Zusatznutzen belegt
Anzahl empfindlicher/druckschmerzhafter Gelenke $\leq 1-28$ untersuchte Gelenke	0,8 [0,6-1,3], p=0,433	Kein Zusatznutzen belegt
Anzahl empfindlicher/druckschmerzhafter Gelenke $\leq 1-68$ untersuchte Gelenke	0,7 [0,4-1,2], p=0,215	Kein Zusatznutzen belegt
Anzahl geschwollener Gelenke $\leq 1-28$ untersuchte Gelenke	1,2 [0,8-1,7], p=0,376	Kein Zusatznutzen belegt
Anzahl geschwollener Gelenke $\leq 1-66$ untersuchte Gelenke	1,1 [0,7-1,6], p=0,655	Kein Zusatznutzen belegt
SF-36		
- Körperlicher Komponentenscore (Verbesserung $\geq 2,5$ )	1,0 [0,8-1,2], p=0,862	Kein Zusatznutzen belegt
- Psychischer Komponentenscore (Verbesserung ≥ 2,5)	1,1 [0,8-1,5], p=0,640	Kein Zusatznutzen belegt
- Vitalität (Verbesserung $\geq 5$ )	1,1 [0,8-1,4], p=0,741	Kein Zusatznutzen belegt
- Körperliche Funktion (Verbesserung $\geq 5$ )	0,9 [0,7-1,1], p=0,292	Kein Zusatznutzen belegt
- Körperliche Schmerzen (Verbesserung $\geq 5$ )	0,9 [0,7-1,2], p=0,445	Kein Zusatznutzen belegt
- Allgemeine Gesundheit (Verbesserung $\geq 5$ )		Kein Zusatznutzen belegt
Europa	0,6 [0,4-0,95], p=0,029	Hinweis auf geringeren Nutzen
USA/Kanada	6,1 [0,9-43,3], p=0,070	Kein Zusatznutzen belegt
Lateinamerika	0,5 [0,2-1,8], p=0,298	Kein Zusatznutzen belegt
Sonstige	3,5 [1,1-10,8], p=0,030	Hinweis auf geringen Zusatznutzen
- Körperliche Rollenfunktion (Verbesserung $\geq 5$ )	1,0 [0,7-1,4], p=0,918	Kein Zusatznutzen belegt
- Emotionale Rollenfunktion (Verbesserung $\geq 5$ )	1,1 [0,7-1,5], p=0,759	Kein Zusatznutzen belegt
- Psychisches Wohlbefinden (Verbesserung ≥ 5)	1,1 [0,8-1,5], p=0,753	Kein Zusatznutzen belegt
- Soziale Funktion (Verbesserung $\geq 5$ )	1,0 [0,7-1,3], p=0,826	Kein Zusatznutzen belegt
FACIT-F (Verbesserung ≥ 4)	1,1 [0,8-1,4], p=0,620	Kein Zusatznutzen belegt
<b>HAQ-DI</b> ≥ 0,22	0,9 [0,7-1,1], p=0,304	Kein Zusatznutzen belegt
HAQ-DI ≥ 0,3	0,9 [0,7-1,2], p=0,369	Kein Zusatznutzen belegt
$HAQ-DI \ge 0.5$	0,9 [0,7-1,2], p=0,435	Kein Zusatznutzen belegt

Stand: 21.04.2017

Tofacitinib (XELJANZ®)

Seite 36 von 442

WLQ		
- Indexscore (Verbesserung ≥ 5)	2,9 [0,9-10,0], p=0,075	Kein Zusatznutzen belegt
- Zeitmanagement (Verbesserung ≥ 13)	1,7 [0,95-2,9], p=0,077	Kein Zusatznutzen belegt
- Körperliche Anforderungen (Verbesserung ≥ 13)	0,5 [0,2-1,3], p=0,151	Kein Zusatznutzen belegt
- Mentale/Interpersonelle Anforderungen (Verbesserung $\geq 13$ )	2,9 [1,2-7,1], p=0,016	Kein Zusatznutzen belegt
- Ergebnisorientierte Anforderungen (Verbesserung ≥ 13)	1,0 [0,6-1,8], p=0,931	Kein Zusatznutzen belegt
MOS Schlafscore (Verbesserung $\geq 6$ )	0,9 [0,6-1,2], p=0,417	Kein Zusatznutzen belegt
Unerwünschte Ereignisse	1,0 [0,8-1,1], p=0,721	Kein Zusatznutzen belegt
Schwerwiegende unerwünschte Ereignisse	1,3 [0,7-2,6], p=0,415	Kein Zusatznutzen belegt
Unerwünschte Ereignisse, die zum Therapieabbruch geführt hatten	0,8 [0,4-1,6], p=0,568	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen	4,0 [0,5-35,5], p=0,208	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung	0,9 [0,7-1,3], p=0,706	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster	0,8 [0,2-3,3], p=0,711	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs	0,3 [0,0-8,2], p=0,503	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs)	0,5 [0,1-5,5], p=0,574	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte	0,5 [0,1-5,5], p=0,574	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut	1,0 [0,1-15,9], p=0,995	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Neutropenie	3,0 [0,1-73,5], p=0,496	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Anämie	3,0 [0,1-73,5], p=0,496	Kein Zusatznutzen belegt

Stand: 21.04.2017

Tofacitinib (XELJANZ®)

Seite 37 von 442

Begründung für die Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens von Tofacitinib im Anwendungsgebiet A – erwachsene Patienten mit mittelschwerer bis schwerer RA, die bisher unzureichend auf MTX und mindestens ein weiteres csDMARD angesprochen haben und für die keine MTX-Unverträglichkeit vorliegt – Teilpopulation "d)-Kombi":

Insgesamt sind bei den dargestellten patientenrelevanten Endpunkten zur Wirksamkeit und gesundheitsbezogenen Lebensqualität von Tofacitinib nur sehr wenige statistisch signifikante Ergebnisse aufgetreten. Dies könnte dadurch begründet sein, dass nur eine Subgruppe der Studie ORAL STANDARD für die Teilpopulation "d)-Kombi" zur Bewertung zur Verfügung steht. Da deren separate Auswertung bei der Planung dieser RCT nicht absehbar war, besitzen diese Analysen keine ausreichende Power zum Nachweis patientenrelevanter Effekte. Hinsichtlich der Verträglichkeit zeigen sich sowohl einige positive Ergebnisse zugunsten von Tofacitinib (+ MTX) im Vergleich zur ZVT (Adalimumab + MTX), als auch einige Trends zuungunsten von Tofacitinib. Hinsichtlich der Verträglichkeit traten insgesamt keinerlei statistisch signifikante Ergebnisse bei den dargestellten Endpunkten auf.

Da es sich bei der Evidenzbasis um eine Teilpopulation der doppelblinden Studie ORAL STANDARD handelt (51% der Patienten), die auf Studienebene und auf der Ebene der meisten dargestellten Endpunkte niedriges Verzerrungspotenzial besitzt ist, für alle nachfolgend diskutierten statistisch signifikanten Ergebnisse die Aussagesicherheit eines "Hinweises" gerechtfertigt.

Hinsichtlich der Wirksamkeit konnte ein statistisch signifikanter Effekt hinsichtlich der **DAS28-Response** (DAS28-4(ESR) < 2,6) zuungunsten von Tofacitinib beobachtet werden. In den durchgeführten Sensitivitätsanalysen erweist sich die Aussagekraft dieses Ergebnisses jedoch nicht als robust, da in sechs der sieben Sensitivitätsanalysen mit unterschiedlichen Imputationsmethoden für fehlende Werte (siehe oben Tabelle 4-16) kein statistisch signifikanter Effekt auftritt. Lediglich bei einer reinen LOCF-Methode, die im konkreten Falle für diesen Endpunkt zu identischen Ergebnissen wie in der primären Analyse führt, ergibt sich zwangsläufig ebenfalls statistische Signifikanz. Die Schätzer für den Behandlungseffekt erweisen sich über alle Analysen hinweg jedoch als robust und weitgehend stabil.

Es wurde zudem ein Beleg auf Effektmodifikation bezüglich des Schweregrades der RA (DAS28-4(ESR)  $\leq 5,1$  vs. > 5,1) festgestellt (Interaktions-p-Wert = 0,029). Ein statistisch signifikanter Effekt zuungunsten von Tofacitinib wurde nur bei den Patienten festgestellt, die bei Baseline eine hohe Krankheitsaktivität aufwiesen (DAS28-4(ESR) > 5,1). In dieser Subgruppe betrug das relatives Risiko für Ansprechen (Response) = 0,4 [0,2-0,9], p=0,035. In der komplementären Subgruppe der Patienten ohne hohe Krankheitsaktivität gibt es einen Trend in die entgegengesetzte Richtung (zugunsten von Tofacitinib), der jedoch nicht statistisch signifikant ist. Dieser signifkante Effekt spiegelt sich jedoch nicht in anderen vergleichbaren Morbiditätsendpunkten, wie beispielsweise SDAI  $\leq 3,3$ , wider. Des Weiteren führt der hohe Anteil von fehlenden Werten (> 30%) dazu, dass eine sinnvolle Interpretation der Ergebnisse des DAS28-4(ESR) (siehe auch Abschnitt 4.3.1.3.1) generell nicht möglich ist.

Deshalb ist in der Gesamteinschätzung der Endpunkt DAS28-4(ESR) für die Ableitung eines Zusatznutzens nicht interpretierbar.

Hinsichtlich der Morbidität trat ein statistisch signifikanter Unterschied zugunsten von Tofacitinib bezüglich der Responder-Rate der **Dimension "Mentale/Interpersonelle Anforderungen" des WLQ** (work limitations questionnaire) auf (relatives Risiko für Response (Verbesserung): 2,9, 95%-Konfidenzintervall [1,2-7,1], p=0,016). Dieser vom Patienten ausgefüllte Fragebogen (WLQ) konnte naturgemäß nur von den Studienpatienten ausgefüllt werden, die im Studienverlauf berufstätig waren. Dadurch ist die Zahl der vorhandenen Datensätze zum WLQ deutlich geringer, als es dem FAS (full analysis set) entsprochen hätte. Da jedoch die Randomisierung in der Studie nicht hinsichtlich des Faktors "Berufstätigkeit" stratifiziert worden war, ist für diesen Endpunkt insgesamt ein hohes Verzerrungspotenzial zu erwarten. Auch der hohe Anteil von fehlenden Werten (ca. 30-35% in den jeweiligen Studienarmen) führt dazu, dass eine sinnvolle Interpretation der Ergebnisse des WLQ nicht möglich ist. Deshalb ist in der Gesamteinschätzung dieser Endpunkt für die Ableitung eines Zusatznutzens nicht interpretierbar.

Hinsichtlich der gesundheitsbezogenen Lebensqualität trat bei einem statistisch nicht signifikanten Unterschied bezüglich der Responder-Rate der Dimension "Allgemeine Gesundheit" des SF-36 ein Beleg auf Effektmodifikation hinsichtlich der Region der Studiendurchführung auf (Interaktions-p-Wert = 0,006). Die Effekte in jeder der vier Regionen (Europa, USA/Kanada, Lateinamerika und 'sonstige') sind dabei sehr heterogen und nur in den Regionen Europa und 'sonstige' statistisch signifikant. Bei der Region sonstige' handelt es sich um die Zusammenfassung von Patienten aus unterschiedlichen Regionen, sofern sie keiner der drei anderen Regionen zuzuordnen waren. Eine plausible Interpretation dieser heterogenen Effekte kann insgesamt nicht gegeben werden, es könnte sich auch um einen falsch positiven Effekt durch die hohe Zahl an Subgruppen-Effekttestungen handeln, da insgesamt 228 Interaktionstests ohne Kontrolle des Fehlers 1. Art für die Teilpopulation "d)-Kombi" vorgenommen wurden. Außerdem liegt in keiner der anderen sieben Einzeldimensionen und auch in beiden Komponentenscores des SF-36 kein weiterer Beleg auf Effektmodifikation bezüglich der Region vor, was ebenfalls die Plausibilität dieses Effektes fraglich erscheinen lässt. Für Europa ergibt sich jedoch ein statistisch signifikanter Effekt zuungunsten von Tofacitinib (relatives Risiko für Response (Verbesserung): 0,6, 95%-Konfidenzintervall [0,4-0,95], p=0,029). Für die zusammengefasste Region ,sonstige' ergibt sich ein statistisch signifikantes Ergebnis zugunsten von Tofacitinib (relatives Risiko für Response (Verbesserung): 3,5, 95%-Konfidenzintervall [1,1-10,8], p=0,030). Wegen der fehlenden Plausibilität werden daher der Hinweis auf einen geringeren Nutzen (Europa) und der Hinweis auf einen geringen Zusatznutzen ("sonstige") nicht interpretiert, sondern das Ergebnis für die gesamte Teilpopulation "d)-Kombi", welches nicht statistisch signifikant ist. Damit liegt für die Dimension "Allgemeine Gesundheit" des SF-36 weder ein belegter Zusatznutzen noch ein belegter geringerer Nutzen von Tofacitinib vor.

Zu keinem der Endpunkte der Kategorie Sicherheit/Verträglichkeit lagen statistisch signifikante Ergebnisse vor.

In der Gesamtschau ist ein positiver Effekt hinsichtlich der gesundheitsbezogenen Lebensqualität für alle Patienten in der Teilpopulation "d)-Kombi" einem negativen Effekt hinsichtlich der Wirksamkeit (DAS28-4(ESR)-Response) in der Subgruppe der bei Baseline schwerer erkrankten Patienten in Relation zu setzen. In der Gesamtschau scheint deshalb das Fehlen eines Nachweises für einen Zusatznutzen bei gleichzeitig ebenfalls fehlendem Nachweis für einen generell geringeren Nutzen von Tofacitinib eine gerechtfertigte Bewertung der vorliegenden Ergebnisse für die Teilpopulation "d)-Kombi". Für die Teilpopulation "d)-Kombi" lässt sich somit kein Zusatznutzen ableiten.

# $\frac{Schlussfolgerungen\ zum\ Zusatznutzen\ und\ zum\ therapeutisch\ bedeutsamen}{Zusatznutzen}$

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

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Anwendungsgebiet A – Population b)-Mono	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population b)-Kombi	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population c)-Mono	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population c)-Kombi	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population d)-Mono	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population d)-Kombi	Kein Zusatznutzen belegt

#### 4.2 Methodik

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

#### 4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

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

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

Die Fragestellung für das vorliegende Nutzendossier zu Tofacitinib in der Indikation für erwachsene Patienten mit moderater bis schwerer rheumatoiden Arthritis (RA) betrifft die Darlegung der verfügbaren aktuellen Daten- und Erkenntnislage zur Nutzenbewertung von Tofacitinib. Für das Anwendungsgebiet A wird die Fragestellung im hier vorliegenden Modul 4A für eine spezifische Teilpopulation im Rahmen der Indikation von Tofacitinib behandelt:

Welcher Zusatznutzen besteht für die Behandlung mit Tofacitinib in Kombination mit MTX (ggf. auch als Monotherapie bei MTX-Unverträglichkeit) im Vergleich zur zweckmäßigen Vergleichstherapie bei erwachsenen Patienten mit moderater bis schwerer rheumatoider Arthritis, die unzureichend auf eine oder mehrere csDMARD-Behandlungen ansprechen oder diese nicht vertragen?

#### Patientenpopulation

Das hier beschriebenen Anwendungsgebiet A umfasst die Patienten im Rahmen der Zulassung von Tofacitinib, die unzureichend auf eine vorhergehende Therapie mit mindestens einem klassischen synthetischen krankheitsmodifizierenden antirheumatischen Arzneimittel (csDMARD) angesprochen oder diese nicht vertragen haben. Im Anwendungsgebiet A dürfen

die Patienten im Indikationsgebiet noch keine Therapie mit (mindestens) einem biologisch hergestellten krankheitsmodifizierenden antirheumatischen Arzneimittel (bDMARD) erhalten haben. Entsprechend nationaler und internationaler Leitlinien kann das Anwendungsgebietes A auch durch die Bezeichnung "Zweitlinientherapie" der RA beschrieben werden. Das Anwendungsgebiet A ist eine Teilpopulation innerhalb der Indikation für Tofacitinib.

Eine zweite Teilpopulation der Indikation von Tofacitinib (Anwendungsgebiet B: Patienten, die bereits auf mindestens einem bDMARD unzureichend angesprochen haben) wird entsprechend in Modul 4B dargestellt.

Die Bezeichnung der relevanten Patienten-Teilpopulationen in diesem Modul 4A übernimmt als vereinfachende Kurzform die Kennzeichnung des G-BA aus der Niederschrift zur Beratung Tofacitinib (1). Demzufolge werden in diesem Modul 4A die folgenden drei Teilpopulationen adressiert:

- **Teilpopulation b**) Patienten ohne ungünstige Prognosefaktoren nach unzureichendem Ansprechen auf ein csDMARD
- **Teilpopulation c)** Patienten mit ungünstigen Prognosefaktoren nach unzureichendem Ansprechen auf ein csDMARD
- **Teilpopulation d**) Patienten nach unzureichendem Ansprechen auf mehr als ein csDMARD.

Als Patienten "mit ungünstigen Prognosefaktoren" werden diejenigen Patienten eingestuft, bei denen entweder (zumindest) Gelenkerosionen vorliegen, die (zumindest) bezüglich der Rheumafaktoren und/oder Antikörper gegen citrullinierte Peptide positiv getestet wurden, oder die (zumindest) eine hohe Krankheitsaktivität (d.h. DAS28 > 5,1) aufweisen.

Die Teilpopulation a) (nach der Bennenung in der Niederschrift des G-BA) umfasst Patienten, die bislang keine krankheitsmodifizierende Therapie erhalten haben. Da diese Patienten nicht durch die Indikation von Tofacitinib abgedeckt sind, wird die Teilpopulation a) nicht in diesem Nutzendossier adressiert.

#### Intervention

Die Intervention besteht aus der Gabe von 5mg BID Tofacitinib (in Kombination mit MTX oder ggfs. als Monotherapie) entsprechend der Fachinformation (2).

#### Vergleichstherapie

Für Patienten ohne ungünstige Prognosefaktoren, die unzureichend auf eine vorangegangene Behandlung mit einem csDMARD ansprachen (Kurzbezeichnung:

Population b)), werden csDMARDs als Mono- oder Kombinationstherapie als zweckmäßige Vergleichstherapie betrachtet (Population "b)-Mono" bzw. "b)-Kombi").

Für Patienten mit ungünstigen Prognosefaktoren, die unzureichend auf ein csDMARD ansprachen (Population c)) bzw. für Patienten, die unzureichend auf mehrere csDMARDs ansprachen (Population d)), wurden bDMARDs in Kombination mit MTX als Vergleichstherapie berücksichtigt (Populationen "c)-Kombi" bzw. "d)-Kombi"). Bei MTX-Unverträglichkeit kann die Vergleichstherapie für die Populationen c) und d) ggf. auch eine bDMARD-Monotherapie sein (Populationen "c)-Mono" bzw. "d)-Mono").

## **Endpunkte**

- Morbidität:
  - o DAS28-4
  - o EULAR-Response
  - o ACR/EULAR-Remission und ACR Response
  - SDAI (Simplified Disease Activity Index)
  - o CDAI (Clinical Disease Activity Index)
  - o Schmerzbeurteilung durch den Patienten
  - o Gesamtbeurteilung der Krankheitsaktivität durch den Patienten
  - o Gesamtbeurteilung der Krankheitsaktivität durch den Arzt
  - o Anzahl empfindlicher/druckschmerzhafter bzw. geschwollener Gelenke
  - o FACIT-F
  - o HAQ-DI
  - o WLO
  - o MOS Schlafscore
- Gesundheitsbezogene Lebensqualität:
  - o SF-36
- Sicherheit/Verträglichkeit:
  - Unerwünschte Ereignisse
  - Schwerwiegende unerwünschte Ereignisse
  - Therapieabbruch wegen unerwünschter Ereignisse
  - Unerwünschte Ereignisse von besonderem Interesse:
    - o Schwere Infektionen
    - o Jedwede Infektion und parasitäre Erkrankung
    - Herpes Zoster
    - Tuberkulose
    - Nicht-Melanom Hautkrebs
    - Gastrointestinale Perforation
    - o Malignität (ohne Nicht-Melanom Hautkrebs)
    - o Erhöhte Leberwerte
    - o Erhöhte Kreatininwerte im Blut
    - Neutropenie

- o Lymphopenie
- o Anämie

Die Auswahl der o.g. unerwünschten Ereignisse von besonderem Interesse erfolgte anhand der Angaben in der Fachinformation und des Risk Management Plans des zu bewertenden Arzneimittels Tofacitinib bzw. den jeweiligen Fachinformationen der zweckmäßigen Vergleichstherapie.

## **Studientyp**

Randomisierte kontrollierte Studie (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.

#### **Patientenpopulation**

Das hier beschriebene Anwendungsgebiet A umfasst die Patienten im Rahmen der Zulassung von Tofacitinib, die unzureichend auf eine vorhergehende Therapie mit mindestens einem csDMARD angesprochen haben. Im Anwendungsgebiet A dürfen die Patienten im Indikationsgebiet noch keine Therapie mit (mindestens) einem bDMARD erhalten haben. Entsprechend nationaler und internationaler Leitlinien kann das Anwendungsgebietes A auch durch die Bezeichnung "Zweitlinientherapie" der RA beschrieben werden. Die Bezeichnung der relevanten Patienten-Teilpopulationen auch in diesem Modul 4A übernimmt als vereinfachende Kurzform die Kennzeichnung des G-BA aus der Niederschrift zur Beratung Tofacitinib (1).

#### <u>Intervention</u>

Die Intervention besteht aus der Gabe von 5mg BID Tofacitinib (in Kombination mit MTX oder ggfs. als Monotherapie) entsprechend der Fachinformation (2).

#### Vergleichstherapie

Für Patienten ohne ungünstige Prognosefaktoren, die unzureichend auf eine vorangegangene Behandlung mit einem csDMARD ansprachen (Kurzbezeichnung: Population b)), werden csDMARDs als Mono- oder Kombinationstherapie als zweckmäßige Vergleichstherapie betrachtet (Population "b)-Mono" bzw. "b)-Kombi").

Für Patienten mit ungünstigen Prognosefaktoren, die unzureichend auf ein csDMARD ansprachen (Population c)) bzw. für Patienten, die unzureichend auf mehrere csDMARDs ansprachen (Population d)), wurden bDMARDs in Kombination mit MTX als Vergleichstherapie berücksichtigt (Populationen "c)-Kombi" bzw. "d)-Kombi"). Bei MTX-Unverträglichkeit kann die Vergleichstherapie für die Populationen c) und d) ggf. auch eine bDMARD-Monotherapie sein (Populationen "c)-Mono" bzw. "d)-Mono").

#### Endpunkte

Es wurden Studien eingeschlossen, für die mindestens zu einem der folgenden patientenrelevanten Endpunkte Ergebnisse vorlagen:

- Mortalität: Gesamtüberleben
- Morbidität
- Gesundheitsbezogene Lebensqualität
- Sicherheit/Verträglichkeit

#### Studientyp

Es wurde nach randomisierten kontrollierten Studien (RCT) gesucht.

#### Studiendauer

Die Mindeststudiendauer für den Einschluss betrug 24 Wochen (6 Monate). Mit diesem Kriterium wird der Position des G-BA gefolgt (1).

## <u>Datenquellen</u>

Zur Nutzenbewertung herangezogen wurden nur Studien, zu denen die Ergebnisse in Form einer Originalpublikation, eines ausführlichen Ergebnisberichts aus einem Studienregister oder eines Studienberichts vorlagen. Ausgeschlossen wurden Sekundärpublikationen, die keine neuen Daten (über die ohnehin in den Studienberichten/Originalpublikationen dargestellten) enthalten.

Im Detail wurden die folgende Ein- bzw. Ausschlusskriterien für die Suche nach im Anwendungsgebiet A relevanten RCT angewandt (siehe Tabelle 4-5).

Tabelle 4-5: Übersicht über Ein- und Ausschlusskriterien der zu berücksichtigenden Studien - RCT

	Einschlusskriterien	E	Ausschlusskriterien	A
Patientenpopulation	Erwachsene Patienten mit moderater bis schwerer rheumatoider Arthritis nach unzureichendem Ansprechen auf mindestens ein csDMARD oder falls dieses nicht vertragen wurde	1	Erwachsene Patienten ohne moderate bis schwere rheumatoide Arthritis oder ohne bisheriges csDMARD-Versagen.	1
Intervention	Tofacitinib 5mg BID (als Monotherapie oder in Kombination mit MTX)	2	Andere Interventionen	2
Vergleichstherapie	Entspricht der zweckmäßigen Vergleichstherapie (ZVT) nach Festlegung des G-BA (1) für mindestens eine der Teilpopulationen b), c) oder d):  • b): csDMARDs als Mono- oder Kombinationstherapie  • c) und d): bDMARDs in Kombination mit MTX bzw. als Monotherapie bei MTX-Unverträglichkeit	3	Entspricht für keine der Teilpopulation b), c) oder d) der Festlegung der ZVT	3
Endpunkte	Ergebnisse zu mind. einem patientenrelevanten Endpunkt bzgl.  • Mortalität  • Morbidität (Symptomatik)  • Lebensqualität  • Sicherheit/Verträglichkeit sind berichtet.	4	Ergebnisse zu keinem Endpunkt bzgl.  • Mortalität • Morbidität (Symptomatik) • Lebensqualität • Sicherheit/Verträglichkeit sind berichtet.	4
Studiendauer	≥ 24 Wochen (bzw. 6 Monate)	5	< 24 Wochen (bzw. 6 Monate)	5
Studientyp	Randomisierte kontrollierte Studie	6	Keine randomisierte kontrollierte Studie	6
Publikationstyp	Studienbericht, Primärpublikation oder Sekundärpublikation mit neuen Daten	7	Sekundärpublikation ohne neue Daten	7
Studienstatus	Studie abgeschlossen	8	Studie läuft noch	8

## 4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen

werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

#### 4.2.3.1 Studien des pharmazeutischen Unternehmers

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

#### 4.2.3.2 Bibliografische Literaturrecherche

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

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

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

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

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

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle

Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Am 07.06.2016 wurde eine systematische Literaturrecherche nach RCT im Indikationsgebiet durchgeführt. Am 02.02.2017 erfolgte eine Update-Recherche. In folgenden Datenbanken wurden hierfür jeweils recherchiert:

- MEDLINE
- EMBASE
- Cochrane Central Register of Controlled Trials

In MEDLINE und EMBASE wurde jeweils ein validierter Filter zur Suche nach RCT angewendet (3).

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

#### 4.2.3.3 Suche in Studienregistern

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

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

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

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

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in 0 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.

Die Suche in den Studienregistern wurde am 07.06.2016 mit den in Anhang 4-B genannten Suchstrategien durchgeführt. Am 02.02.2017 erfolgte eine Update-Recherche. Folgende Register wurden jeweils in die Recherche eingeschlossen:

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

#### 4.2.3.4 Selektion relevanter Studien

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

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

Nach Erhalt der Volltexte wurden diese ebenso von zwei Personen unabhängig voneinander bezüglich ihrer Relevanz durchgesehen und im Fall einer positiven Einordnung für die Nutzenbewertung herangezogen. Eine Liste der im Volltext gesichteten und ausgeschlossenen Studien ist mit Angabe des spezifischen Ausschlussgrunds in Anhang 4-C aufgeführt.

Ebenso wurde bei der Selektion der Treffer aus den Studienregistern verfahren. Diese zusätzlichen Quellen wurden in einem zweiten Schritt ebenfalls – sofern verfügbar – als Volltexte beschafft bzw. auf vorhandene Ergebnisdarstellungen in den Studienregistern untersucht. Eine Liste der ausgeschlossenen Studien ist mit Angabe des spezifischen Ausschlussgrunds in Anhang 4-D aufgeführt.

#### 4.2.4 Bewertung der Aussagekraft der Nachweise

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

## A: Verzerrungsaspekte der Ergebnisse auf Studienebene

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

## B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

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

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

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

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

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als "hoch" soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

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

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

#### 4.2.5 Informations synthese und -analyse

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

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten "Charakteristika der in die Bewertung eingeschlossenen Studien" und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)<sup>1</sup>. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-<sup>2</sup> bzw. STROBE-Statements<sup>3</sup> folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

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

Für die eingeschlossene randomisierte kontrollierte klinische Studie werden das Design und die Methodik gemäß der Anforderungen nach CONSORT dargestellt. In Anhang 4-E wird die für die Nutzenbewertung herangezogene RCT anhand der CONSORT-Items 2b bis 14 inklusive eines studienspezifischen *Flow-Charts* beschrieben.

\_

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

<sup>&</sup>lt;sup>2</sup> Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.

<sup>&</sup>lt;sup>3</sup> Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

## 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 Siedie für dieBewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Die Studienpopulationen der eingeschlossenen randomisierten klinischen Studie A3921064 (ORAL STANDARD) werden durch die folgenden zu Studienbeginn erhobenen Patientencharakteristika beschrieben:

- Geschlecht
- Alter
- Gewicht
- BMI
- Ethniziät
- Region
- Krankheitsdauer
- ESR-Wert (*Erythrocyte Sedimentation Rate*)
- CRP-Wert (C-reaktives Protein)
- DAS28-4
- Anti-CCP
- Rheumafaktor

Gemäß der Arzneimittel-Nutzenbewertungsverordnung sollen bei der Nutzenbewertung die folgenden Nutzenaspekte angemessen berücksichtigt werden: Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung der Lebensdauer, Verringerung von Nebenwirkungen sowie Verbesserung der Lebensqualität. Im vorliegenden

Dossier werden die Ergebnisse patientenrelevanter Endpunkte dargestellt, die diese Voraussetzung erfüllen. Da nur eine Studie in diese Nutzenbewertung eingeschlossen wird, ist die Beschreibung unterschiedlicher Operationalisierungen zwischen verschiedenen Studien obsolet.

Nachfolgend sind die Endpunkte aufgeführt, die in diesem Nutzendossier dargestellt sind:

#### Morbidität:

- DAS28-4
- EULAR-Response
- ACR/EULAR-Remission und ACR Response
- SDAI (Simplified Disease Activity Index)
- CDAI (Clinical Disease Activity Index)
- Schmerzbeurteilung durch den Patienten
- Gesamtbeurteilung der Krankheitsaktivität durch den Patienten
- Gesamtbeurteilung der Krankheitsaktivität durch den Arzt
- Anzahl empfindlicher/druckschmerzhafter bzw. geschwollener Gelenke
- FACIT-F
- HAQ-DI
- WLQ
- MOS Schlafscore

#### Gesundheitsbezogene Lebensqualität:

• SF-36

#### Sicherheit/Verträglichkeit:

- Unerwünschte Ereignisse
- Schwerwiegende unerwünschte Ereignisse
- Therapieabbruch wegen unerwünschter Ereignisse
- Unerwünschte Ereignisse von besonderem Interesse:
  - o Schwere Infektionen
  - o Jedwede Infektion und parasitäre Erkrankung
  - Herpes Zoster
  - Tuberkulose
  - Nicht-Melanom Hautkrebs
  - Gastrointestinale Perforation
  - o Malignität (ohne Nicht-Melanom Hautkrebs)
  - o Erhöhte Leberwerte
  - o Erhöhte Kreatininwerte im Blut

- o Neutropenie
- o Lymphopenie
- o Anämie

#### Patientenrelevanz der dargestellten Endpunkte:

Im Folgenden wird für die einzelnen Endpunkte jeweils beschrieben, weshalb sie als patientenrelevant einzustufen sind.

#### DAS28-4:

Der DAS28-4 ist ein gewichteter Score für die Erfassung der Krankheitsaktivität, der sich aus den folgenden Einzelkomponenten zusammensetzt:

- Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 28 Gelenken)
- Anzahl geschwollener Gelenke (von insgesamt 28 Gelenken)
- ESR- bzw. CRP-Wert
- Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA).

Laut G-BA Beratung ist der DAS-28 als patientenrelevant einzustufen (1). Die Ableitung der Schwellenwerte des DAS28 für Remission (< 2,8) und niedrige Krankheitsaktivität ( $\le$  3,2) ist in (4) beschrieben.

#### **EULAR-Response:**

EULAR-Response misst die Veränderung der Krankheitsaktivität im Verlauf. Ein Patient wird einer der drei Kategorien eingestuft: keine Verbesserung, moderate Verbesserung, gute Verbesserung (jeweils gegenüber Baseline). Die drei Kategorien beruhen auf den DAS28-4-Werten und sind anhand des folgenden Schemas definiert (4, 5):

Tabelle 4-6: Definition der EULAR-Response

aktueller DAS28-4 Wert	DAS-28-4 Änderung gegenüber Baseline			
DAS26-4 Welt	> 1,2	$> 0.6 \text{ und} \le 1.2$	≤ 0,6	
≤ 3,2	gute Verbesserung	moderate Verbesserung	keine Verbesserung	
$>$ 3,2 und $\leq$ 5,1	moderate Verbesserung	moderate Verbesserung	keine Verbesserung	
>5,1	moderate Verbesserung	keine Verbesserung	keine Verbesserung	

Die Validierung der Schwellenwerte orientiert sich an der Arbeit von van Gestel 1996 (5) und wurde später aktualisiert und angepasst (4). Aus dieser Operationalisierung ergibt sich, dass deren Kriterien eine detailliertere Diskrimination bezüglich eines Therapieansprechens ermöglichen als es die dichotome Betrachtung der DAS28-4 Remission leisten kann. Da es sich um eine alternative Operationalisierung des patientenrelevanten Endpunkts DAS-28 handelt, ist auch der Endpunkt EULAR-Response als patientenrelevant anzusehen.

## ACR/EULAR-Remission und ACR Response:

Der ACR ist ein Instrument, welches vom American College of Rheumatology definiert wurde zur Beurteilung der Wirksamkeit. Der ACR existiert in verschiedenen Varianten (hier ACR50 und ACR70), allgemein berechnet er sich aus den folgenden Einzelkomponenten:

- 1. Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 68 Gelenken)
- 2. Anzahl geschwollener Gelenke (von insgesamt 66 Gelenken)
- 3. Schmerzbeurteilung durch den Patienten (visuelle Analogskala)
- 4. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
- 5. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PhysGA)
- 6. CRP-Wert
- 7. HAQ-DI

Eine Remission gemäß ACR/EULAR-Kriterien ist erreicht, wenn in den obigen Kriterien 1., 2., 4. und 6. eine Verbesserung vorliegt (Anzahl Gelenke jeweils  $\leq 1$ , CRP  $\leq 1$  mg/dl, PatGA  $\leq 10$ ).

Ein Ansprechen bzgl. des ACR50 liegt vor, wenn der Patient in den ersten der beiden Einzelkomponenten eine mind. 50%-ige Verbesserung erfährt sowie jeweils mind. eine 50%-ige Verbesserung in mind. 3 der 5 weiteren Einzelkomponenten. Die Definition des ACR70 gilt analog.

Verbesserungen um 50% oder 70% gehören zu den relevanten Maßen zur Beurteilung von Behandlungseffekten im Behandlungsalltag und in klinischen Zulassungsstudien (6, 7). In diversen Cochrane-Reports zur Bewertung der Evidenz von DMARDs einschließlich zweier Netzwerk-Metaanalysen zur Bewertung von Biologika (8) bzw. Biologika und Tofacitinib (9) bei Patienten, die auf MTX oder andere klassische DMARDs nicht angesprochen hatten, wird ACR50 als primärer Outcome-Parameter herangezogen. Veränderungen des ACR von 50% oder sogar 70% sind entsprechend der obigen Definitionen als mit Sicherheit vom Patienten wahrnehmbar einzuschätzen. Obwohl der G-BA die klinische Relevanz dieser relativen Veränderungen, falls sie allein auftreten würden, kritisch sieht (1), werden diese aus den genannten Gründen als patientenrelevant angesehen. Analoge Effekte in weiteren patientenrelevanten Endpunkten sollten dennoch vorliegen, um diese Einschätzung zu erhärten.

#### SDAI:

Der SDAI ist ein Summenscore für die Erfassung der Krankheitsaktivität, der sich aus den folgenden Einzelkomponenten zusammensetzt:

- Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 28 Gelenken)
- Anzahl geschwollener Gelenke (von insgesamt 28 Gelenken)
- Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
- Gesamtbeurteilung der Krankheitsaktivität durch den Arzt (PhysGA)
- CRP-Wert.

Der SDAI (Simplified Disease Activity Index) ist ein validiertes und im Vergleich zum DAS28-4 Instrument zur Messung der Krankheitsaktivität von Patienten mit rheumatoider Arthritis in klinischen Studien (10-12). Da der Score überwiegend direkt patientenrelevante Krankheitsaktivitäts-Merkmale abbildet, wird er als patientenrelevant eingestuft. Auch für diesen Endpunkt gilt jedoch die Aussage des G-BA (1), dass das Muster mehrerer patientenrelevanter Endpunkte für die Beurteilung des Zusatznutzens heranzuziehen ist und bei nur punktuellen Effekten im SDAI zusätzlich diskutiert werden muss, welches der Wert des zusätzlichen Informationsgehaltes (hier u.a. gegenüber dem DAS-28) ist.

#### CDAI:

Der CDAI (*Clinical Disease Activity Index*) ist ein Summenscore für die Erfassung der Krankheitsaktivität, der im klinischen Alltag gegenüber dem SDAI bevorzugt angewendet wird, da die Notwenigkeit der Erhebung des CRP entfällt (13). Der CDAI setzt sich aus den folgenden Einzelkomponenten zusammen:

- Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 28 Gelenken)
- Anzahl geschwollener Gelenke (von insgesamt 28 Gelenken)
- Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
- Gesamtbeurteilung der Krankheitsaktivität durch den Arzt (PhysGA).

Aus Gründen der Praktikabilität und der Praxistauglichkeit wurde der CDAI entwickelt, der dem SDAI sehr ähnlich ist und auf die Erhebung des CRP-Werts verzichtet. Der CDAI wird aus analogen Überlegungen und mit ähnlichen Einschränkungen wie zuvor für den SDAI als patientenrelevant eingestuft.

## Schmerzbeurteilung durch den Patienten:

Die Schmerzbeurteilung durch den Patienten wird anhand einer visuellen Analogskala erfasst, bei der der Patient die Intensität seiner Schmerzen auf einer Skala von 0 (keinerlei Schmerz) bis 100 (größte Schmerzen) beurteilt. Es handelt sich um ein validiertes Instrument in der vorliegenden Indikation (14). Die Intensität der durch RA bedingten Schmerzen wird direkt vom Patienten erfahren und beurteilt und ist somit unmittelbar patientenrelevant. Auch laut G-BA Beratung ist der Endpunkt "Schmerzen" als patientenrelevant einzustufen (1).

## Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA):

Der PatGA (*Patient Global Assessment of Arthritis*) ist ein Instrument für die Erfassung der Krankheitsaktivität, bei der der Patient auf einer visuellen Analogskala die Frage "Wie schätzen Sie heute Ihre Krankheitsaktivität ein?" auf einer Skala von 0 (keinerlei Krankheitsaktivität) bis 100 (höchste Krankheitsaktivität) beurteilt. Es handelt sich um ein validiertes Instrument in der vorliegenden Indikation (14, 15). Die Auswirkung der Krankheitsaktivität wird vom Patienten erfahren und beurteilt und ist somit unmittelbar patientenrelevant. Auch laut G-BA Beratung ist die Erfassung der Krankheitsaktivität als Endpunkt als patientenrelevant einzustufen (1).

## Gesamtbeurteilung der Krankheitsaktivität durch den Arzt (PhysGA):

Der PhysGA (*Physician Global Assessment of Arthritis*) ist ein Instrument für die Erfassung der Krankheitsaktivität, bei der der Arzt die Krankheitsaktivität des Patienten auf einer Skala von 0 (keinerlei Krankheitsaktivität) bis 100 (höchste Krankheitsaktivität) beurteilt. Laut G-BA Beratung ist prinzipiell die Krankheitsaktivität als Endpunkt als patientenrelevant einzustufen (1). Die Beurteilung durch den Arzt ist jedoch nicht immer konform mit der des Patienten. Insbesondere da die PhysGA eine Komponente in einigen der dargestellten Endpunkte enthalten ist, ist die separate Darstellung dieses Endpunktes zur Beurteilung des Gesamtbildes hinsichtlich des Zusatznutzens sehr wichtig, weshalb er hier zur Beurteilung patientenrelevanter Effekte auch dargestellt wird.

#### Anzahl empfindlicher/druckschmerzhafter bzw. geschwollener Gelenke:

Die Anzahl der empfindlichen/druckschmerzhaften bzw. geschwollenen Gelenke wird mittels zweier verschiedener Ansätze bestimmt. Zum einen von insgesamt 28 Gelenken und zum anderen von insgesamt 68 (empfindliche/druckschmerzhafte) bzw. 66 (geschwollene) Gelenken. Der Patient selbst beurteilt den Status jedes dieser möglicherweise betroffenen Gelenke und daher ist jede Operationalisierung dieses Endpunktes unmittelbar patientenrelevant (1).

#### SF-36:

Der SF-36 ist ein generisches Instrument zur Erfassung der gesundheitsbezogenen Lebensqualität (16). Er besteht aus acht Domänen und zwei Komponentenscores:

- Körperlicher Komponentenscore
- Psychischer Komponentenscore
- Vitalität
- Körperliche Funktion
- Körperliche Schmerzen
- Allgemeine Gesundheit
- Körperliche Rollenfunktion
- Emotionale Rollenfunktion
- Psychisches Wohlbefinden
- Soziale Funktion.

Der SF-36 ist für die Indikation der rheumatoiden Arthritis validiert (17, 18). Die MCID liegt für die beiden Komponentenscores bei  $\geq 2,5$  Punkten und für die übrigen Domänen bei  $\geq 5$  Punkten (19, 20). Laut G-BA Beratung ist der SF-36 als patientenrelevant einzustufen (1).

#### FACIT-F:

Der FACIT-F (Functional Assessment of Chronic Illness Therapy - Fatigue) ist ein Instrument zur Erfassung der Erschöpfung (Fatigue). Hierbei handelt es sich um ein validiertes Messinstrument (21). Die MCID liegt bei  $\geq$  4 Punkten (20). Fatigue wird direkt vom Patienten erfahren und beurteilt und ist somit unmittelbar patientenrelevant. Auch laut G-BA Beratung ist Fatigue als patientenrelevanter Endpunkt einzustufen (1).

#### HAQ-DI:

Der HAQ-DI (*Health Assessment Questionnaire – Disability Index*) ist ein Instrument zur Erfassung der Behinderung durch RA (22, 23). Es werden acht alltägliche Funktionsbereiche (Anziehen, Aufstehen, Essen, Gehen, Körperpflege, Gegenstände reichen, Greifen, weitere Tätigkeiten) abgefragt und in einem Score zusammengefasst. Für jeden der acht Funktionsbereiche werden die Werte 0 (keine Einschränkung), 1 (leichte Einschränkung), 2 (starke Einschränkung) oder 3 (unfähig, die Tätigkeit auszuführen) vergeben. Der HAQ-DI Score berechnet sich dann als Mittelwert aus diesen acht Werten. Es handelt sich um einen in der Indikation RA validierten Score (24-26). Es werden direkt vom Patienten das Vorhandensein und die Ausprägung der körperlichen Einschränkung dokumentiert, der Endpunkt ist somit unmittelbar patientenrelevant. Als MCID gilt eine Verbesserung um mindestens 0,22 Punkte (6, 27), in diesem Dossier werden zudem Responderanalysen zu weiteren (größeren) Schwellenwerten (0,3 und 0,5 Punkte) dargestellt.

#### WLQ:

Der WLQ (*Work Limitations Questionnaire*) ist ein Instrument zur Erfassung der durch Krankheit verursachten Einschränkungen im Arbeitsleben (28). Der WLQ besteht aus vier Einzeldimensionen und einem Indexscore zur Erfassung des allgemeinen Arbeitsverlusts:

- Indexscore
- Zeitmanagement
- Körperliche Anforderungen
- Mentale/Interpersonelle Anforderungen
- Ergebnisorientierte Anforderungen.

Bei dem WLQ handelt es sich um einen validierten Fragebogen (28, 29). Es werden direkt vom Patienten das Vorhandensein und die Ausprägung der Einschränkungen im Arbeitsleben dokumentiert, der Endpunkt ist somit unmittelbar patientenrelevant. Für die Einzeldimensionen wird jeweils eine MCID von 13 Punkten, für den Indexscore eine MCID von 5 Punkten verwendet (30). Aus dem Kontext ergibt sich bereits, dass der WLQ nur von berufstätigen RA-Patienten ausgefüllt werden kann.

#### MOS Schlafscore:

Patienten mit rheumatoider Arthritis haben oft eine schlechte Schlafqualität und folglich eine verminderte Lebensqualität (31). Chronische Schlafstörungen sind zudem mit erhöhter Mortalität assoziiert (32). Schlafprobleme werden direkt vom Patienten erfahren und beurteilt, und dieser Endpunkt ist somit unmittelbar patientenrelevant. Der MOS (*Medical Outcome Study*) Schlafscore ist ein validiertes Instrument zur Erfassung von Schlafproblemen (33, 34). Der Score besteht aus sieben Subskalen und einem Summenscore, der die Schlafprobleme quantifiziert. In diesem Dossier werden die Ergebnisse des Summenscores dargestellt. Als Schwellenwert gilt eine Verbesserung um mind. 6 Punkten für im Sinne einer MCID als relevant (35, 36).

#### Daten zur Sicherheit/Verträglichkeit:

Die Darstellung der auftretenden unerwünschten Ereignisse (nach SGB V) während der Behandlung bzw. in einem definierten Nachbeobachtungszeitraum wird vom IQWiG als "Ermittlung des Schadens medizinischer Interventionen" bezeichnet. Der Nachweis der Verringerung der Nebenwirkungen gilt generell als eine Verminderung der Morbidität sowie eine Verbesserung der Lebensqualität (37). Die Patientenrelevanz dieser Endpunkte ist als allgemein anerkannt in der Nutzenbewertung zu betrachten. § 2 Abs. 3 der Arzneimittel-Nutzenbewertungsverordnung nennt die Verringerung der Nebenwirkungen explizit als patientenrelevanten Effekt. Der G-BA benennt ausführlich die darzustellenden Kategorien der Nebenwirkungen (unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse, Therapieabbrüche aufgrund von unerwünschten Ereignissen sowie unerwünschte Ereignisse von besonderem Interesse) (1), die wegen der Tatsache, dass der Patient sie erlebt und berichtet, unmittelbar patientenrelevant sind.

Die Auswahl der unerwünschten Ereignisse von besonderem Interesse erfolgte anhand der Angaben in der Fachinformation und des Risk Management Plans des zu bewertenden Arzneimittels Tofacitinib bzw. den jeweiligen Fachinformationen der zweckmäßigen Vergleichstherapie.

#### Dargestellte Effektschätzer

Als Effektschätzer auf Studienebene wird für dichotome Endpunkte (z.B. Responderanalysen, Sicherheitsendpunkte) das Relative Risiko inklusive des zugehörigen Konfidenzintervalls und des p-Werts berichtet. Zudem werden für alle dichotomen Endpunkte auf Ebene der Gesamtstudienpopulation das Odds Ratio sowie die absolute Risikoreduktion (jeweils inkl. 95%-Konfidenzintervall und p-Wert) angegeben. Die Interpretation der Ergebnisse findet primär anhand des Relativen Risikos statt. Im Falle, dass in mindestens einer Behandlungsgruppe keinerlei Ereignisse bezüglich des betrachteten Endpunktes aufgetreten sind, wird 0,5 als Kontinuitätskorrektur zu den betreffenden Zellen addiert. In diesen Einzelfällen wurden die Effektschätzer ad hoc für die Darstellungen in diesem Modul 4 des Nutzendossiers berechnet und können ggf. von den Angaben in den Zusatzanalysen für die Nutzenbewertung (38-41) geringfügig abweichen.

Wenn für kontinuierliche Endpunkte, die jeweils zu Studienbeginn und -ende erhoben wurden, auf individueller Ebene validierte Relevanzschwellen bekannt sind, werden diese Endpunkte dichotomisiert und als Responderanalysen mittels Relativer Risiken (und Odds Ratios bzw. absoluter Risikoreduktion auf Ebene der Teilpopulationen) ausgewertet.

Für Endpunkte, die im zeitlichen Verlauf mittels *Time-To-Event* Analysen ausgewertet werden (das betrachtete Ereignis im Verlauf ist die jeweilige "Response"), werden neben der Angabe der medianen Zeiten als Effektschätzer Hazard Ratios mit 95%-Konfidenzintervallen und p-Werten berichtet. Als Testverfahren zur Bestimmung der p-Werte kommen Cox Proportional Hazards Modelle zur Anwendung. Zudem werden Kaplan-Meier-Plots stratifiziert nach den Behandlungsarmen dargestellt.

#### 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<sup>4</sup> erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär 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<sup>5</sup> 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 I² und des statistischen Tests auf Vorliegen von Heterogenität<sup>6</sup> erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

Es wurde keine Meta-Analyse durchgeführt, da lediglich ein RCT für die Darstellung des Zusatznutzens eingeschlossen werden konnte.

#### 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

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

<sup>&</sup>lt;sup>5</sup> Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2008. S. 243-296.

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

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 primären Analyse für diese Nutzenbewertung werden die Responderanalysen gemäß des NRI-Ansatzes (nonresponder imputation) ausgewertet, der auch im statistischen Analyseplan der Studie geplant war und im Studienbericht dargestellt wurde. Bei diesem (konservativen) NRI-Ansatz wird für die Patienten mit fehlenden Werten zunächst LOCF angewandt. Patienten, die die Studie abgebrochen haben, wurden ab dem Studienzeitpunkt zu dem für sie keine Werte mehr verfügbar sind als Nonresponder gezählt. Als Sensitivitätsanalyse werden für die Ergebnisse der Gesamtpopulation für ausgewählte Endpunkte zusätzlich die Ergebnisse der beiden Methoden:

- keine Imputation
- Last Observation Carried Forward (LOCF)

angewendet und dargestellt.

Diese beiden Sensitivitätsanalysen werden für die folgenden Endpunkte in diesem Modul 4A berichtet:

- DAS28-4(ESR)
- EULAR-Response
- ACR/EULAR-Remission
- SF-36 Komponentenscores
- FACIT-F
- HAQ-DI
- MOS Schlafscore.

Für diese und für alle anderen Endpunkte sind die Sensitivitätsanalysen im Detail zu finden in (38).

Es wurden weitere vier Sensitivitätsanalysen für alle berichteten Endpunkte durchgeführt (38). Sie werden in Modul 4A aus Gründen der Übersichtlichkeit nicht ausführlich dargestellt – mit einer Ausnahme: für den Endpunkt DAS28-4(ESR) < 2,6 werden alle durchgeführten Analysen gezeigt, um die Robustheit dieses speziellen Ergebnisses ausführlicher zu untersuchen. Diese weiteren Sensitivitätsanalysen umfassen die folgenden methodischen Ansätze:

- "reine" NRI-Impuation 1: Alle Patienten mit fehlenden Werten werden als Nonresponder gewertet.
- "reine" NRI-Imputation 2:
   Wie zuvor, wobei Patienten, für die keinerlei Messwert vorliegt, ausgeschlossen werden. In der Analyse sind demnach nur Patienten, für die mindestens ein Messwert (Baseline oder später) vorliegt.
- "ZVT" Imputationsverfahren Für Patienten mit fehlenden Werten wird in beiden Behandlungsgruppen jeweils diejenige Rate an Nonrespondern imputiert, die derjenigen aus der Kontrollgruppe, d.h. der Gruppe, die in der Studie mit ZVT behandelt wird, entspricht.
- "Worst-case" Imputationsverfahren: Wie oben, jedoch wird jeweils Werte gemäß der Rate der "schlechteren" Behandlungsgruppe (d.h. diejenige mit der niedrigeren Responserate) imputiert. Die Behandlungsgruppe mit der "schlechteren" Rate wechselt entsprechend je nach Endpunkt.

Nur für den Endpunkt DAS28-4(ESR) wird zusätzlich die folgende Sensitivitätsanalyse durchgeführt und dargestellt:

• Lineare Interpolation der fehlenden DAS28-4(ESR)-Werte mittels der DAS28-4(CRP)-Werte. Die Daten werden gemäß der Formel:

```
\begin{aligned} DAS28-4(ESR) &= X + DAS-28-4(CRP) \text{ [im Tofacitinib-Arm]} \\ DAS28-4(ESR) &= Y + DAS-28-4(CRP) \text{ [im Kontrollarm]} \\ &\text{ersetzt.} \end{aligned}
```

#### 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.

Im Anschluss werden die im Modul 4A dargestellten Subgruppenanalysen beschrieben. Folgende Subgruppen wurden für alle patientenrelevanten Endpunkte analysiert:

- Alter ( $\le 65 / > 65 \text{ Jahre}$ )
- Geschlecht (männlich / weiblich)
- Region (Europa / USA und Kanada / Lateinamerika / Sonstige)
- Schweregrad nach DAS28-4(ESR) ( $\leq 5,1/>5,1$ )

Die Merkmale Alter, Geschlecht und Schweregrad sind laut Verfahrensordnung gefordert. Zusätzlich werden potentielle geographische Unterschiede durch den Faktor "Region" untersucht.

Eine detailliertere Unterteilung der Studienpopulation auf Länder- oder Zentrumsebene wird darüber hinaus nicht als sinnvoll erachtet, da die Anzahl der Patienten in den jeweiligen Subgruppen zu gering für eine Interpretation von Behandlungseffekten ist. Insgesamt sind durch die dargestellten Merkmale alle sozialrechtlich erforderlichen Kriterien enthalten, um eventuelle Effektmodifikationen, die relevant für die vorliegende Nutzenbewertung sind, abbilden und interpretieren zu können.

Subgruppenanalysen werden für alle Endpunkte der Morbidität, der gesundheitsbezogenen Lebensqualität und der Sicherheit/Verträglichkeit durchgeführt. Lediglich für den Endpunkt DAS28-4(CRP) wurden keine Subgruppenanalysen berechnet. Dies begründet sich darin, dass der analoge Endpunkt DAS28-4(ESR) auf Subgruppenebene analysiert wurde und die Ergebnisse in diesem Modul 4A dargestellt werden. Wegen der größeren Bedeutung des Parameters ESR (Blutsenkungsgeschwindigkeit) für den Versorgungsalltag wurde diesem der

Vorzug gegeben und die Interpretation daher auch primär darauf bezogen. Außerdem ist bekannt, dass Tofacitinib direkt den CRP-Wert beeinflusst und von daher das auch auf den CRP aufbauende Responder-Kriterium stärker verzerrt sein wird. Aus diesem Grund erscheint eine weitere Auffächerung der Ergebnisse für den DAS28-4(CRP) hinsichtlich der Subgruppen als verzichtbar.

Die Analysen auf Subgruppenebene wurden *post hoc* im Rahmen dieser Dossiererstellung angefertigt. *A priori* waren keine Subgruppenanalysen für die eingeschlossene Studie geplant. Sämtliche Subgruppenergebnisse für diese Nutzenbewertung sind zu finden in (38-41).

Effekte zwischen verschiedenen Subgruppen werden mittels Interaktionstests bewertet. Zur Ableitung von Belegen oder Hinweisen auf Effektmodifikation werden die p-Werte der Interaktionstests herangezogen. Dabei besteht bei einem p-Wert < 0,05 ein Beleg und bei 0,05 ≤ p < 0,20 ein Hinweis auf Effektmodifikation. Dieses Vorgehen wird wie gefordert gewählt, obwohl darauf hingewiesen werden muss, dass bei der Vielzahl von Subgruppenanalysen (Zahl der betrachteten Endpunkte mal Anzahl der Subgruppen-Variablen) hier das Problem des multiplen Testens zu beachten ist, welchem auch für solche Zusatzanalysen durch geeignete Adjustierung des Signifikanzniveaus begegnet werden müsste, um die Rate der "falsch positiven" Subgruppeneffekte kontrollieren zu können. Eine solche Adjustierung wurde gemäß IQWiG-Methodik (37) nicht angewandt. Jedoch ist die Tatsache des erheblich erhöhten Fehlerrisikos 1. Art generell für die Interpretation der Relevanz und Plausibilität bestimmter auffälliger Subgruppeneffekte zu berücksichtigen.

Im Falle von Belegen oder Hinweisen auf Effektmodifikation (Interaktions-p-Wert < 0,20) werden die Ergebnisse auf Subgruppenebene im vorliegenden Modul 4A dieses Nutzendossiers für die jeweiligen Endpunkte dargestellt. Liegt hingegen kein Hinweis auf eine Interaktion zwischen dem Behandlungseffekt und dem Faktor der Subgruppe vor (Interaktions-p-Wert  $\ge 0,20$ ), wird auf die ausführliche Darstellung der Ergebnisse auf Subgruppenebene verzichtet. Diese Subgruppenergebnisse sind jedoch vollständig zu finden in den Zusatzanalysen (40, 41).

Bei einem Beleg auf Effektmodifikation werden die Ergebnisse ausschließlich auf Subgruppenniveau interpretiert. Falls die Subgruppenergebnisse jedoch nach der Definition des IQWiG als gleichgerichtet einzustufen sind (37), erfolgt trotzdem eine Interpretation auf dem Niveau der Gesamtpopulation.

Liegt lediglich ein Hinweis auf Effektmodifikation vor, werden die Ergebnisse weiterhin bezüglich der Gesamtpopulation interpretiert, wobei in die Diskussion des Zusatznutzens die Tatsache des Hinweises auf Effektmodifikation einfließt. Hierbei kommt es darauf an, ob die Effektschätzer in den Subgruppen gleichgerichtet sind oder nicht. In der Regel wird in einer Situation, in der die Ergebnisse in den Subgruppen nicht gleichgerichtet sind, der eventuelle Zusatznutzen in der Gesamtpopulation abgewertet werden, wobei die medizinische und statistische Plausibilität für den Hinweis auf Effektmodifikation sowie der Grad der Nicht-Gleichgerichtetheit berücksichtigt werden.

Die Schwellenwerte/Kategorien für Subgruppen, die anhand eines stetigen bzw. kategoriellen Merkmals, d. h. Alter, Region oder Schweregrad, gebildet wurden, wurden nach medizinischer Rationale im Vorfeld dieser Dossiererstellung festgelegt. Die Ergebnisdarstellung in den Subgruppen erfolgt analog zu derjenigen auf Ebene der Gesamtstudienpopulation mit der Ausnahme, dass keine *Time-to-Event* Analysen für Responderanalysen berichtet werden.

#### 4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich<sup>7</sup>. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als "Mixed-Treatment-Comparison(MTC)-Meta-Analysen"<sup>8</sup>, "Multiple-Treatment-Meta-Analysen"<sup>9</sup> oder auch "Netzwerk-Meta-Analysen"<sup>10</sup>, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

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

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.

<sup>&</sup>lt;sup>7</sup> Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. Health Technol Assess 2005; 9(26): 1-148.

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

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

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

<sup>&</sup>lt;sup>11</sup> B. Schöttker, D. Lühmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen<sup>12</sup>.

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

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

Es wurde kein indirekter Vergleich für die Teilpopulationen c)-Kombi und d)-Kombi durchgeführt, da eine direkt vergleichende randomisierte Studie vorliegt.

Für die Teilpopulationen b)-Mono, b)-Kombi, c)-Mono und d)-Mono konnte trotz des Fehlens direkt vergleichender Evidenz aus RCT kein adjustierter indirekter Vergleich durchgeführt werden, da keine der in Frage kommenden Studien mit Tofacitinib Patienten ohne das Vorliegen ungünstiger Prognosefaktoren eingeschlossen hatte (für b)-Mono und b)-Kombi).

Des Weiteren liegen für die Teilpopulationen c)-Mono und d)-Mono keine Daten mit ausreichender, als patientenrelevant eingestufter Therapiedauer von mindestens 6 Monaten unter einem möglichen "Brückenkomparator" (es kommt aufgrund der bekannten Tofacitinib-Studien nur Placebo dafür in Betracht) vor. Nicht-adjustierte indirekte Vergleiche für diese

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

Teilpopulationen wurden wegen ihrer limitierten Aussagekraft als nicht zielführend für die Nutzenbewertung eingestuft und daher ebenfalls nicht durchgeführt.

## 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 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapicarme
A3921019	Ja	ja	abgeschlossen	12 Wochen (6 Wochen	Tofacitinib <sup>a</sup> 5 mg BID Tofacitinib <sup>a</sup> 15 mg BID
				Behandlung + 6 Wochen Follow-Up)	Tofacitinib <sup>a</sup> 30 mg BID Placebo
A3921025	Ja	ja	abgeschlossen	6 Monate	Tofacitinib <sup>a</sup> 1 mg BID  → Tofacitinib 5 mg BID <sup>b, c</sup>
					Tofacitinib <sup>a</sup> 3 mg BID  → Tofacitinib 5 mg BID <sup>b, c</sup>
					Tofacitinib <sup>a</sup> 5 mg BID
					Tofacitinib <sup>a</sup> 10 mg BID
					Tofacitinib <sup>a</sup> 15 mg BID
					Tofacitinib <sup>a</sup> 20 mg QD  → Tofacitinib 5 mg BID <sup>b, c</sup>
					Placebo → Tofacitinib 5 mg BID <sup>b, c</sup>
					- jeweils zuzüglich MTX als weiterbestehende Vortherapie
A3921035	Ja	ja	abgeschlossen	6 Monate	Tofacitinib <sup>a</sup> 1 mg BID  → Tofacitinib 5 mg BID <sup>b</sup>
					Tofacitinib <sup>a</sup> 3 mg BID  → Tofacitinib 5 mg BID <sup>b, c</sup>
					Tofacitinib <sup>a</sup> 5 mg BID
					Tofacitinib <sup>a</sup> 10 mg BID
					Tofacitinib <sup>a</sup> 15 mg BID
					Adalimumab 40 mg jede 2. Woche SC → Tofacitinib 5 mg BID <sup>b</sup>
					Placebo -> Tofacitinib 5 mg BID <sup>b, c</sup>
A3921039	Ja	ja	abgeschlossen	12 Wochen	Tofacitinib <sup>a</sup> 1 mg BID
					Tofacitinib <sup>a</sup> 3 mg BID
					Tofacitinib <sup>a</sup> 5 mg BID
					Tofacitinib <sup>a</sup> 10 mg BID
					Placebo
					- jeweils zuzüglich

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapicarme
					MTX als weiterbestehende Vortherapie
A3921040	Ja	ja	abgeschlossen	12 Wochen	Tofacitinib <sup>a</sup> 1 mg BID Tofacitinib <sup>a</sup> 3 mg BID Tofacitinib <sup>a</sup> 5 mg BID Tofacitinib <sup>a</sup> 10 mg BID Tofacitinib <sup>a</sup> 15 mg BID Placebo
A3921044 (ORAL SCAN)	Ja	ja	abgeschlossen	6 Monate	Tofacitinib <sup>a</sup> 5 mg BID  Tofacitinib <sup>a</sup> 10 mg BID  Placebo → Tofacitinib 5 mg BID <sup>b, c</sup> Placebo → Tofacitinib 10 mg BID <sup>b, c</sup> - jeweils zuzüglich  MTX als  weiterbestehende  Vortherapie
A3921045 (ORAL SOLO)	Ja	ja	abgeschlossen	6 Monate	Tofacitinib <sup>a</sup> 5 mg BID  Tofacitinib <sup>a</sup> 10 mg BID  Placebo → Tofacitinib 5 mg BID <sup>d</sup> Placebo → Tofacitinib 10 mg BID <sup>d</sup>
A3921046 (ORAL SYNC)	Ja	ja	abgeschlossen	12 Monate	Tofacitinib <sup>a</sup> 5 mg BID  Tofacitinib <sup>a</sup> 10 mg BID  Placebo → Tofacitinib 5 mg BID <sup>b, c</sup> Placebo → Tofacitinib 10 mg BID <sup>b, c</sup> - jeweils zuzüglich csDMARD als weiterbestehende Vortherapie
A3921064 (ORAL STANDARD)	ja	ja	abgeschlossen	12 Monate	Tofacitinib <sup>a</sup> 5 mg BID  Tofacitinib <sup>a</sup> 10 mg BID  Placebo → Tofacitinib 5 mg BID <sup>b, c</sup> Placebo → Tofacitinib 10 mg BID <sup>b, c</sup> Adalimumab 40 mg jede 2. Woche SC  - jeweils zuzüglich  MTX als

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
					weiterbestehende Vortherapie
A3921133	Ja	ja	laufend	bis 1.500 Patienten für 3 Jahre beobachtet worden sind	Tofacitinib <sup>a</sup> 5 mg BID Tofacitinib <sup>a</sup> 10 mg BID TNF-Inhibitor (Adalimumab jede 2. Woche SC (US, Puerto Rico, Kanada) oder Etanercept 50 mg wöchentlich SC (alle anderen Länder))
A3921187 (ORAL STRATEGY)	nein	ja	abgeschlossen, noch keine Ergebnisse vorliegend	12 Monate	Tofacitinib 5 mg BID Tofacitinib 5 mg BID + MTX Adalimumab 40 mg jede 2. Woche SC + MTX
A3921237	Ja	ja	abgeschlossen	14 Wochen	Tofacitinib <sup>a</sup> 5 mg BID Placebo - jeweils zuzüglich MTX als weiterbestehende Vortherapie

<sup>&</sup>lt;sup>a</sup> CP-690,550 ist die im Studienprotokoll vor der Vergabe des internationalen Freinamens (INN) verwendte Bezeichnung für Tofacitinib

Geben Sie an, welchen Stand die Information in Tabelle 4-7 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 oben dargestellte Studienstatus bildet den Stand zum 08.02.2017 ab.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-7 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

<sup>&</sup>lt;sup>b</sup> Nach 3 Monaten wurden die Patienten bei Nichtansprechen auf die jeweilige Dosis von Tofacitinib umgestellt.

<sup>&</sup>lt;sup>c</sup> Nach 6 Monaten wurden alle Patienten der Placebo-Arme umgestellt.

<sup>&</sup>lt;sup>d</sup> Nach 3 Monaten wurden die Patienten auf die jeweilige Dosis von Tofacitinib umgestellt.

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

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
A3921019	Vergleich nicht gegen die zweckmäßige Vergleichstherapie; Studiendauer zu gering
A3921025	Vergleich nicht gegen die zweckmäßige Vergleichstherapie, Dauer der Komparator-Therapie (Placebo) zu kurz
A3921035	Dauer der Behandlung mit der zweckmäßigen Vergleichstherapie zu kurz (3 Monate)
A3921039	Vergleich nicht gegen die zweckmäßige Vergleichstherapie; Studiendauer zu gering
A3921040	Vergleich nicht gegen die zweckmäßige Vergleichstherapie; Studiendauer zu gering
A3921044 (ORAL SCAN)	Vergleich nicht gegen die zweckmäßige Vergleichstherapie; Dauer der Komparator-Therapie (Placebo) zu kurz
A3921045 (ORAL SOLO)	Vergleich nicht gegen die zweckmäßige Vergleichstherapie; Dauer der Komparator-Therapie (Placebo) zu kurz
A3921046 (ORAL SYNC)	Vergleich nicht gegen die zweckmäßige Vergleichstherapie; Dauer der Komparator-Therapie (Placebo) zu kurz
A3921133	Studie noch nicht abgeschlossen
A3921187 (ORAL STRATEGY)	Studie kürzlich klinisch abgeschlossen, jedoch keine Ergebnisse vorliegend
A3921237	Vergleich nicht gegen die zweckmäßige Vergleichstherapie; Studiendauer zu gering

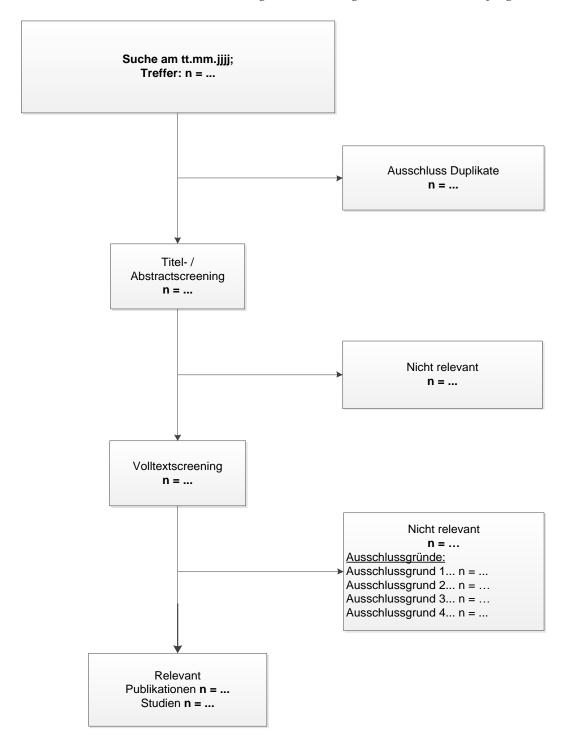
# 4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

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.



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

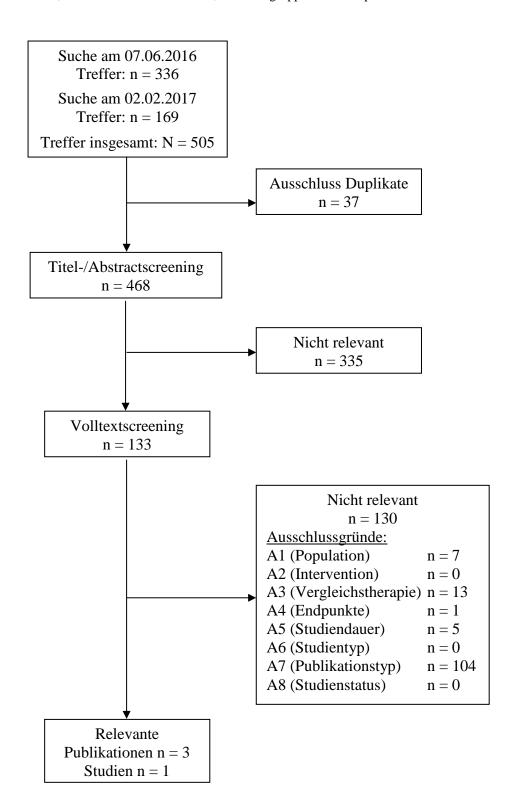


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

# 4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-7) 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 4-9: 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 <sup>a</sup> )	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Status (abgeschlossen/ abgebrochen/ laufend)
A3921064 (ORAL STANDARD)	clinicaltrials.gov [NCT00853385] (42) ICTRP (43, 44) EU-CTR [2008-008338- 35] (45) PharmNet.Bund (46)	ja	ja	abgeschlossen
A3921133	clinicaltrials.gov [NCT02092467] (47) ICTRP (48) EU-CTR [2013-003177- 99] (49)	ja	nein	laufend
A3921187 (ORAL STRATEGY)	clinicaltrials.gov [NCT02187055] (50) ICTRP (51) EU-CTR [2014-000358- 13] (52) PharmNet.Bund (53)	ja	nein	abgeschlossen, noch keine Ergebnisse vorliegend

 $a: Zit at \ des \ Studienregister eintrags \ sowie, \ falls \ vorhanden, \ der \ im \ Studienregister \ aufgelisteten \ Berichte \ "" ber Studiendesign \ und/oder \ -ergebnisse.$ 

Geben Sie an, welchen Stand die Information in Tabelle 4-9 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 in den Studienregistern wurde am 07.06.2016 durchgeführt und am 02.02.2017 aktualisiert.

# 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 Zwischenzeilen- überschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-10: Studienpool –	- RCT mit dem zi	u bewertenden Arzneimittel

	Studienkategorie			verfügbare Quellen <sup>a</sup>		
Studie	Studie zur Zulassung des zu bewertenden	gesponserte Studie <sup>b</sup>	Studie Dritter	Studien- bericht	Register- eintrag <sup>c</sup>	Publikation
	Arzneimittels (ja/nein)	(ja/nein)	(ja/nein)	(ja/nein [Zitat])	(ja/nein [Zitat])	(ja/nein [Zitat])
A3921064 (ORAL STANDARD)	ja	ja	nein	ja (54)	ja (42-46)	ja (55-57)

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.

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

Tabelle 4-11: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
A3921064 (ORAL STANDARD)	RCT, Phase III, doppelblind, parallel, placebokontrolliert	Erwachsene Patienten mit aktiver RA, die Methotrexat als stabile Begleitmedikation erhalten und noch nicht auf einem b- DMARD"versagt" haben	<ul> <li>Tofacitinib 5 mg BID (n=204)</li> <li>Tofacitinib 10 mg BID (n=201)<sup>a</sup></li> <li>Placebo → Tofacitinib 5 mg BID (n=56)<sup>a,b</sup></li> <li>Placebo → Tofacitinib 10 mg BID (n=52)<sup>a,b</sup></li> <li>Adalimumab 40 mg jede 2. Woche (n=204)</li> </ul>	1 Jahr Behandlung	115 Zentren weltweit 05/2009 – 03/2011	ACR20, ACR50,ACR70, HAQ-DI, DAS28-4(ESR), DAS28-4(CRP), DAS28-3(ESR), DAS28-3(CRP), patientenberichtete Endpunkte (SF-36, WLQ, EQ-5D, MOS Schlafscore, FACIT), unerwünschte Ereignisse, Pharmakokinetik

<sup>&</sup>lt;sup>a</sup> Da diese Studienarme für die Bewertung des Zusatznutzens nicht relevant sind, werden im Folgenden ausschließlich die beiden relevanten Studienarme berichtet.

Tofacitinib (XELJANZ®)

<sup>&</sup>lt;sup>b</sup> Nach 3 Monaten wurden die Patienten bei Nichtansprechen auf die jeweilige aktive Substanz umgestellt. Nach 6 Monaten wurden alle Patienten der Placebo-Arme umgestellt.

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

Studie	Tofacitinib 5 mg	Adalimumab 40 mg	ggf. weitere Spalten mit Behandlungscharakteristika z.B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
A3921064 (ORAL STANDARD)	Tofacitinib 5 mg BID oral +	Adalimumab 40 mg SC jede 2. Woche	Bestehende, stabile Behandlung mit Methotrexat als Begleitmedikation
	Placebo-Injektionen SC jede 2. Woche	Placebo-Tabletten BID	

Aufgrund der Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA (1) für Anwendungsgebiet A besteht die Notwendigkeit, die Studienpopulation in zunächst drei Teilpopulationen aufzuteilen. Der Einfachheit und Übersichtlichkeit halber wird in diesem Dossier die vom G-BA gewählte kurze Bezeichnung dieser Teilpopulationen beibehalten, d.h. es handelt sich innerhalb des Anwendungsgebietes A um die Teilpopulationen b), c) und d) (siehe dazu auch Modul 3A des Nutzendossiers).

Diese drei Teilpopulationen müssen jeweils weiterhin dahingehend unterschieden und getrennt betrachtet werden, ob die Patienten jeweils unter einer MTX-Unverträglichkeit leiden bzw. eine Weiterbehandlung mit MTX ungeeignet wäre oder nicht. Im Fall einer MTX-Unverträglichkeit (oder Nicht-Eignung für eine MTX-Weiterbehandlung) kommt für diese Patienten eine Behandlung mit MTX nicht mehr in Frage, und sie kommen für eine Monotherapie mit Tofacitinib bzw. der jeweiligen ZVT in Frage. Im anderen Fall kann die laut Indikation für Tofacitinib vorgesehene Kombinationsbehandlung mit MTX erfolgen und in Analogie soll dann auch die ZVT als Kombinationstherapie mit MTX stattfinden. Daraus folgen die sechs Teilpopulationen, die es gestatten, das Anwendungsgebiet A komplett zu betrachten (d.h. es gibt keine Patienten in Anwendungsgebiet A in der Indikation für Tofacitinib, die nicht durch eine der sechs Teilpopulationen umfasst wären).

Es folgt eine Auflistung der sechs in Anwendungsgebiet A zu betrachtenden Teilpopulationen:

- Teilpopulation b): Patienten bei denen keine ungünstigen Prognosefaktoren vorliegen nach unzureichendem Ansprechen auf MTX
  - ➤ Patienten der Teilpopulation b) mit MTX-Unverträglichkeit (**Teilpopulation** "b)-Mono")
  - ➤ Patienten der Teilpopulation b) ohne MTX-Unverträglichkeit (**Teilpopulation** "b)-Kombi")

- Teilpopulation c): Patienten mit ungünstigen Prognosefaktoren nach unzureichendem Ansprechen auf MTX
  - Patienten der Teilpopulation c) mit MTX-Unverträglichkeit (Teilpopulation ,,c)-Mono")
  - > Patienten der Teilpopulation c) ohne MTX-Unverträglichkeit (**Teilpopulation** "c)-Kombi")

Als Patienten "mit ungünstigen Prognosefaktoren" werden diejenigen Patienten eingestuft, bei denen entweder (zumindest) Gelenkerosionen vorliegen, die (zumindest) bezüglich der Rheumafaktoren und/oder Antikörper gegen citrullinierte Peptide positiv getestet wurden, oder die (zumindest) eine hohe Krankheitsaktivität (d.h. DAS28 > 5,1) aufweisen.

- Teilpopulation d): Patienten nach unzureichendem Ansprechen auf mehrere csDMARD (einschließlich MTX)
  - Patienten der Teilpopulation d) mit MTX-Unverträglichkeit (Teilpopulation "d)-Mono")
  - Patienten der Teilpopulation d) ohne MTX-Unverträglichkeit (**Teilpopulation** "d)-Kombi")

Für Anwendungsgebiet A konnte nur eine RCT identifiziert werden (Studie ORAL STANDARD). Aufgrund der Ein- und Ausschlusskriterien der Studie und der Patientencharakteristika wurde geprüft, für welche der zuvor aufgelisteten Teilpopulationen Ergebnisse aus der Studie abgeleitet werden können. Unmittelbar ergab sich, dass die Patienten in allen Studienarmen in Kombination mit MTX behandelt wurden. Damit kommen also nur noch die Teilpopulationen "b)-Kombi", "c)-Kombi" und "d)-Kombi" in Betracht. Eine weitere detaillierte Betrachtung der Vortherapien vor Studieneinschluss ergab, dass ein Teilkollektiv der Studie ORAL STANDARD die Kriterien für die Teilpopulation "c)-Kombi" und ein anderes Teilkollektiv die Kriterien für "d)-Kombi" erfüllt. Darüber hinaus gab es noch einige wenige Patienten, die für keines der möglichen Teilpopulationen aus Anwendungsgebiet A passend waren (in Modul 4 B wird dargestellt, dass diese Patienten auch nicht für eine Evidenzdarstellung in Anwendungsgebiet B geeignet waren). Für die Teilpopulation "b)-Kombi" konnte mithin keinerlei verfügbare Evidenz aus RCT identifiziert werden.

Bezüglich aller Teilpopulationen, für die keinerlei Evidenzbasis aus RCT mit Tofacitinib im Indikationsgebiet identifiziert werden konnte, wurde nicht nach nicht-randomisierten Studien bzw. weiterer Evidenz gesucht, da diese – selbst bei Vorhandensein – als nicht belastbar bezüglich der Begründung eines Zusatznutzens angesehen wurde. Dies betrifft die Teilpopulationen "b)-Mono", "b)-Kombi", "c)-Mono" und "d)-Mono".

Im Folgenden werden die beiden Teilkollektive der Studie ORAL STANDARD jeweils getrennt dargestellt, die zur Bewertung des Zusatznutzens für die Teilpopulationen "c)-Kombi" und "d)-Kombi" herangezogen werden können. Für die Teilpopulation c)-Kombi beträgt der Anteil der Patienten 39% der Gesamt-Studienpopulation der ORAL STANDARD, für d)-Kombi sind es 51%. Es werden jeweils nur die für die Nutzenbewertung relevanten Studienarme entsprechend der Indikation für Tofacitinib bzw. für die ausgewählte ZVT dargestellt: "Tofacitinib 5 mg BID + MTX" sowie "Adalimumab + MTX". Zur Vereinfachung wird im weiteren Verlauf des Moduls 4A jedoch nur noch von "Tofacitinib + MTX" gesprochen und automatisch vorausgesetzt, dass es sich dabei ausschließlich um die Dosis Tofacitinib 5 mg BID entsprechend der Tofacitinib-Indikation – und damit dem bestimmungsgemäßen Gebrauch – handelt.

In der folgenden Tabelle 4-13 sind die Charakteristika der Studienpatienten für die beiden Teilpopulationen (c)-Kombi und d)-Kombi) der einzigen für Anwendungsgebiet A relevanten Studie ORAL STANDARD zusammengestellt.

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

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

Studie A3921064 ORAL STANDARD	-	c)-Kombi: D in Vortherapie	Population d)-Kombi: MTX + mind. ein weiteres csDMARD in Vortherapie		
OKAL STANDARD	<b>111 (111 (112 111 111 111 111 111 111 11</b>	z m (oronorupic			
	Tofacitinib + MTX N=81	Adalimumab + MTX N=76	Tofacitinib + MTX N=102	Adalimumab + MTX N=104	
Geschlecht: n (%)					
männlich	13 (16,0)	12 (15,8)	13 (12,7)	26 (25,0)	
weiblich	68 (84,0)	64 (84,2)	89 (87,3)	78 (75,0)	
Alter [Jahre]: Mittelwert ± SD	55,2 ± 12,5	52,2 ± 12,1	52,1 ± 11,0	53,3 ± 11,0	
Alter: n (%)					
≤ 65 Jahre	62 (76,5)	64 (84,2)	89 (87,3)	93 (89,4)	
> 65 Jahre	19 (23,5)	12 (15,8)	13 (12,7)	11 (10,6)	
Gewicht [kg]: Mittelwert ± SD	$75,3 \pm 17,3$	$75,9 \pm 16,3$	$69,6 \pm 20,0$	$69,5 \pm 15,2$	
BMI [kg/m <sup>2</sup> ]: Mittelwert ± SD	$28,6 \pm 6,3$	$28,7 \pm 6,0$	$25,9 \pm 6,8$	$25.8 \pm 4.8$	
Ethnizität: n (%)					
Weiße	64 (79,0)	58 (76,3)	74 (72,5)	74 (71,2)	
Schwarze	2 (2,5)	2 (2,6)	2 (2,0)	0 (0,0)	
Asiaten	7 (8,6)	7 (9,2)	19 (18,6)	19 (18,3)	
Sonstige	8 (9,9)	9 (11,8)	7 (6,9)	11 (10,6)	
Region: n (%)					
Europa	42 (51,9)	41 (53,9)	64 (62,7)	61 (58,7)	
USA/Kanada	19 (23,5)	13 (17,1)	9 (8,8)	11 (10,6)	
Lateinamerika	11 (13,6)	10 (13,2)	7 (6,9)	11 (10,6)	
Sonstige	9 (11,1)	12 (15,8)	22 (21,6)	21 (20,2)	
Krankheitsdauer [Jahre]: Mittelwert ± SD	$5,1 \pm 6,7$	5,1 ± 6,4	$9,2 \pm 7,6$	$10,1 \pm 7,5$	
ESR [mm/Std]: Mittelwert ± SD	$48,8 \pm 23,3$	47,0 ± 22,4	47,3 ± 23,5	50,1 ± 24,9	
CRP [mg/L]: Mittelwert ± SD	$14,6 \pm 20,8$	$18,8 \pm 24,1$	$14,5 \pm 15,8$	$18,0 \pm 22,7$	

Tofacitinib (XELJANZ®)

DAS28-4(ESR): n (%)				
< 2,6	0 (0,0)	0 (0,0)	0 (0,0)	0 (0,0)
2,6-3,2	0 (0,0)	0 (0,0)	0 (0,0)	0 (0,0)
$> 3,2 - \le 5,1$	6 (7,4)	5 (6,6)	4 (3,9)	10 (9,6)
> 5,1	71 (87,7)	67 (88,2)	96 (94,1)	91 (87,5)
unbekannt	4 (4,9)	4 (5,3)	2 (2,2)	3 (2,9)
DAS28-4(ESR): Mittelwert $\pm$ SD	$6,6 \pm 1,0$	$6,4 \pm 0,9$	$6,6 \pm 0,9$	$6,3 \pm 0,9$
DAS28-4(CRP): Mittelwert $\pm$ SD	5,7 ± 1,0	5,7 ± 1,0	$5,7 \pm 0,9$	$5,5 \pm 1,0$
Anzahl empfindlicher/druckschmerzhafter Gelenke (28): Mittelwert $\pm$ SD	$16,2 \pm 7,3$	15,4 ± 7,0	$15,8\pm6,5$	$13,6 \pm 7,1$
Anzahl empfindlicher/druckschmerzhafter Gelenke (68): Mittelwert $\pm$ SD	$30,2 \pm 16,0$	$28,6 \pm 16,2$	$27,4 \pm 14,3$	25,0 ± 14,4
Anzahl geschwollener Gelenke (28): Mittelwert ± SD	$11,4 \pm 5,4$	$11,1 \pm 5,5$	$11,8 \pm 5,1$	$11,3 \pm 5,8$
Anzahl geschwollener Gelenke (66): Mittelwert ± SD	$16,3 \pm 8,8$	$15,9 \pm 8,2$	$17.0 \pm 9.0$	$16,3 \pm 8,5$
Anti-CCP: n (%)				
positiv	49 (60,5)	56 (73,7)	82 (80,4)	79 (76,0)
negativ	31 (38,3)	20 (26,3)	20 (19,6)	24 (23,1)
unbekannt	1 (1,2)	0 (0,0)	0 (0,0)	1 (1,0)
Rheumafaktor: n (%)				
positiv	51 (63,0)	47 (61,8)	70 (68,6)	74 (71,2)
negativ	26 (32,1)	27 (35,5)	32 (31,4)	30 (28,8)
unbekannt	4 (4,9)	2 (2,6)	0 (0,0)	0 (0,0)
HAQ-DI: Mittelwert ± SD	$1,4 \pm 0,6$	$1,6 \pm 0,6$	$1,5 \pm 0,6$	$1,5 \pm 0,6$

Quelle: (39)

Nach Abschluss der Studie ORAL STANDARD und nach Fertigstellung des Studienberichts wurde ein Zentrum aufgrund des Bekanntwerdens bedeutender GCP-Verstöße aus sämtlichen Wirksamkeitsanalysen ausgeschlossen. Dieses Zentrum hatte insgesamt neun Patienten in die fünf Behandlungsgruppen eingeschlossen. In den Teilpopulationen und den beiden relevanten Therapiearmen, die in diesem Dossier dargestellt sind, betrifft das konkret fünf Patienten (Tofacitinib: n=3, Adalimumab: n=2), die nachträglich aus allen Analysen zur Wirksamkeit ausgeschlossen wurden. Für die Analysen zur Sicherheit/Verträglichkeit sind die Patienten weiterhin eingeschlossen.

Die Patienten dieses Zentrums sind in allen Analysen des klinischen Studienberichts (CSR) der Studie ORAL STANDARD enthalten, so dass sich von daher eine Abweichung gegenüber dem CSR ergibt.

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.

Da nur eine RCT als Evidenzbasis für Anwendungsgebiet A zur Verfügung steht, können keine Unterschiede zwischen Studien diskutiert werden. Es wird jedoch kurz auf potentielle Unterschiede zwischen den beiden für die Bewertung des Zusatznutzens separat dargestellten Teilpopulationen dieser Studie – das sind c)-Kombi und d)-Kombi – eingegangen. Außerdem wird die Vergleichbarkeit der Studienarme innerhalb der beiden Teilpopulationen diskutiert.

Die beiden Teilpopulationen c)-Kombi und d)-Kombi unterscheiden sich hinsichtlich der mittleren Erkrankungsdauer, die in Teilpopulation d)-Kombi im Mittel 4 bis 5 Jahre länger und damit fast doppelt so groß wie in Teilpopulation c)-Kombi ist. Da das mittlere Alter zwischen den beiden Teilpopulationen ungefähr vergleichbar ist, wird daraus erkennbar, dass die Teilpopulation d)-Kombi im Vergleich zu c)-Kombi einen deutlich früheren mittleren Krankheitsbeginn aufweisen. Insofern könnte man allein aufgrund dieser wenigen Charakteristika einschätzen, dass sich die Patienten der Teilpopulation d)-Kombi von denen der Teilpopulation c)-Kombi unterscheiden.

Im Folgenden wird noch kurz auf die Vergleichbarkeit der Studienarme innerhalb der beiden Teilpopulationen eingegangen.

# Population c)-Kombi:

Die Patienten sind weitestgehend vergleichbar in beiden Studienarmen bzgl. demographischer Merkmale und Krankheitsschwere zu Studienbeginn. Im Mittel sind die Patienten im Tofacitinib-Arm um drei Jahre älter als im Adalimumab-Arm (55,2 vs. 52,2 Jahre). Die

mittleren Baseline-CRP-Werte sind im Adalimumab-Arm höher (18,8 mg/L vs. 14,6 mg/L), und der Anteil der Patienten mit positivem Anti-CCP ist ebenfalls erhöht im Adalimumab-Arm (73,7% vs. 60,5%).

# Population d)-Kombi:

Die Patienten sind weitestgehend vergleichbar in beiden Studienarmen bzgl. demographischer Merkmale und Krankheitsschwere zu Studienbeginn. Der Anteil der männlichen Patienten ist im Adalimumab-Arm etwa doppelt so hoch wie im Tofacitinib-Arm (25,0% vs. 12,7%), und die mittleren Baseline-CRP-Werte sind leicht höher im Adalimumab-Arm (18,0 vs. 14,5 mg/L).

Die Patientenpopulationen der Studie insgesamt und ihrer beiden dargestellten Teilpopulationen entsprechen dem deutschen Versorgungskontext, da die Studie überwiegend in Europa und Nordamerika durchgeführt wurde und keine Patientencharakteristika der Studienkollektive vorliegen, die auf eine Abweichung vom deutschen Versorgungskontext hindeuten würden. Die Dosierung des zu bewertenden Arzneimittels sowie der zweckmäßigen Vergleichstherapie wurden jeweils entsprechend des in Deutschland zutreffenden bestimmungsgemäßen Gebrauchs verabreicht. Insofern ist die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext gegeben.

# Analysepopulationen:

*Full Analysis Set* (FAS): Diese Population umfasst alle Patienten, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben. Des Weiteren mussten die Patienten mindestens einen Wert nach Baseline vorliegen haben für die Analyse der jeweiligen Endpunkte.

Safety Analysis Population: Diese Population umfasst alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben.

## 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 4-14: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

	gu -sg		Verblindur	ng	ige		ial
Studie	Adäquate Erzeugung der Randomisierungs sequenz	Verdeckung der Gruppenzuteilung	Patient	Behandelnde Personen	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzia auf Studienebene
A3921064							
(ORAL STANDARD)	ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Randomisierung erfolgte zentral, jedem Patienten wurde nach dem Screening-Besuch Randomisierungsnummer mit der dem Randomisierungsplan entsprechenden Behandlung zugeteilt. Die Verdeckung der Gruppenzuteilung wurde durch den Einsatz eines zentralen, unabhängigen Interactive Voice Response System (IVRS) gewährleistet.

Die Studie wurde doppelblind durchgeführt. Weder der behandelnde Arzt noch der Patient oder der Sponsor hatten Kenntnis über die tatsächlich verabreichte Medikation.

Die im Studienbericht präsentierten Ergebnisse folgen den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial der Studie beeinflussen könnten.

Das Verzerrungspotenzial auf Studienebene wird 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.

Tabelle 4-15: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Endpunkt	untersucht in Studie A3921064 (ORAL STANDARD)
DAS28-4(ESR) nach 12 Monaten	ja
DAS28-4(CRP) nach 12 Monaten	ja
EULAR-Response nach 12 Monaten	ja
ACR/EULAR-Remission und ACR Responsenach 12 Monaten	ja
SDAI nach 12 Monaten	ja
CDAI nach 12 Monaten	ja
Schmerzbeurteilung nach 12 Monaten durch den Patienten	ja
Gesamtbeurteilung der Krankheitsaktivität nach 12 Monaten durch den Patienten	ja
Gesamtbeurteilung der Krankheitsaktivität nach 12 Monaten durch den Arzt	ja
Anzahl empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten	ja
Anzahl geschwollener Gelenke nach 12 Monaten	ja
SF-36 nach 12 Monaten	ja
FACIT-F nach 12 Monaten	ja
HAQ-DI nach 12 Monaten	ja
WLQ nach 12 Monaten	ja
MOS Schlafscore nach 12 Monaten	ja
Unerwünschte Ereignisse bis 12 Monate	ja
Schwerwiegende unerwünschte Ereignisse bis 12 Monate	ja
Unerwünschte Ereignisse, die bis Monat 12 zum Therapieabbruch geführt hatten	ja
Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht- Melanom Hautkrebs) bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte bis 12 Monate	ja

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

Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie bis 12 Monate	ja
Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12 Monate	ja

#### **4.3.1.3.1** DAS28-4(ESR) nach 12 Monaten – RCT

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

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

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

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

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

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

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

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

Tabelle 4-16: Operationalisierung von DAS28-4(ESR) nach 12 Monaten

Studie	Operationalisierung
A3921064	Der DAS28-4(ESR) ist ein gewichteter Score für die Erfassung der Krankheitsaktivität, der sich aus den folgenden Einzelkomponenten zusammensetzt:
	1. Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 28 Gelenken)
	<ul><li>2. Anzahl geschwollener Gelenke (von insgesamt 28 Gelenken)</li><li>3. ESR-Wert</li></ul>
	4. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
	Die Einzelkomponenten und somit auch der DAS28-4(ESR) wurden zu jedem Studienzeitpunkt erhoben. Der DAS28-4(ESR) hat einen Wertebereich von 0 (keinerlei Krankheitsaktivität) bis 9,4 (höchste Krankheitsaktivität).
	Es werden Responderanalysen zu Monat 12 mit zwei verschiedenen Schwellenwerten dargestellt. Ein Patient gilt demnach als Responder, falls sein DAS28-4(ESR) < 2,6 bzw. ≤ 3,2 nach 12 Monaten Studiendauer beträgt. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.
	Als Sensitivitätsanalysen werden zusätzlich die Effektschätzer gezeigt, die sich ohne Imputation ergeben sowie die, die mittels LOCF-Imputation resultieren. Für die Teilpopulation d)-Kombi werden zudem sämtliche Ergebnisse der durchgeführten Sensitivitätsanalysen gezeigt, um die Robustheit dieses speziellen Ergebnisses ausführlicher darzulegen.
	Alle Effektschätzer werden inklusive ihres 95%-Konfidenzintervalls und p-Werts angegeben.
	Der Endpunkt wird für die FAS Population analysiert. ESR wurde nur in solchen Zentren erhoben, bei denen die Voraussetzungen für eine direkte Weiterleitung des Testergebnisses an das Zentrallabor gegeben waren, um die Verblindung der Studie zu gewährleisten. Daher ist die Zahl der Patienten für diesen Endpunkt etwas niedriger als das eigentliche FAS Kollektiv (> 30% der Patienten fehlen in der Analyse). Dieser Aspekt führt dazu, dass dieser Endpunkt im Folgenden als hochverzerrt eingestuft wird und dass eine sinnvolle Interpretation der Ergebnisse des DAS-28-4(ESR) nicht möglich ist.

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 4-17: Bewertung des Verzerrungspotenzials für DAS28-4(ESR) nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	unklar	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Der kombinierte Endpunkt setzt sich aus vier Einzelkomponenten zusammen, die bezüglich des Verzerrungspotenzials einheitlich eingestuft werden können. Auf Studienebene wurde das Verzerrungspotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. ESR wurde nur in solchen Zentren erhoben, bei denen die Voraussetzungen für eine direkte Weiterleitung des Testergebnisses an das Zentrallabor gegeben waren, um die Verblindung der Studie zu gewährleisten. Es wird eingeschätzt, dass der hohe Anteil von Patienten, für die keine Ergebnisse zu diesem Endpunkt vorliegen (> 30% des FAS in den jeweiligen Studienarmen) dazu führt, dass eine adäquate Umsetzung des ITT-Prinzips nicht mehr gegeben ist und dadurch eine sinnvolle Interpretation der Ergebnisse des DAS28-4(ESR) nicht möglich ist.

Des Weiteren existieren keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als hoch eingestuft.

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

Tabelle 4-18: Ergebnisse für DAS28-4(ESR) < 2,6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	
	Anzahl Responder /N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	)-Kombi: nur e	in csDMARD in	Vortherapie			
A3921064	9/70 (12,9)		11/64 (17,2)		0,8 [0,3-1,7], p=0,484 0,7 [0,3-1,9], p=0,484 -4,3% [-16,5%-7,8%], p=0,484	
		NA [NA-NA]		NA [NA-NA]	1,2 [0,6-2,5], p=0,631	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e		
A3921064	9/91 (9,9)		20/93 (21,5)		0,5 [0,2-0,96], p=0,038 0,4 [0,2-0,9], p=0,034 -11,6% [-22,0%;-1,3%], p=0,028	
		NA [NA-NA]		NA [NA-NA]	0,6 [0,3-1,2], p=0,170	
Ū	n: RR = Re		o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-	

# Population c)-Kombi:

Die Responderraten sind höher unter Adalimumab + MTX als unter Tofacitinib + MTX (17,2% vs. 12,9%). Der Unterschied ist jedoch nicht statistisch signifikant.

#### Population d)-Kombi:

Die Responderraten sind unter Adalimumab + MTX höher als unter Tofacitinib + MTX (21,5% vs. 9,9%). Der Unterschied ist statistisch signifikant zugunsten von Adalimumab + MTX (p=0,038).

Eine sinnvolle Interpretation dieses Endpunkts ist für beide Teilpopulationen jedoch nicht möglich aufgrund der niedrigen Patientenzahl in Bezug zum FAS, welche ein hohes Verzerrungspotenzial zur Folge hat.

Abbildung 2: Kaplan-Meier Kurven für den Endpunkt DAS28-4(ESR) < 2,6 nach 12 Monaten (Population c)-Kombi)

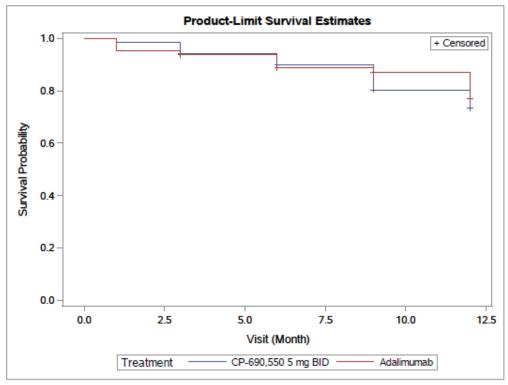


Abbildung 3: Kaplan-Meier Kurven für den Endpunkt DAS28-4(ESR) < 2,6 nach 12 Monaten (Population d)-Kombi)

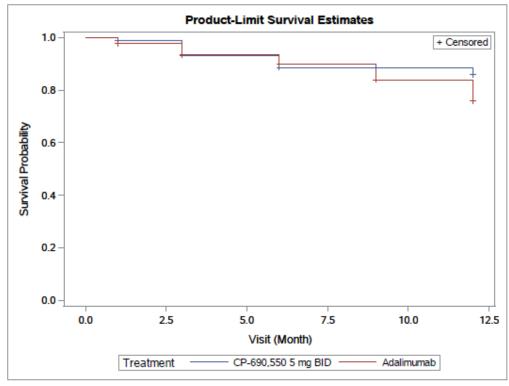


Tabelle 4-19: Ergebnisse für DAS28-4(ESR) ≤ 3,2 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab ATX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	
	Anzahl Responder /N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	-Kombi: nur ei	in csDMARD in	Vortherapie			
A3921064	16/70		19/64		0,8 [0,4-1,4], p=0,370	
	(22,9)		(29,7)		0,7 [0,3-1,5], p=0,370	
					-6,8% [-21,7%-8,1%], p=0,369	
		NA		NA	1,0 [0,6-1,8], p=0,925	
		[9,0-NA]		[12,0-NA]		
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e		
A3921064	17/91		24/93		0,7 [0,4-1,3], p=0,250	
	(18,7)		(25,8)		0,7 [0,3-1,3], p=0,247	
					-7,1% [-19,1%-4,8%], p=0,243	
		NA		NA	1,0 [0,6-1,6], p=0,984	
		[12-NA]		[NA-NA]	- •	
Quelle: (40)						
U			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%	

# Population c)-Kombi:

Die Responderraten sind höher unter Adalimumab + MTX als unter Tofacitinib + MTX (29,7% vs. 22,9%). Der Unterschied ist jedoch nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderraten sind höher unter Adalimumab + MTX als unter Tofacitinib + MTX (25,8% vs. 18,7%). Der Unterschied ist jedoch nicht statistisch signifikant.

Eine sinnvolle Interpretation dieses Endpunkts ist für beide Teilpopulationen jedoch nicht möglich aufgrund der niedrigen Patientenzahl in Bezug zum FAS, welche ein hohes Verzerrungspotenzial zur Folge hat.

Abbildung 4: Kaplan-Meier Kurven für den Endpunkt DAS28-4(ESR)  $\leq$  3,2 nach 12 Monaten (Population c)-Kombi)

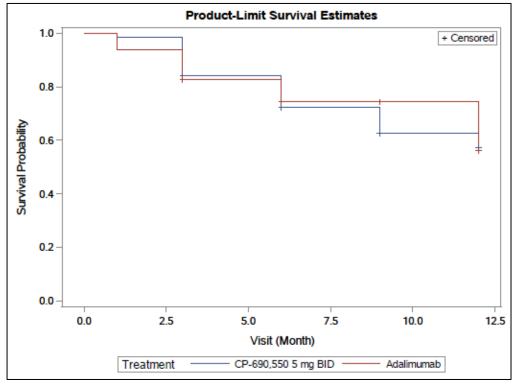


Abbildung 5: Kaplan-Meier Kurven für den Endpunkt DAS28-4(ESR) ≤ 3,2 nach 12 Monaten (Population d)-Kombi)

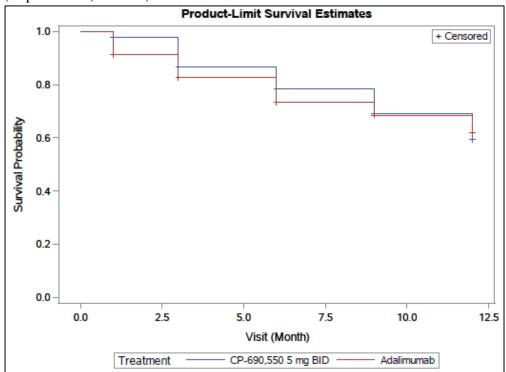


Tabelle 4-20: Sensitivitätsanalysen für DAS28-4(ESR) < 2,6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib +MTX	Adalimumab +MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)-K	ombi: nur ein csDMARD in V	ortherapie	
keine Imputation	9/53 (17,0)	11/50 (22,0)	0,8 [0,3-1,7], p=0,522 0,7 [0,3-1,9], p=0,521 -5,0% [-20,3%-10,3%], p=0,520
LOCF	14/70 (20,0)	11/64 (17,2)	1,2 [0,6-2,4], p=0,677 1,2 [0,5-2,9], p=0,677 2,8% [-10,4%-16,0%], p=0,675
Population d)-K	ombi: mehr als ein csDMARD	in Vortherapie	
keine Imputation	9/71 (12,7)	19/74 (16,7)	0,5 [0,4-1,02], p=0,056 0,4 [0,2-1,01], p=0,051 -13,0% [-25,6%;-0,4%], p=0,043
LOCF	9/91 (9,9)	20/93 (21,5)	0,5 [0,2-0,96], p=0,038 0,4 [0,2-0,9], p=0,034 -11,6% [-22,0%;-1,3%], p=0,028
"reine" NRI- Imputation 1	9/102 (8,8)	19/104 (18,3)	0,5 [0,2-1,02], p=0,055 0,4 [0,2-1,01], p=0,052 -9,5% [-18,7%;-0,2%], p=0,045
"reine" NRI- Imputation 2	9/93 (9,7)	19/94 (20,2)	0,5 [0,2-1,0], p=0,051 0,4 [0,2-0,99], p=0,048 -10,5% [-20,6%;-0,4%], p=0,041
"ZVT" Imputations- verfahren	17/102 (16,7)	27/104 (26,0)	0,6 [0,4-1,1], p=0,109 0,6 [0,3-1,1], p=0,106 -9,3% [-20,4%-1,8%], p=0,101
"Worst- case" Imputations- verfahren	13/102 (12,8)	23/104 (22,1)	0,6 [0,3-1,1], p=0,083 0,5 [0,2-1,1], p=0,080 -9,4% [-19,6%-0,9%], p=0,074
lineare Interpolation	11/78 (14,1)	22/84 (26,2)	0,5 [0,3-1,04], p=0,064 0,5 [0,2-1,03], p=0,060 -12,1% [-24,3%-0,1%], p=0,052
_		OR = Odds Ratio, 95%-KI ried Forward, NA = nicht ange	= 95%-Konfidenzintervall, MTX ebbar

Tofacitinib (XELJANZ®)

#### Population c)-Kombi:

Die Analyse ohne imputierte Werte liefert ebenfalls einen numerischen Trend zugunsten von Adalimumab + MTX. Mit LOCF kehrt sich dieser Trend zugunsten von Tofacitinib + MTX um. Beide beobachteten Unterschiede zwischen den Therpiearmen sind nicht statistisch signifikant.

# Population d)-Kombi:

Für alle der sieben gerechneten Sensitivitätsanalysen zeigen sich jeweils höhere Responderraten im Adalimumab-Arm, was auch in der primären Analyse (NRI) zu sehen war. Lediglich mittels LOCF ist dieser Unterschied zudem statistisch signifikant (p=0,038). In allen anderen Fällen liegen jeweils keine statistischen Signifikanzen vor.

Eine sinnvolle Interpretation dieses Endpunkts ist für beide Teilpopulationen jedoch nicht möglich aufgrund der niedrigen Patientenzahl in Bezug zum FAS, welche ein hohes Verzerrungspotenzial zur Folge hat.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

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** DAS28-4(CRP) nach 12 Monaten – RCT

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

Tabelle 4-21: Operationalisierung von DAS28-4(CRP) nach 12 Monaten

Studie	Operationalisierung
A3921064	Der DAS28-4(CRP) ist ein gewichteter Score für die Erfassung der Krankheitsaktivität, der sich aus den folgenden Einzelkomponenten zusammensetzt:
	1. Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 28 Gelenken)
	2. Anzahl geschwollener Gelenke (von insgesamt 28 Gelenken)
	3. CRP-Wert
	4. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
	Die Einzelkomponenten und somit auch der DAS28-4(CRP) wurden zu jedem Studienzeitpunkt erhoben. Der DAS28-4(CRP) hat einen Wertebereich von 0 (keinerlei Krankheitsaktivität) bis 9,4 (höchste Krankheitsaktivität).
	Es werden Responderanalysen zu Monat 12 mit zwei verschiedenen Schwellenwerten dargestellt. Ein Patient gilt demnach als Responder, falls sein DAS28-4(CRP) < 2,6 bzw. ≤ 3,2 nach 12 Monaten Studiendauer beträgt. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.
	Der Endpunkt wird für die FAS Population analysiert.

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

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

Tabelle 4-22: Bewertung des Verzerrungspotenzials für DAS28-4(CRP) nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der kombinierte Endpunkt setzt sich aus vier Einzelkomponenten zusammen, die bezüglich des Verzerrungspotenzials einheitlich eingestuft werden können. Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-23: Ergebnisse für DAS28-4(CRP) < 2,6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	in csDMARD in	Vortherapie		
A3921064	25/79		21/75		1,1 [0,7-1,8], p=0,622
	(31,6)		(28,0)		1,2 [0,6-2,4], p=0,621
					3,7% [-10,8%-18,1%], p=0,621
		12,0		NA	1,3 [0,8-2,2], p=0,236
		[9,0-NA]		[12,0-NA]	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	26/100		28/103		1,0 [0,6-1,5], p=0,849
	(26,0)		(27,2)		0,9 [0,5-1,8], p=0,849
					-1,2% [-13,3%-11,0%], p=0,849
		NA		NA	1,2 [0,8-1,8], p=0,462
		[9,0-NA]		[12,0-NA]	
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

# Population c)-Kombi:

Die Responderraten sind höher unter Tofacitinib + MTX als unter Adalimumab + MTX (31,7% vs. 28,0%). Der Unterschied ist jedoch nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderraten sind vergleichbar in beiden Studienarmen (26,0% bzw. 27,2%). Es gibt keinen statistisch signifikanten Unterschied.

Abbildung 6: Kaplan-Meier Kurven für den Endpunkt DAS28-4(CRP) < 2,6 nach 12 Monaten (Population c)-Kombi)

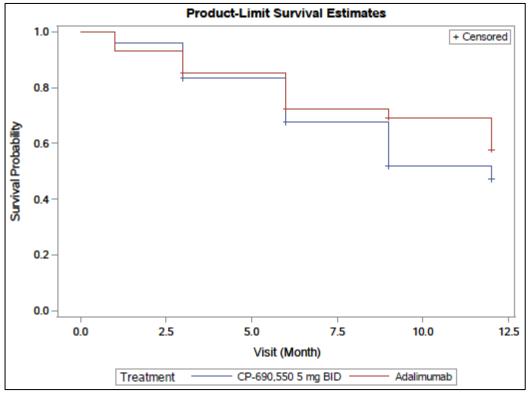
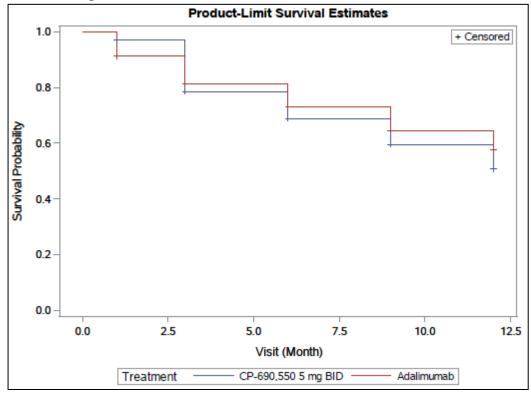


Abbildung 7: Kaplan-Meier Kurven für den Endpunkt DAS28-4(CRP) < 2,6 nach 12 Monaten (Population d)-Kombi)



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

Tabelle 4-24: Ergebnisse für DAS28-4(CRP) ≤ 3,2 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			imumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/ N	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population	c)-Kombi: nur e	in csDMARD in	Vortherapie		
A392106 4	40/79 (50,6)		34/75 (45,3)		1,1 [0,8-1,6], p=0,512 1,2 [0,7-2,3], p=0,511 5,3% [-10,5%-21,1%], p=0,510
		6,0 [6,0-9,0]		9,0 [6,0-12,0]	1,4 [0,9-2,1], p=0,117
Population	d)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A392106 4	44/100 (44,0)		44/103 (42,7)		1,0 [0,8 -1,4], p=0,854 1,1 [0,6-1,8], p=0,854 1,3% [-12,4%-14,9%], p=0,854
		6,0 [6,0-9,0]		9,0 [6,0-12,0]	1,1 [0,7-1,5], p=0,785
_	gen: RR = Re		o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

# Population c)-Kombi:

Die Responderraten sind höher unter Tofacitinib + MTX als unter Adalimumab + MTX (50,6% vs. 45,3%). Der Unterschied ist jedoch nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderraten sind vergleichbar in beiden Studienarmen (44,0% bzw. 42,7%). Es gibt keinen statistisch signifikanten Unterschied.

Abbildung 8: Kaplan-Meier Kurven für den Endpunkt DAS28-4(CRP)  $\leq$  3,2 nach 12 Monaten (Population c)-Kombi)

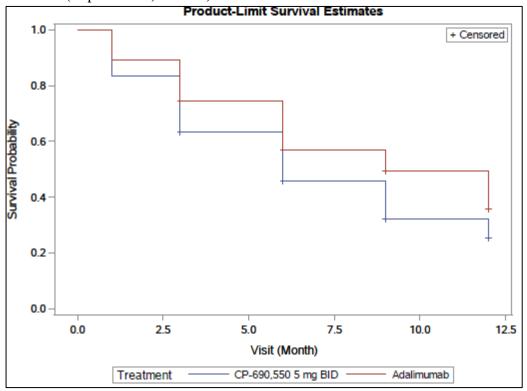
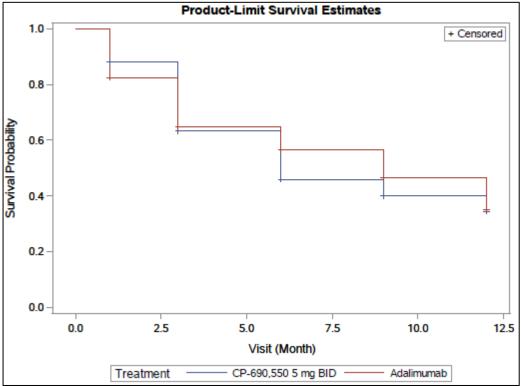


Abbildung 9: Kaplan-Meier Kurven für den Endpunkt DAS28-4(CRP)  $\leq$  3,2 nach 12 Monaten (Population d)-Kombi)



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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

**Operationalisierung** 

Studie

### 4.3.1.3.3 EULAR-Response nach 12 Monaten – RCT

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

Tabelle 4-25: Operationalisierung von EULAR-Response nach 12 Monaten

200000	o per automanister	8				
A3921064	einer der drei Kat guteVerbesserung	EULAR-Response misst die Veränderung der Krankheitsaktivität im Verlauf. Ein Patient wird einer der drei Kategorien eingestuft: keine Verbesserung, moderate Verbesserung, gute Verbesserung (jeweils gegenüber Baseline). Die drei Kategorien sind anhand des folgenden Schemas definiert:				
	aktueller DAS28-4 Wert	DAS-28-4 Änderung gegenüber Baseline				
	DAS20-4 Welt	> 1,2	$> 0.6 \text{ und} \le 1.2$	≤ 0,6		
	≤ 3,2	gute Verbesserung	moderate Verbesserung	keine Verbesserung		
	$>$ 3,2 und $\leq$ 5,1 moderate Verbe		moderate Verbesserung	keine Verbesserung		
	>5,1	moderate Verbesserung keine Verbesserung keine Verbesserung				

Der DAS28-4 wurde zu jedem Studienzeitpunkt erhoben (s. Abschnitte zu DAS28-4). Er wurde mittels zweier verschiedener Operationalisierungen ausgewertet: ESR und CRP, folglich wird auch die EULAR-Response in diesem Abschnitt für beide Varianten dargestellt.

Es werden Responderanalysen zu Monat 12 für "gute Verbesserung" sowie für "gute oder moderate Verbesserung" dargestellt. Fehlende Werte werden gemäß der sog. *Nonresponder Imputation* (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.

Zusätzlich werden Hazard Ratios für *Time-to-Event* Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis (*event*) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.

Als Sensitivitätsanalysen werden zusätzlich die Effektschätzer gezeigt, die sich ohne Imputation ergeben sowie die, die mittels LOCF-Imputation resultieren.

Der Endpunkt wird für die FAS Population analysiert.

ESR wurde nur in solchen Zentren erhoben, bei denen die Voraussetzungen für eine direkte Weiterleitung des Testergebnisses an das Zentrallabor gegeben waren, um die Verblindung der Studie zu gewährleisten. Daher ist die Zahl der Patienten für die EULAR Response mittels ESR deutlich niedriger als für die Response mittels CRP bzw. als das eigentliche FAS Kollektiv (< 30% Unterschied). Dieser Aspekt führt dazu, dass der Endpunkt EULAR-Response(ESR) im Folgenden als hochverzerrt eingestuft wird und dass eine sinnvolle Interpretation der Ergebnisse des DAS-28-4(ESR) nicht möglich ist.

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

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

Tabelle 4-26: Bewertung des Verzerrungspotenzials für EULAR-Response nach 12 Monaten 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
A3921064	niedrig	ja	unklar (ESR)	ja	ja	hoch (ESR)
A3921064	niedrig	ja	ja (CRP)	ja	ja	niedrig (CRP)

Begründen Sie für jede Studie die abschließende Einschätzung.

Die kombinierten Endpunkte DAS28-4 setzen sich aus jeweils 4 Einzelkomponenten zusammen, die bezüglich des Verzerrungspotenzials einheitlich eingestuft werden können. Der Endpunkt DAS28-4(ESR) wurde auf Endpunktebene als mit hohem Verzerrungspotenzial behaftet eingestuft, der DAS28-4(CRP) als mit niedrigem Verzerrungspotenzial behaftet. Das Kriterium EULAR-Response ändert nichts an dieser Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung der EULAR-Response mittels CRP wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt.

ESR konnte nur in solchen Zentren erhoben werden, bei denen die Voraussetzungen für eine direkte Weiterleitung des Testergebnisses an das Zentrallabor gegeben waren, um die Verblindung der Studie zu gewährleisten. Es wird eingeschätzt, dass der hohe Anteil von Patienten, für die keine Ergebnisse zu diesem Endpunkt vorliegen (> 30% des FAS in den jeweiligen Studienarmen) dazu führt, dass eine adäquate Umsetzung des ITT-Prinzips nicht mehr gegeben ist und dadurch eine sinnvolle Interpretation der Ergebnisse der EULAR-Response bei Verwendung der ESR nicht möglich ist.

Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich für die Operationalisierung mittels CRP als niedrig eingestuft, mittels ESR als hoch.

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 EULAR-Response (ESR) nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib +MTX			imumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	)-Kombi: nur e	in csDMARD in	Vortherapie			
gute Ver- besserung	16/70		19/64		0,8 [0,4-1,4], p=0,370	
o esserang	(22,9)		(29,7)		0,7 [0,3-1,5], p=0,370 -6,8% [-21,7%-8,1%], p=0,369	
		NA [9,0-NA]		NA [12,0-NA]	1,0 [0,6-1,8], p=0,943	
gute oder	50/70		41/64		1,1 [0,9-1,4], p=0,366	
moderate	(71,4)		(64,1)		1,4 [0,7-2,9], p=0,362	
Ver- besserung					7,4% [-8,5%-23,2%], p=0,361	
		1,0 [1,0-3,0]		1,0 [1,0-3,0]	1,0 [0,7-1,4], p=1,000	
Population d	)-Kombi: mehi	als ein csDMA	RD in Vortherapi	e		
gute Ver-	17/91		24/93		0,7 [0,4-1,3], p=0,250	
besserung	(18,7)		(25,8)		0,7 [0,3-1,3], p=0,247	
					-7,1% [-19,1%-4,8%], p=0,243	
		NA [12,0-NA]		NA [NA-NA]	1,0 [0,6-1,6], p=0,998	
gute oder	63/91		63/93		1,0 [0,8-1,2], p=0,828	
moderate	(69,2)		(67,7)		1,1 [0,6-2,0], p=0,828	
Ver- besserung					1,5% [-11,9%-14,9%], p=0,828	
		3,0		1,0	0,9 [0,6-1,1], p=0,274	
		[1,0-3,0]		[1,0-3,0]		
Quelle: (40)						
			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%	

# <u>Population c)-Komb</u>i:

Die Responderrate für "gute Verbesserung" sind höher im Adalimumab-Arm (29,7% vs. 22,9%); für "gute oder moderate Verbesserung" ist die Rate höher im Tofacitinib-Arm (71,4% vs. 64,1%). Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

### Population d)-Kombi:

Für "gute Verbesserung" liegt eine höhere Responderrate unter Adalimumab + MTX vor (25,8% vs. 18,7%). Für "gute oder moderate Verbesserung" sind die Raten vergleichbar in beiden Studienarmen (69,2% bzw. 67,7%). Es liegen keine statistisch signifikanten Unterschiede vor.

Eine sinnvolle Interpretation dieses Endpunkts ist für beide Teilpopulationen jedoch nicht möglich aufgrund der niedrigen Patientenzahl in Bezug zum FAS, welche ein hohes Verzerrungspotenzial zur Folge hat.

Abbildung 10: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (ESR) nach 12 Monaten ("gute Verbesserung", Population c)-Kombi)

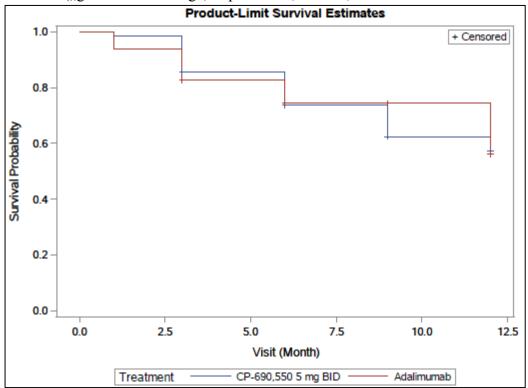


Abbildung 11: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (ESR) nach 12 Monaten ("gute Verbesserung", Population d)-Kombi)

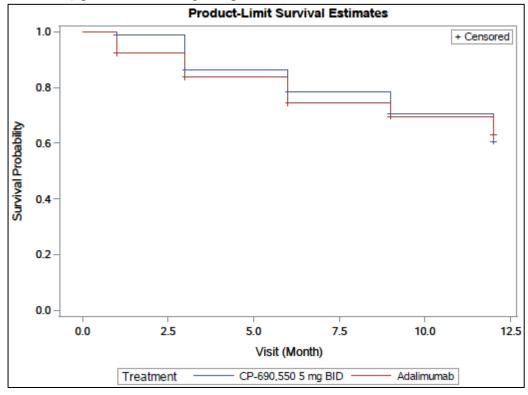


Abbildung 12: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (ESR) nach 12 Monaten ("gute oder moderate Verbesserung", Population c)-Kombi)

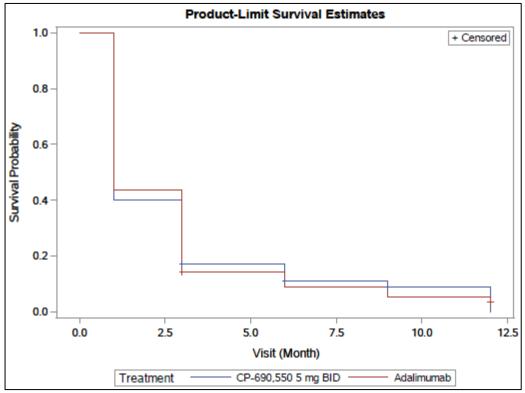


Abbildung 13: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (ESR) nach 12 Monaten ("gute oder moderate Verbesserung", Population d)-Kombi)

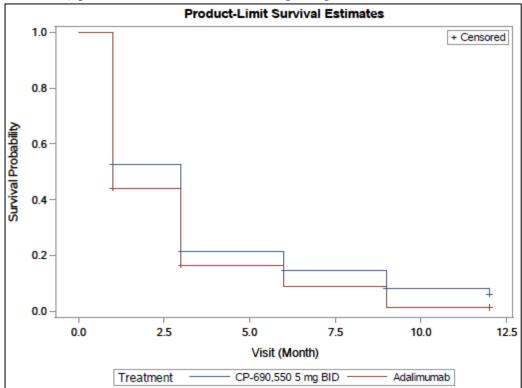


Tabelle 4-28: Sensitivitätsanalysen für EULAR-Response (ESR) nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064		Tofacitinib +MTX	Adalimumab +MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
		Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)	-Kombi: nur MTX in Vorther	apie		
keine	gute Verbesserung	16/53 (30,2)	18/50 (36,0)	0,8 [0,5-1,5], p=0,532 0,8 [0,3-1,8], p=0,531 -5,8% [-24,0%-12,4%], p=0,531
Imputation	gute oder moderate Verbesserung	49/53 (92,5)	40/50 (80,0)	1,2 [0,99-1,4], p=0,074 3,1 [0,9-10,5], p=0,075 12,5% [-0,7%-25,6%], p=0,064
	gute Verbesserung	22/70 (31,4)	20/64 (31,3)	1,0 [0,6-1,7], p=0,982 1,0 [0,5-2,1], p=0,982 0,2% [-15,6%-15,9%], p=0,982
LOCF	gute oder moderate Verbesserung	61/70 (87,1)	48/64 (75,0)	1,2 [0,98-1,4], p=0,079 2,3 [0,9-5,6], p=0,076 12,1% [-1,1%-25,3%], p=0,071
Population d	)-Kombi: mehr als ein csDMA	RD in Vortherapi	e	
keine	gute Verbesserung	17/71 (23,9)	23/74 (31,1)	0,8 [0,5-1,3], p=0,340 0,7 [0,3-1,5], p=0,338 -7,1% [-21,6%-7,3%], p=0,334
Imputation	gute oder moderate Verbesserung	62/71 (87,3)	63/74 (85,1)	1,0 [0,9-1,2], p=0,702 1,2 [0,5-3,1], p=0,703 2,2% [-9,0%-13,4%], p=0,702
	gute Verbesserung	18/91 (19,8)	28/93 (30,1)	0,7 [0,4-1,1], p=0,111 0,6 [0,3-1,1], p=0,108 -10,3% [-22,7%-2,1%], p=0,103
LOCF	gute oder moderate Verbesserung	71/91 (78,0)	76/93 (81,7)	1,0 [0,8-1,1], p=0,532 0,8 [0,4-1,6], p=0,532 -3,7% [-15,3%-7,9%], p=0,531
Quelle: (40)				

Quelle: (40)

Abkürzungen: RR = Relatives Risiko, OR = Odds Ratio, 95%-KI = 95%-Konfidenzintervall, MTX =

Methotrexat, LOCF = Last Observation Carried Forward, NA = nicht angebbar

### Population c)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

### Population d)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

Eine sinnvolle Interpretation dieses Endpunkts ist für beide Teilpopulationen jedoch nicht möglich aufgrund der niedrigen Patientenzahl in Bezug zum FAS, welche ein hohes Verzerrungspotenzial zur Folge hat.

Tabelle 4-29: Ergebnisse für EULAR-Response (CRP) nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib +MTX			imumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	)-Kombi: nur e	in csDMARD in	Vortherapie			
gute Ver- besserung	39/79 (49,4)		33/75 (44,0)		1,1 [0,8-1,6], p=0,506 1,2 [0,7-2,3], p=0,505 5,4% [-10,4%-21,1%], p=0,504	
		6,0 [6,0-9,0]		9,0 [6,0-12,0]	1,3 [0,9-2,0], p=0,157	
gute oder	53/79		52/75		1,0 [0,8-1,2], p=0,765	
moderate Ver- besserung	(67,1)		(69,3)		0,9 [0,5-1,8], p=0,765 -2,2% [-17,0%-12,5%], p=0,765	
		1,0 [NA-NA]		1,0 [NA-NA]	1,0 [0,7-1,3], p=0,759	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e		
gute Ver- besserung	42/99 (42,4)		41/103 (39,8)		1,1 [0,8-1,5], p=0,705 1,1 [0,6-2,0], p=0,705 2,6% [-11,0%-16,2%], p=0,705	
		6,0 [6,0-9,0]		9,0 [6,0-12,0]	1,1 [0,8-1,6], p=0,515	
gute oder	70/99		79/103		0,9 [0,8-1,1], p=0,336	
moderate Ver- besserung	(70,7)		(76,7)		0,7 [0,4-1,4], p=0,334 -6,0% [-18,1%-6,1%], p=0,333	
		1,0 [NA-NA]		1,0 [NA-NA]	1,0 [0,8-1,3], p=0,979	
_	n: RR = Re		o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%	

### Population c)-Kombi:

Für "gute Verbesserung" liegt eine höhere Responderrate unter Adalimumab + MTX vor (44,0% vs. 30,4%). Für "gute oder moderate Verbesserung" sind die Raten vergleichbar in beiden Studienarmen (67,1% vs. 69,3%). Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

## Population d)-Kombi:

Für "gute Verbesserung" liegt eine höhere Responderrate unter Tofacitinib + MTX vor (42,4% vs. 39,8%). Für "gute oder moderate Verbesserung" ist die Rate höher unter Adalimumab + MTX (76,7% vs. 70,7%). Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

Abbildung 14: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (CRP) nach 12 Monaten ("gute Verbesserung", Population c)-Kombi)

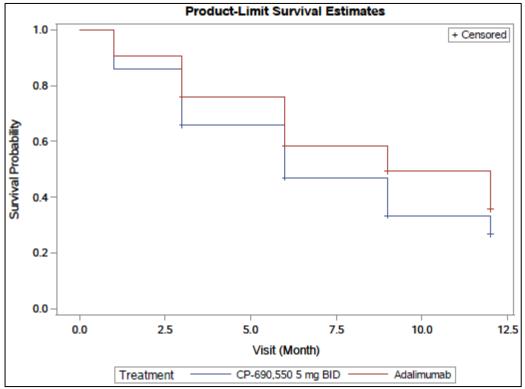


Abbildung 15: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (CRP) nach 12 Monaten ("gute Verbesserung", Population d)-Kombi)

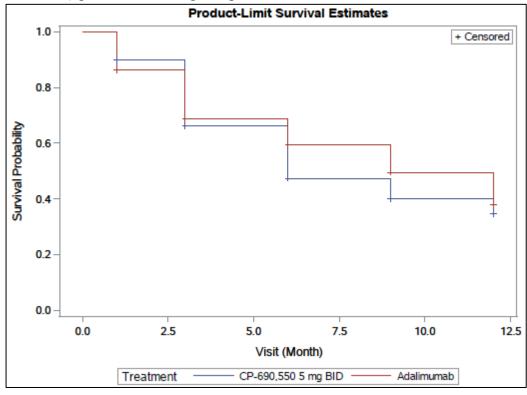


Abbildung 16: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (CRP) nach 12 Monaten ("gute oder moderate Verbesserung", Population c)-Kombi)

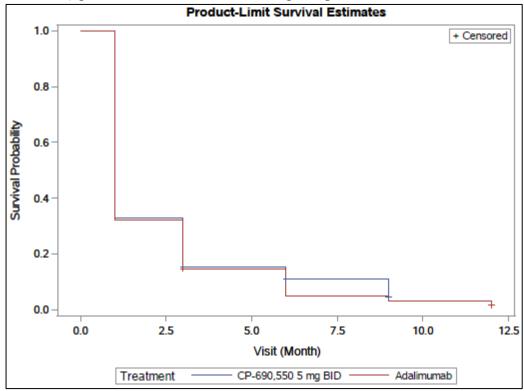
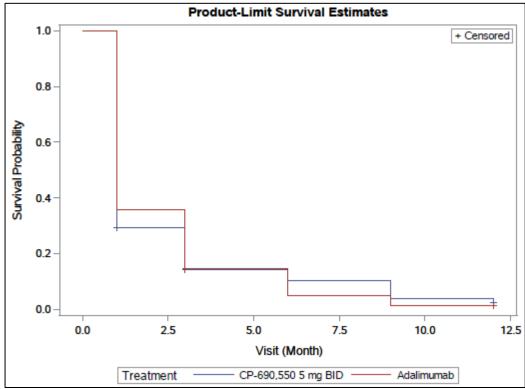


Abbildung 17: Kaplan-Meier Kurven für den Endpunkt EULAR-Response (CRP) nach 12 Monaten ("gute oder moderate Verbesserung", Population d)-Kombi)



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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

### 4.3.1.3.4 ACR/EULAR-Remission und ACR Response nach 12 Monaten – RCT

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

Tabelle 4-30: Operationalisierung von ACR Response und Remission

Studie	Operationalisierung
A3921064	Der ACR ist ein Instrument, welches vom American College of Rheumatology definiert wurde zur Beurteilung der Wirksamkeit. Der ACR existiert in verschiedenen Varianten (hier ACR50 und ACR70), allgemein berechnet er sich aus den folgenden Einzelkomponenten:
	1. Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 68 Gelenken)
	2. Anzahl geschwollener Gelenke (von insgesamt 66 Gelenken)
	3. Schmerzbeurteilung durch den Patienten (visuelle Analogskala)
	4. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
	5. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PhysGA)
	6. CRP-Wert
	7. HAQ-DI
	Ein Ansprechen bzgl. des ACR50 liegt vor, wenn der Patient in den ersten der beiden Einzelkomponenten eine mind. 50%-ige Verbesserung erfährt sowie jeweils mind. eine 50%-ig Verbesserung in mind. 3 der 5 weiteren Einzelkomponenten. Die Definition des ACR70 gilt analog.
	Die Erhebung aller Einzelkomponenten und des ACR erfolgte zu jedem Studienzeitpunkt.
	Eine Remission gemäß ACR/EULAR-Kriterien war in diesem Fall definiert, wenn in den obigen Kriterien 1., 2., 4. und 6. eine Verbesserung vorlag (Anzahl Gelenke jeweils $\leq$ 1, CRP $\leq$ 1 mg/dl, PatGA $\leq$ 10).
	Es werden Responderanalysen zu Monat 12 sowohl für den ACR50 als auch für den ACR70 und ACR/EULAR-Remission berichtet. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.
	Für die ACR/EULAR-Remission werden als Sensitivitätsanalysen zusätzlich die Effektschätzer gezeigt, die sich ohne Imputation ergeben sowie die, die mittels LOCF-Imputation resultieren.
	Der Endpunkt wird für die FAS Population analysiert.

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 4-31: Bewertung des Verzerrungspotenzials für ACR/EULAR-Remission und ACR Response nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der kombinierte Endpunkt setzt sich aus sieben Einzelkomponenten zusammen, die bezüglich des Verzerrungspotenzials einheitlich eingestuft werden können. Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-32: Ergebnisse für ACR70 Response nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	in csDMARD in	Vortherapie		
A3921064	22/79		12/75		1,7 [0,9-3,3], p=0,084
	(27,8)		(16,0)		2,0 [0,9-4,5], p=0,080
					11,9% [-1,1%-24,8%], p=0,072
		NA		NA	1,5 [0,9-2,6], p=0,134
		[12,0-NA]		[NA-NA]	
Population d	)-Kombi: mehr	als ein csDMAl	RD in Vortherapi	e	
A3921064	24/100		20/103		1,2 [0,7-2,1], p=0,430
	(24,0)		(19,4)		1,3 [0,7-2,6], p=0,429
					4,6% [-6,8%-15,9%], p=0,428
		NA		NA	1,5 [0,9-2,4], p=0,129
		[12,0-NA]		[NA-NA]	
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

### Population c)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (27,9% vs. 16,0%). Der Unterschied ist jedoch nicht statistisch signifikant.

#### Population d)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (24,0% vs. 19,4%). Der Unterschied ist jedoch nicht statistisch signifikant.

Abbildung 18: Kaplan-Meier Kurven für den Endpunkt ACR70 Response nach 12 Monaten (Population c)-Kombi)

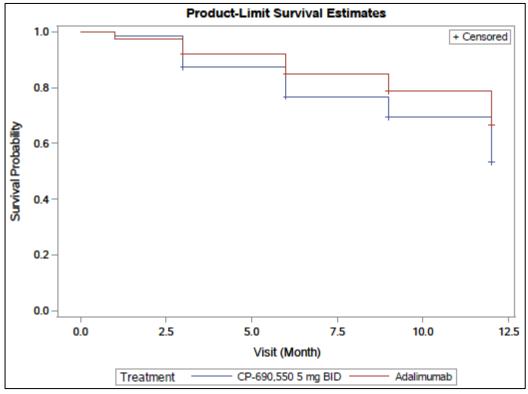


Abbildung 19: Kaplan-Meier Kurven für den Endpunkt ACR70 Response nach 12 Monaten (Population d)-Kombi)

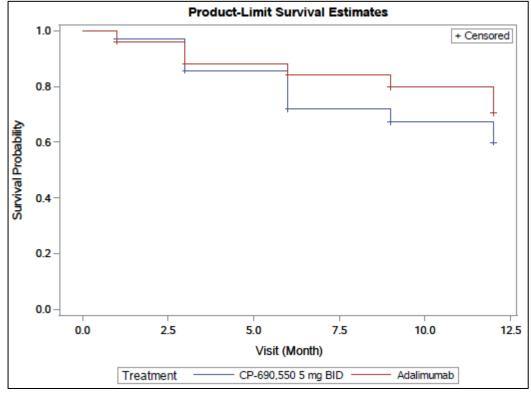


Tabelle 4-33: Ergebnisse für ACR50 Response nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			imumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder /N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	-Kombi: nur e	in csDMARD in	Vortherapie		
A3921064	37/79 (46,8)		26/75 (34,7)		1,4 [0,9-2,0], p=0,130 1,7 [0,9-3,2], p=0,126
	(40,8)		(34,7)		12,2% [-3,2%-27,6%], p=0,121
		6,0		9,0	1,4 [0,9-2,0], p=0,143
		[6,0-9,0]		[6,0-12,0]	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	42/100		43/103		1,0 [0,7-1,4], p=0,971
	(42,0)		(41,7)		1,0 [0,6-1,8], p=0,971
					0,3% [-13,3%-13,8%], p=0,971
		6,0		9,0	1,0 [0,7-1,4], p=0,917
		[3,0-12,0]		[6,0-12,0]	
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

# Population c)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (46,8% vs. 34,7%). Der Unterschied ist jedoch nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderraten sind vergleichbar in beiden Studienarmen (42,0% bzw. 41,7%). Es liegt kein statistisch signifikanter Unterschied vor.

Abbildung 20: Kaplan-Meier Kurven für den Endpunkt ACR50 Response nach 12 Monaten (Population c)-Kombi)

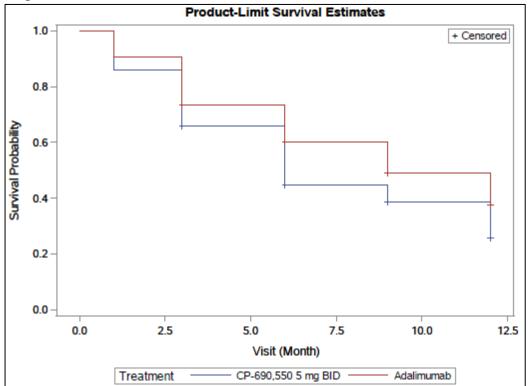


Abbildung 21: Kaplan-Meier Kurven für den Endpunkt ACR50 Response nach 12 Monaten (Population d)-Kombi)

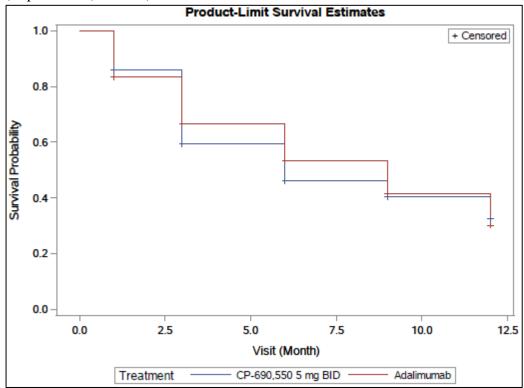


Tabelle 4-34: Ergebnisse für ACR/EULAR-Remission nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab ATX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	n csDMARD in	Vortherapie		
A3921064	7/79		4/75		1,7 [0,5-5,5], p=0,402
	(8,9)		(5,3)		1,7 [0,5-6,2], p=0,400
					3,5% [-4,5%-11,6%], p=0,392
		NA		NA	1,2 [0,5-2,7], p=0,702
		[NA-NA]		[NA-NA]	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	11/100		10/103		1,1 [0,5-2,6], p=0,763
	(11,0)		(9,7)		1,2 [0,5-2,8], p=0,763
					1,3% [-7,1%-9,7%], p=0,763
		NA		NA	1,5 [0,7-2,9], p=0,288
		[NA-NA]		[NA-NA]	- 4
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

# Population c)-Kombi:

Die Responderraten sind höher unter Tofacitinib + MTX (8,9% bzw. 5,3%). Es liegt kein statistisch signifikanter Unterschied vor.

# Population d)-Kombi:

Die Responderraten sind vergleichbar in beiden Studienarmen (11,0% bzw. 9,7%). Es liegt kein statistisch signifikanter Unterschied vor.

Abbildung 22: Kaplan-Meier Kurven für den Endpunkt ACR/EULAR-Remission nach 12 Monaten (Population c)-Kombi)

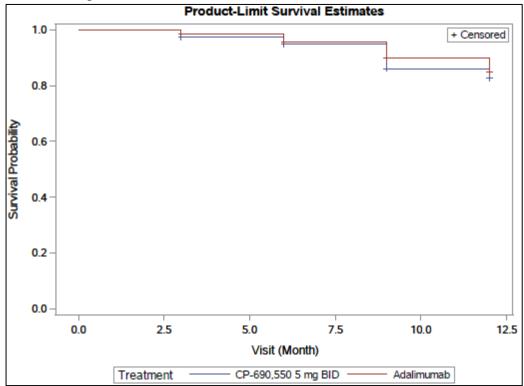


Abbildung 23: Kaplan-Meier Kurven für den Endpunkt ACR/EULAR-Remission nach 12 Monaten (Population d)-Kombi)

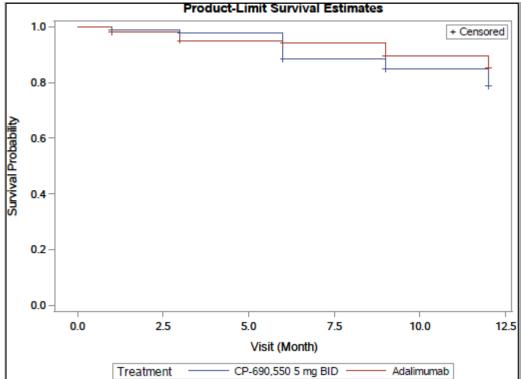


Tabelle 4-35: Sensitivitätsanalysen für ACR/EULAR-Remission nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib +MTX	Adalimumab +MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)-l	Kombi: nur ein csDMARD in V	ortherapie	
keine Imputation	6/60 (10,0)	4/61 (6,6)	1,5 [0,5-5,1], p=0,496 1,6 [0,4-5,9], p=0,495 3,4% [-6,4%-13,3%], p=0,492
LOCF	8/79 (10,1)	5/75 (6,7)	1,5 [0,5-4,4], p=0,445 1,6 [0,5-5,1], p=0,443 3,5% [-5,3%-12,2%], p=0,437
Population d)-	Kombi: mehr als ein csDMARI	) in Vortherapie	
keine Imputation	11/79 (13,9)	10/84 (11,9)	1,2 [0,5-2,6], p=0,701 1,2 [0,5-3,0], p=0,701 2,0% [-8,3%;12,3%], p=0,701
LOCF	13/100 (13,0)	12/103 (11,7)	1,1 [0,5-2,3], p=0,770 1,1 [0,5-2,6], p=0,770 1,4% [-7,7%;10,4%], p=0,770
Quelle: (38) Abkürzungen:	RR = Relatives Risiko, (	OR = Odds Ratio, 95%-KI	= 95%-Konfidenzintervall, MTX =

Abkürzungen: RR = Relatives Risiko, OR = Odds Ratio, 95%-KI = 95%-Konfidenzintervall, MTX = Methotrexat, LOCF = Last Observation Carried Forward, NA = nicht angebbar

### Population c)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

#### Population d)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

#### **4.3.1.3.5 SDAI** nach **12 Monaten** – **RCT**

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

Tabelle 4-36: Operationalisierung von SDAI nach 12 Monaten

Studie	Operationalisierung
A3921064	Der SDAI ist ein Summenscore für die Erfassung der Krankheitsaktivität, der sich aus den folgenden Einzelkomponenten zusammensetzt:
	1. Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 28 Gelenken)
	2. Anzahl geschwollener Gelenke (von insgesamt 28 Gelenken)
	3. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
	4. Gesamtbeurteilung der Krankheitsaktivität durch den Arzt (PhysGA)
	5. CRP-Wert
	Die Einzelkomponenten und somit auch der SDAI wurden zu jedem Studienzeitpunkt erhoben. Der SDAI hat einen Wertebereich von 0 (keinerlei Krankheitsaktivität) bis 86 (höchste Krankheitsaktivität).
	Es werden Responderanalysen zu Monat 12 mit zwei verschiedenen Schwellenwerten dargestellt. Ein Patient gilt demnach als Responder, falls sein SDAI ≤ 3,3 (dies ist das Kriterium für "Remission") bzw. < 11 nach 12 Monaten Studiendauer beträgt. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.
	Der Endpunkt wird für die FAS Population analysiert.

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

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

Tabelle 4-37: Bewertung des Verzerrungspotenzials für SDAI nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der kombinierte Endpunkt setzt sich aus fünf Einzelkomponenten zusammen, die bezüglich des Verzerrungspotenzials einheitlich eingestuft werden können. Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-38: Ergebnisse für SDAI ≤ 3,3 (Remission) nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX		Adalimumab +MTX		Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	n csDMARD in	Vortherapie		
A3921064	12/79		6/75		1,9 [0,8-4,8], p=0,176
	(15,2)		(8,0)		2,1 [0,7-5,8], p=0,172
					7,2% [-2,8%-17,2%], p=0,160
		NA		NA	1,6 [0,8-3,2], p=0,229
		[NA-NA]		[NA-NA]	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	14/100		17/103		0,9 [0,4-1,6], p=0,621
	(14,0)		(16,5)		0,8 [0,4-1,8], p=0,620
					-2,5% [-12,4%-7,4%], p=0,619
		NA		NA	1,0 [0,6-2,0], p=0,893
		[NA-NA]		[NA-NA]	-
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

### Population c)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (15,2% vs. 8,0%). Der Unterschied ist jedoch nicht statistisch signifikant.

#### Population d)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (16,5% vs. 14,0%). Der Unterschied ist jedoch nicht statistisch signifikant.

Abbildung 24: Kaplan-Meier Kurven für den Endpunkt SDAI  $\leq$  3,3 nach 12 Monaten (Population c)-Kombi)

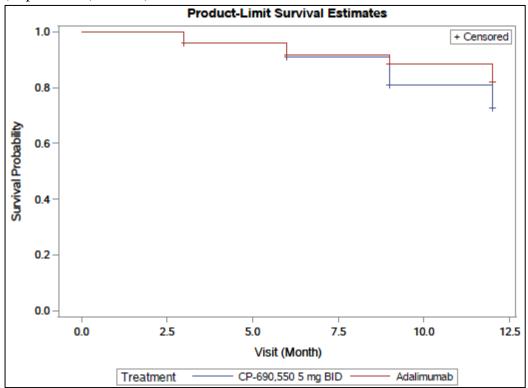


Abbildung 25: Kaplan-Meier Kurven für den Endpunkt SDAI  $\leq$  3,3 nach 12 Monaten (Population d)-Kombi)

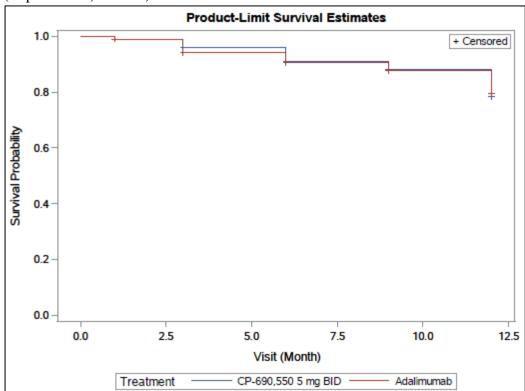


Tabelle 4-39: Ergebnisse für SDAI ≤ 11 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX		Adalimumab +MTX		Behandlungseffekt (Tofacitinib vs. Adalimumab)	
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	)-Kombi: nur ei	in csDMARD in	Vortherapie			
A3921064	39/79		32/75		1,2 [0,8-1,6], p=0,407	
	(49,4)		(42,7)		1,3 [0,7-2,5], p=0,405	
					6,7% [-9,0%-22,4%], p=0,403	
		9,0		9,0	1,3 [0,9-2,0], p=0,197	
		[6,0-9,0]		[6,0-12,0]		
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e		
A3921064	45/100		37/103		1,3 [0,9-1,8], p=0,190	
	(45,0)		(35,9)		1,5 [0,8-2,6], p=0,188	
					9,1% [-4,4%-22,5%], p=0,186	
		6,0		9,0	1,1 [0,8-1,6], p=0,616	
		[6,0-12,0]		[9,0-12,0]	- •	
Quelle: (40)						
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-	

# Population c)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (49,4% vs. 42,7%). Der Unterschied ist jedoch nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (45,0% vs. 35,9%). Der Unterschied ist jedoch nicht statistisch signifikant.

Abbildung 26: Kaplan-Meier Kurven für den Endpunkt SDAI ≤ 11 nach 12 Monaten (Population c)-Kombi)

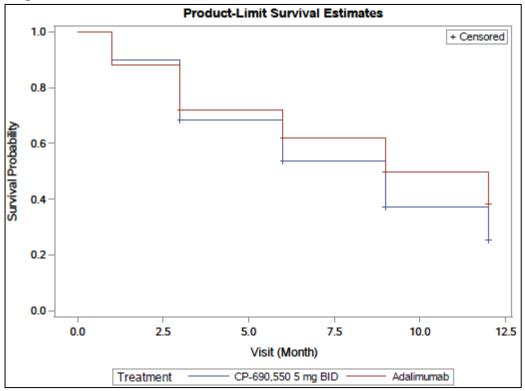
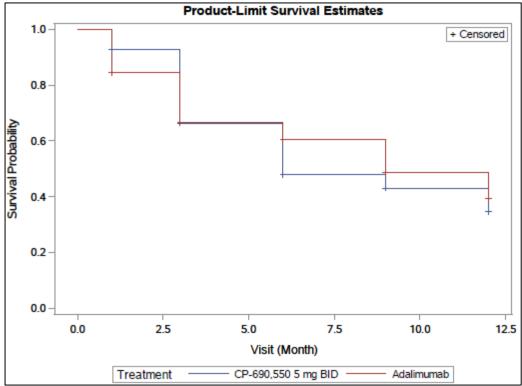


Abbildung 27: Kaplan-Meier Kurven für den Endpunkt  $SDAI \leq 11$  nach 12 Monaten (Population d)-Kombi)



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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

#### **4.3.1.3.6** CDAI nach 12 Monaten – RCT

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

Tabelle 4-40: Operationalisierung von CDAI nach 12 Monaten

Studie	Operationalisierung
A3921064	Der CDAI ist ein Summenscore für die Erfassung der Krankheitsaktivität, der im klinischen Alltag gegenüber dem SDAI bevorzugt angewendet wird, da die Notwenigkeit der Erhebung des CRP entfällt. Der CDAI setzt sich aus den folgenden Einzelkomponenten zusammen:
	1. Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 28 Gelenken)
	2. Anzahl geschwollener Gelenke (von insgesamt 28 Gelenken)
	3. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
	4. Gesamtbeurteilung der Krankheitsaktivität durch den Arzt (PhysGA)
	Die Einzelkomponenten und somit auch der CDAI wurden zu jedem Studienzeitpunkt erhoben. Der CDAI hat einen Wertebereich von 0 (keinerlei Krankheitsaktivität) bis 76 (höchste Krankheitsaktivität).
	Es werden Responderanalysen zu Monat 12 mit zwei verschiedenen Schwellenwerten dargestellt. Ein Patient gilt demnach als Responder, falls sein CDAI < 2,8 bzw. < 10 nach 12 Monaten Studiendauer beträgt. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.
	Der Endpunkt wird für die FAS Population analysiert.

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

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

Tabelle 4-41: Bewertung des Verzerrungspotenzials für CDAI nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der kombinierte Endpunkt setzt sich aus vier Einzelkomponenten zusammen, die bezüglich des Verzerrungspotenzials einheitlich eingestuft werden können. Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-42: Ergebnisse für CDAI ≤ 2,8 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX		Adalimumab +MTX		Behandlungseffekt (Tofacitinib vs. Adalimumab)	
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	)-Kombi: nur ei	in csDMARD in	Vortherapie			
A3921064	12/79		7/75		1,6 [0,7-3,9], p=0,276	
	(15,2)		(9,3)		1,7 [0,7-4,7], p=0,274	
					5,9% [-4,4%-16,2%], p=0,265	
		NA		NA	1,9 [0,9-4,1], p=0,099	
		[NA-NA]		[NA-NA]		
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e		
A3921064	14/100		14/103		1,0 [0,5-2,1], p=0,933	
	(14,0)		(13,6)		1,0 [0,5-2,3], p=0,933	
					0,4% [-9,1%-9,9%], p=0,933	
		NA		NA	1,0 [0,5-1,9], p=0,971	
		[NA-NA]		[NA-NA]		
Quelle: (40)						
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-	

### Population c)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (15,2% vs. 9,3%). Der Unterschied ist jedoch nicht statistisch signifikant.

#### Population d)-Kombi:

Die Responderraten sind vergleichbar in beiden Studienarmen (14,0% bzw. 13,6%). Es liegt kein statistisch signifikanter Unterschied vor.

Abbildung 28: Kaplan-Meier Kurven für den Endpunkt CDAI  $\leq$  2,8 nach 12 Monaten (Population c)-Kombi)

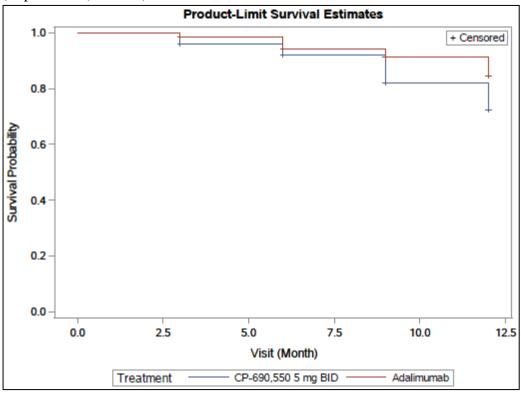


Abbildung 29: Kaplan-Meier Kurven für den Endpunkt CDAI ≤ 2,8 nach 12 Monaten (Population d)-Kombi)

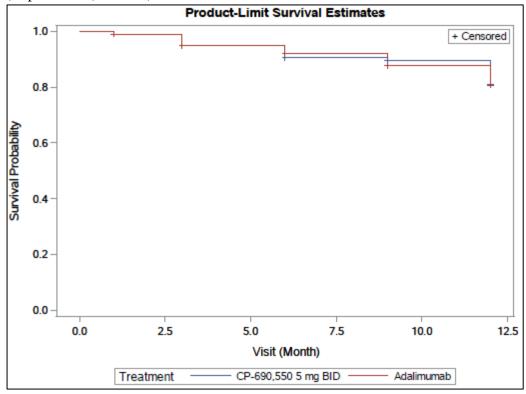


Tabelle 4-43: Ergebnisse für CDAI ≤ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX		Adalimumab +MTX		Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder /N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur e	in csDMARD in	Vortherapie		
A3921064	40/79		30/75		1,3 [0,9-1,8], p=0,190
	(50,6)		(40,0)		1,5 [0,8-2,9], p=0,186
					10,6% [-5,0%-26,3%], p=0,183
		9,0		12,0	1,3 [0,9-1,9], p=0,230
		[6,0-12,0]		[9,0-NA]	
Population d	)-Kombi: mehr	als ein csDMAl	RD in Vortherapi	e	
A3921064	42/100		37/103		1,2 [0,8-1,7], p=0,376
	(42,0)		(35,9)		1,3 [0,7-2,3], p=0,375
					6,1% [-7,3%-19,5%], p=0,374
		9,0		9,0	1,1 [0,7-1,5], p=0,731
		[6,0-12,0]		[9,0-12,0]	- *
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

# Population c)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (50,6% vs. 40,0%). Der Unterschied ist jedoch nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (42,0% vs. 35,9%). Der Unterschied ist jedoch nicht statistisch signifikant.

Abbildung 30: Kaplan-Meier Kurven für den Endpunkt CDAI  $\leq 10$  nach 12 Monaten (Population c)-Kombi)

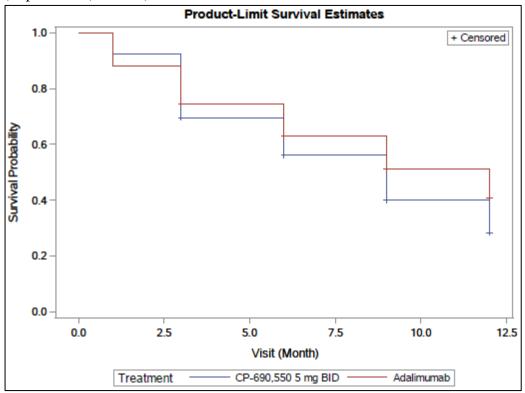
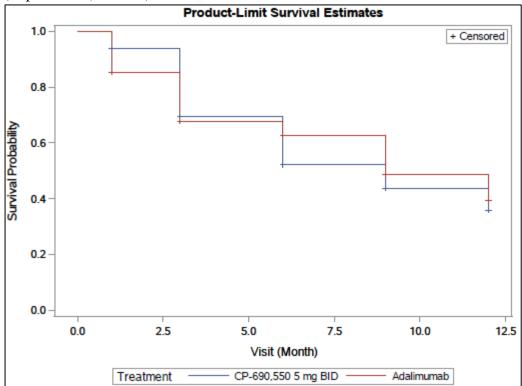


Abbildung 31: Kaplan-Meier Kurven für den Endpunkt CDAI ≤ 10 nach 12 Monaten (Population d)-Kombi)



Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.7 Schmerzbeurteilung nach 12 Monaten durch den Patienten – RCT

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

Tabelle 4-44: Operationalisierung von Schmerzbeurteilung nach 12 Monaten durch den Patienten

Studie	Operationalisierung
A3921064	Die Schmerzbeurteilung durch den Patienten wird anhand einer visuellen Analogskala erfasst, bei der der Patient die Intensität seiner Schmerzen auf einer Skala von 0 (keinerlei Schmerz) bis 100 (größte Schmerzen) beurteilt.
	Die Schmerzbeurteilung durch den Patienten wurde zu jedem Studienzeitpunkt erhoben.
	Es werden Responderanalysen zu Monat 12 dargestellt. Ein Patient gilt als Responder, falls sich sein Wert um mind. 10 Punkte auf der Skala gegenüber Baseline verbessert hat. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der Definition.
	Der Endpunkt wird für die FAS Population analysiert.

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

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

Tabelle 4-45: Bewertung des Verzerrungspotenzials für Schmerzbeurteilung nach 12 Monaten durch den Patienten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber (in diesem Fall der Patient) keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

Tabelle 4-46: Ergebnisse für Schmerzbeurteilung nach 12 Monaten durch den Patienten ≥ 10 aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab ATX	Behandlungseffekt (Tofacitinib vs. Adalimumab)				
	Anzahl Responder N	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert				
					HR [95%-KI], p-Wert				
Population c)-Kombi: nur ein csDMARD in Vortherapie									
A3921064	46/79		47/75		0,9 [0,7-1,2], p=0,573				
	(58,2)		(62,7)		0,8 [0,4-1,6], p=0,574				
					-4,4% [-19,9%-11,0%], p=0,573				
		1,0		1,0	1,0 [0,7-1,5], p=0,842				
		[1,0-3,0]		[1,0-3,0]					
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e					
A3921064	62/100		65/103		1,0 [0,8-1,2], p=0,871				
	(62,0)		(63,1)		1,0 [0,5-1,7], p=0,871				
					-1,1% [-14,4%-12,2%], p=0,871				
		1,0		1,0	0,9 [0,7-1,3], p=0,681				
		[NA-NA]		[NA-NA]					
Quelle: (40)									
_	Abkürzungen: RR = Relatives Risiko, OR = Odds Ratio, HR = Hazard Ratio, 95%-KI = 95%-Konfidenzintervall, MTX = Methotrexat, NA = nicht angebbar								

# Population c)-Kombi:

Die Responderraten sind höher unter Adalimumab + MTX (62,7% bzw. 58,2%). Es liegt kein statistisch signifikanter Unterschied vor.

## Population d)-Kombi:

Die Responderraten sind vergleichbar in beiden Studienarmen (62,0% bzw. 63,1%). Es liegt kein statistisch signifikanter Unterschied vor.

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

Abbildung 32: Kaplan-Meier Kurven für den Endpunkt Schmerzbeurteilung nach 12 Monaten durch den Patienten ≥ 10 (Population c)-Kombi)

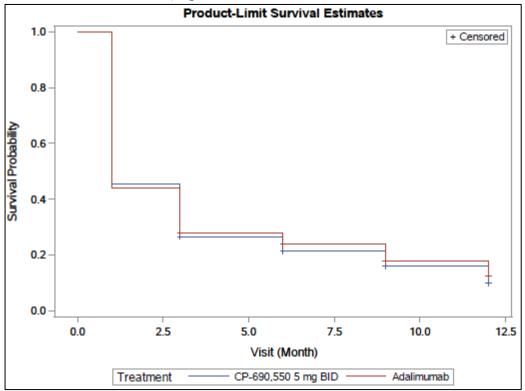
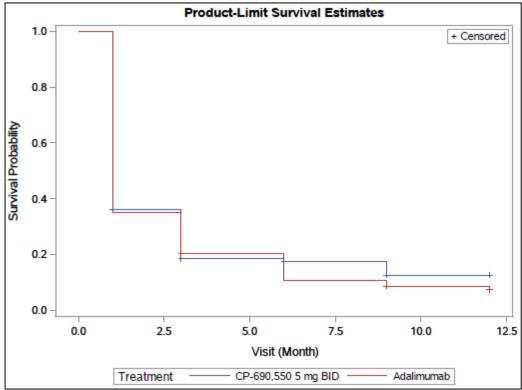


Abbildung 33: Kaplan-Meier Kurven für den Endpunkt Schmerzbeurteilung nach 12 Monaten durch den Patienten  $\geq$  10 (Population d)-Kombi)



Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

# 4.3.1.3.8 Gesamtbeurteilung der Krankheitsaktivität nach 12 Monaten durch den Patienten (PatGA) – RCT

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

Tabelle 4-47: Operationalisierung von PatGA nach 12 Monaten

Studie	Operationalisierung
A3921064	Der PatGA ( <i>Patient Global Assessment of Arthritis</i> ) ist ein Instrument für die Erfassung der Krankheitsaktivität, bei der der Patient auf einer visuellen Analogskala die Frage "Wie schätzen Sie heute Ihre Krankheitsaktivität ein?" auf einer Skala von 0 (keinerlei Krankheitsaktivität) bis 100 (höchste Krankheitsaktivität) beurteilt.
	Der PatGA wurde zu jedem Studienzeitpunkt erhoben.
	Es werden Responderanalysen zu Monat 12 dargestellt. Ein Patient gilt als Responder, falls sich sein Wert um mind. 10 Punkte auf der Skala gegenüber Baseline verbessert hat. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der Definition.
	Der Endpunkt wird für die FAS Population analysiert.

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 PatGA nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber (in diesem Fall der Patient) keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

Tabelle 4-49: Ergebnisse für PatGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab ATX	Behandlungseffekt (Tofacitinib vs. Adalimumab)		
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert		
					HR [95%-KI], p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie							
A3921064	46/79		48/75		0,9 [0,7-1,2], p=0,463		
	(58,2)		(64,0)		0,8 [0,4-1,5], p=0,463		
					-5,8% [-21,1%-9,6%], p=0,462		
		1,0		1,0	1,0 [0,7-1,4], p=0,928		
		[1,0-3,0]		[1,0-3,0]			
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e			
A3921064	64/100		61/103		1,1 [0,9-1,3], p=0,485		
	(64,0)		(59,2)		1,2 [0,7-2,2], p=0,485		
					4,8% [-8,6%-18,1%], p=0,484		
		1,0		1,0	1,0 [0,8-1,4], p=0,876		
		[NA-NA]		[1,0-3,0]	-		
Quelle: (40)							
_			o, OR = Odds NA = nicht ange		: Hazard Ratio, 95%-KI = 95%-		

# Population c)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (64,0% vs. 58,2%). Der Unterschied ist nicht statistisch signifikant.

## Population d)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (64,0% vs. 59,2%). Der Unterschied ist jedoch nicht statistisch signifikant.

Abbildung 34: Kaplan-Meier Kurven für den Endpunkt PatGA ≥ 10 nach 12 Monaten (Population c)-Kombi)

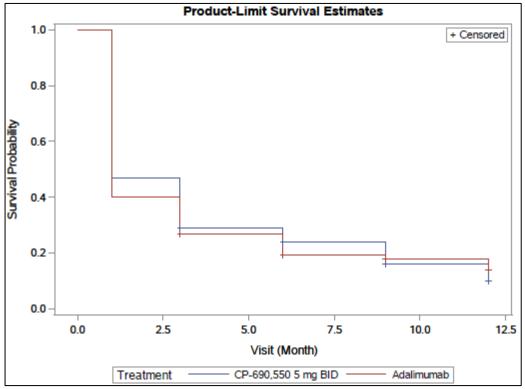
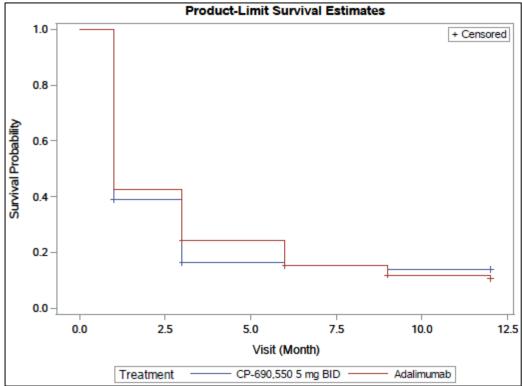


Abbildung 35: Kaplan-Meier Kurven für den Endpunkt PatGA ≥ 10 nach 12 Monaten (Population d)-Kombi)



Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

# 4.3.1.3.9 Gesamtbeurteilung der Krankheitsaktivität nach 12 Monaten durch den Arzt (PhysGA) – RCT

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

Tabelle 4-50: Operationalisierung von PhysGA

Studie	Operationalisierung
A3921064	Der PhysGA ( <i>Physician Global Assessment of Arthritis</i> ) ist ein Instrument für die Erfassung der Krankheitsaktivität, bei der der Arzt die Krankheitsaktivität des Patienten auf einer Skala von 0 (keinerlei Krankheitsaktivität) bis 100 (höchste Krankheitsaktivität) beurteilt.
	Der PhysGA wurde zu jedem Studienzeitpunkt erhoben.
	Es werden Responderanalysen zu Monat 12 dargestellt. Ein Patient gilt als Responder, falls sich sein Wert um mind. 10 Punkte auf der Skala gegenüber Baseline verbessert hat. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der Definition.
	Der Endpunkt wird für die FAS Population analysiert.

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 4-51: Bewertung des Verzerrungspotenzials für PhysGA nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

Tabelle 4-52: Ergebnisse für Phys $GA \ge 10$  nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab ATX	Behandlungseffekt (Tofacitinib vs. Adalimumab)		
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert		
					HR [95%-KI], p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie							
A3921064	58/78 (74,4)		58/75 (77,3)		1,0 [0,8-1,2], p=0,667 0,9 [0,4-1,8], p=0,668 -3,0% [-16,5%-10,6%], p=0,667		
		1,0 [NA-NA]		1,0 [NA-NA]	1,0 [0,7-1,3], p=0,833		
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e			
A3921064	70/99 (70,7)		76/103 (73,8)		1,0 [0,8-1,1], p=0,626 0,9 [0,5-1,6], p=0,625 -3,1% [-15,4%-9,3%], p=0,625		
		1,0 [NA-NA]		1,0 [NA-NA]	0,9 [0,7-1,2], p=0,432		
_	en: RR = Re		o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-		

# Population c)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (77,3% vs. 74,4%). Der Unterschied ist nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (73,8% vs. 70,7%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 36: Kaplan-Meier Kurven für den Endpunkt Phys $GA \ge 10$  nach 12 Monaten (Population c)-Kombi)

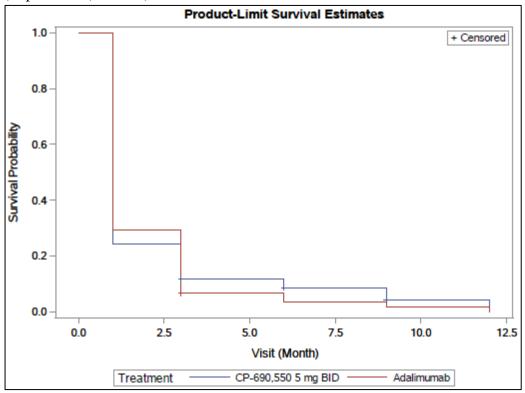
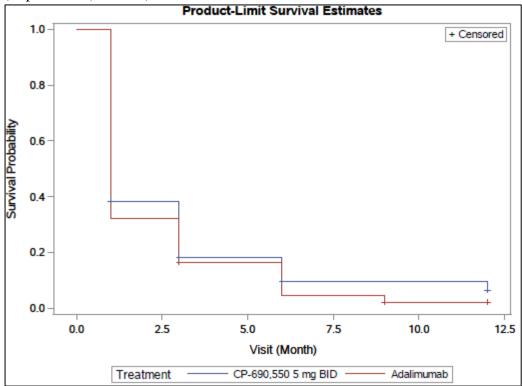


Abbildung 37: Kaplan-Meier Kurven für den Endpunkt Phys $GA \ge 10$  nach 12 Monaten (Population d)-Kombi)



Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.10 Anzahl empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten – RCT

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

Tabelle 4-53: Operationalisierung von Anzahl empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten

Studie	Operationalisierung
A3921064	Die Anzahl der empfindlichen/druckschmerzhaften Gelenke wird mittels zweier verschiedener Ansätze bestimmt. Zum einen von insgesamt 28 Gelenken und zum anderen von insgesamt 68 Gelenken.
	Anzahl empfindlicher/druckschmerzhafter Gelenke (28):
	Hierzu werden folgende Gelenke auf beiden Körperseiten untersucht: Knie-, Schulter-, Ellbogen-, Handgelenk, Fingergrundgelenke und proximale Interphalangealgelenke.
	Anzahl empfindlicher/druckschmerzhafter Gelenke (68):
	Hierzu werden folgende Gelenke auf beiden Körperseiten untersucht:
	Oberkörper: Temporomandibular-, Sternoklavikular-, Akromioklavikulargelenk
	Obere Extremitäten: Schulter-, Ellbogen-, Handgelenk, Fingergrundgelenke, Daumeninterphalangealgelenk, proximale und distale Interphalangealgelenke
	<u>Untere Extremitäten</u> : Hüft-, Knie-, Sprung-, Fußwurzelgelenk, Zehengrundgelenke, Interphalangealgelenk des großen Zehs, proximale und distale Interphalangealgelenke.
	Für beide Varianten (28 bzw. 68 Gelenke) gilt, dass ein Patient als Responder gezählt wird, falls sich in der klinischen Untersuchung maximal ein Gelenk als empfindlich/druckschmerzhaft zeigt.
	Es werden jeweils Responderanalysen zu Monat 12 berichtet. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.
	Der Endpunkt wird für die FAS Population analysiert.

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 4-54: Bewertung des Verzerrungspotenzials für Anzahl empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

Tabelle 4-55: Ergebnisse für Anzahl empfindlicher/druckschmerzhafter Gelenke (28) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX			mumab ATX	Behandlungseffekt (Tofacitinib vs. Adalimumab)		
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert		
					HR [95%-KI], p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie							
A3921064	24/79		22/75		1,0 [0,6-1,7], p=0,887		
	(30,4)		(29,3)		1,1 [0,5-2,1], p=0,887		
					1,1% [-13,4%-15,5%], p=0,887		
		12,0		NA	1,3 [0,8-2,1], p=0,238		
		[9,0-NA]		[12,0-NA]			
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e			
A3921064	27/100		33/103		0,8 [0,6-1,3], p=0,433		
	(27,0)		(32,0)		0,8 [0,4-1,4], p=0,432		
					-5,0% [-17,6%-7,5%], p=0,431		
		12,0		NA	1,1 [0,7-1,7], p=0,625		
		[9,0-NA]		[9,0-NA]	-		
Quelle: (40)							
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-		

# Population c)-Kombi:

Die Responderraten sind vergleichbar in beiden Studienarmen (30,4% bzw. 29,3%). Es liegt kein statistisch signifikanter Unterschied vor.

## Population d)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (32,0% vs. 27,0%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 38: Kaplan-Meier Kurven für den Endpunkt empfindlicher/druckschmerzhafter Gelenke (28) ≤ 1 nach 12 Monaten (Population c)-Kombi)

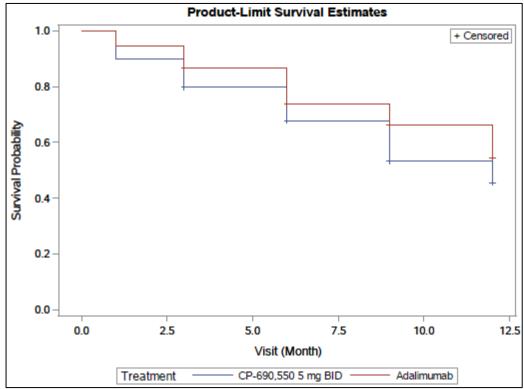


Abbildung 39: Kaplan-Meier Kurven für den Endpunkt empfindlicher/druckschmerzhafter Gelenke (28) ≤ 1 nach 12 Monaten (Population d)-Kombi)

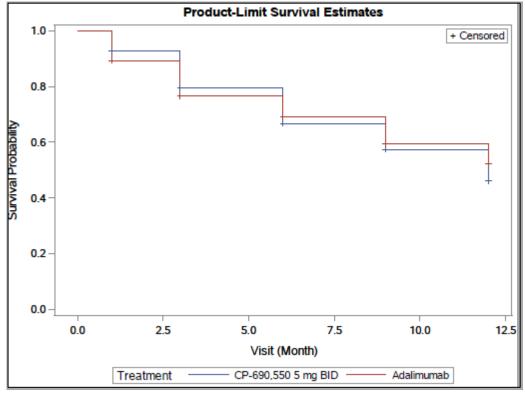


Tabelle 4-56: Ergebnisse für Anzahl empfindlicher/druckschmerzhafter Gelenke (68) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie Tofacitinib +MTX				imumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur e	in csDMARD in	Vortherapie		
A3921064	15/79 (19,0)		16/75 (21,3)		0,9 [0,5-1,7], p=0,717 0,9 [0,4-1,9], p=0,717
	(19,0)		(21,3)		-2,4% [-15,0%-10,3%], p=0,717
		NA		NA	1,6 [0,9-2,8], p=0,147
		[12,0-NA]		[NA-NA]	
Population d	)-Kombi: mehr	als ein csDMAl	RD in Vortherapi	e	
A3921064	18/100		26/103		0,7 [0,4-1,2], p=0,215
	(19,0)		(25,2)		0,7 [0,3-1,3], p=0,212
					-7,2% [-18,5%-4,0%], p=0,208
		NA		NA	0,8 [0,5-1,4], p=0,530
		[NA-NA]		[NA-NA]	
Quelle: (40)					
•			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

# Population c)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (21,3% vs. 19,0%). Der Unterschied ist nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (25,2% vs. 19,0%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 40: Kaplan-Meier Kurven für den Endpunkt empfindlicher/druckschmerzhafter Gelenke (68) ≤ 1 nach 12 Monaten (Population c)-Kombi)

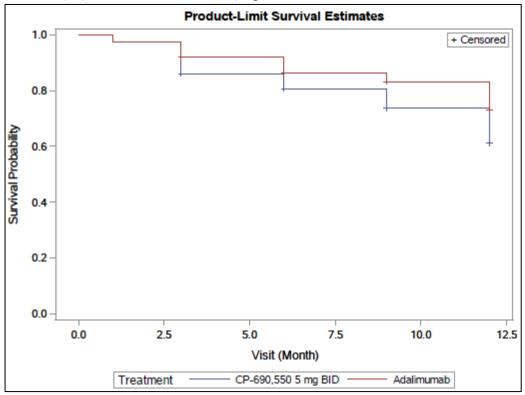
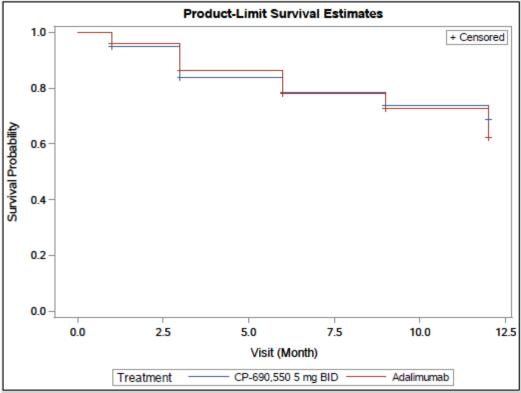


Abbildung 41: Kaplan-Meier Kurven für den Endpunkt empfindlicher/druckschmerzhafter Gelenke (68)  $\leq$  1 nach 12 Monaten (Population d)-Kombi)



Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.11 Anzahl geschwollener Gelenke nach 12 Monaten – RCT

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

Tabelle 4-57: Operationalisierung von Anzahl geschwollener Gelenke nach 12 Monaten

Studie	Operationalisierung					
A3921064	Die Anzahl der geschwollenen Gelenke wird mittels zweier verschiedener Ansätze bestimmt. Zum einen von insgesamt 28 Gelenken und zum anderen von insgesamt 66 Gelenken.					
	Anzahl geschwollener Gelenke (28):					
	Hierzu werden folgende Gelenke auf beiden Körperseiten untersucht: Knie-, Schulter-, Ellbogen-, Handgelenk, Fingergrundgelenke und proximale Interphalangealgelenke.					
	Anzahl geschwollener Gelenke (66):					
	Hierzu werden folgende Gelenke auf beiden Körperseiten untersucht:					
	Oberkörper: Temporomandibular-, Sternoklavikular-, Akromioklavikulargelenk					
	Obere Extremitäten: Schulter-, Ellbogen-, Handgelenk, Fingergrundgelenke, Daumeninterphalangealgelenk, proximale und distale Interphalangealgelenke					
	<u>Untere Extremitäten</u> : Knie-, Sprung-, Fußwurzelgelenk, Zehengrundgelenke, Interphalangealgelenk des großen Zehs, proximale und distale Interphalangealgelenke.					
	Für beide Varianten (28 bzw. 66 Gelenke) gilt, dass ein Patient als Responder gezählt wird, falls sich in der klinischen Untersuchung maximal ein Gelenk als empfindlich/druckschmerzhaft zeigt.					
	Es werden jeweils Responderanalysen zu Monat 12 berichtet. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.					
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.					
	Der Endpunkt wird für die FAS Population analysiert.					

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 4-58: Bewertung des Verzerrungspotenzials für Anzahl geschwollener Gelenke nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

Tabelle 4-59: Ergebnisse für Anzahl geschwollener Gelenke (28)  $\leq$  1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX		Adalimumab +MTX		Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	in csDMARD in	Vortherapie		
A3921064	36/79		34/75		1,0 [0,7-1,4], p=0,977
	(45,6)		(45,3)		1,0 [0,5-1,9], p=0,977
					0,2% [-15,5%-16,0%], p=0,977
		9,0		9,0	1,2 [0,8-1,7], p=0,481
		[6,0-12,0]		[6,0-12,0]	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	42/100		37/103		1,2 [0,8-1,7], p=0,376
	(42,0)		(35,9)		1,3 [0,7-2,3], p=0,375
					6,1% [-7,3%-19,5%], p=0,374
		9,0		12,0	1,4 [0,9-2,0], p=0,114
		[6,0-12,0]		[9,0-NA]	-
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

# Population c)-Kombi:

Die Responderraten sind vergleichbar in beiden Studienarmen (45,6% bzw. 45,3%). Es liegt kein statistisch signifikanter Unterschied vor.

## Population d)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (42,0% vs. 35,9%). Der Unterschied ist nicht statistisch signifikant.

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

Abbildung 42: Kaplan-Meier Kurven für den Endpunkt geschwollener Gelenke (28) ≤ 1 nach 12 Monaten (Population c)-Kombi)

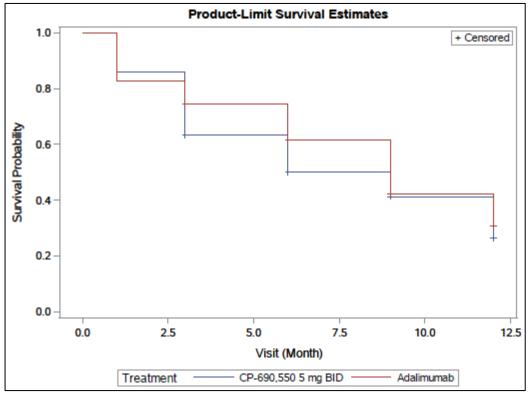


Abbildung 43: Kaplan-Meier Kurven für den Endpunkt geschwollener Gelenke (28)  $\leq$  1 nach 12 Monaten (Population d)-Kombi)

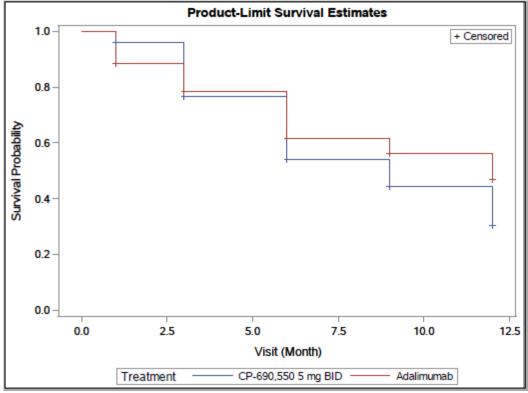


Tabelle 4-60: Ergebnisse für Anzahl geschwollener Gelenke (66) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib +MTX		Adalimumab +MTX		Behandlungseffekt (Tofacitinib vs. Adalimumab)	
	Anzahl Responder N	Median (Monate) [95%-KI]	Anzahl Responder/ N	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	)-Kombi: nur ei	in csDMARD in	Vortherapie			
A3921064	33/79		28/75		1,1 [0,8-1,7], p=0,574	
	(41,8)		(37,3)		1,2 [0,6-2,3], p=0,574	
					4,4% [-11,0%-19,9%], p=0,573	
		9,0		12,0	1,2 [0,8-1,8], p=0,343	
		[6,0-12,0]		[9,0-12,0]		
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e		
A3921064	35/100		33/103		1,1 [0,7-1,6], p=0,655	
	(35,0)		(32,0)		1,1 [0,6-2,1], p=0,655	
					3,0% [-10,0%-16,0%], p=0,655	
		12,0		NA	1,3 [0,9-2,0], p=0,157	
		[9,0-NA]		[12,0-NA]		
Quelle: (40)						
Ū			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%	

# Population c)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (41,8% vs. 37,3%). Der Unterschied ist nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (35,0% vs. 32,0%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 44: Kaplan-Meier Kurven für den Endpunkt geschwollener Gelenke (66) ≤ 1 nach 12 Monaten (Population c)-Kombi)

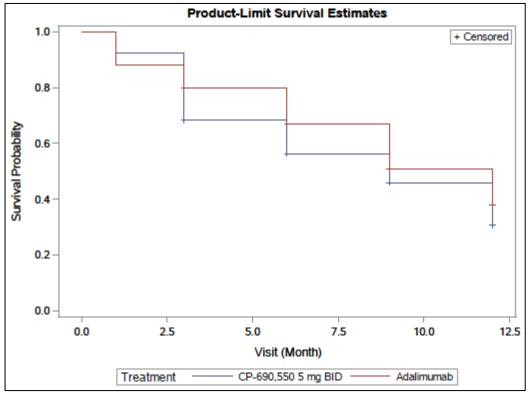
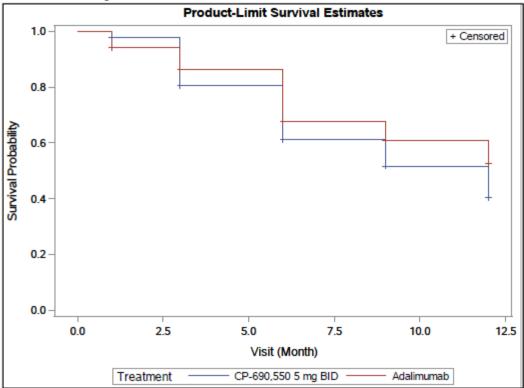


Abbildung 45: Kaplan-Meier Kurven für den Endpunkt geschwollener Gelenke (66) ≤ 1 nach 12 Monaten (Population d)-Kombi)



Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

#### 4.3.1.3.12 SF-36 nach 12 Monaten – RCT

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

Tabelle 4-61: Operationalisierung von SF-36 nach 12 Monaten

Studie	Operationalisierung						
A3921064	Der SF-36 ist ein Instrument zur Erfassung der gesundheitsbezogenen Lebensqualität. Er besteht aus acht Domänen und zwei Komponentenscores:						
	Körperlicher Komponentenscore						
	Psychischer Komponentenscore						
	• Vitalität						
	Körperliche Funktion						
	Körperliche Schmerzen						
	Allgemeine Gesundheit						
	Körperliche Rollenfunktion						
	Emotionale Rollenfunktion						
	Psychisches Wohlbefinden						
	• Soziale Funktion						
	Die Fragebögen wurden von den Patienten jeweils zu Beginn jeden Visits ausgefüllt. Die einzelnen Domänen und Komponentenscores haben jeweils einen Wertebereich von 0 bis 100; je höher der Wert, desto besser die Lebensqualität.  Es werden jeweils Responderanalysen zu Monat 12 dargestellt mit den Schwellenwerten 2,5 (Komponentenscores) bzw. 5 (pro Domäne). Fehlende Werte werden gemäß der sog. Nonresponder Imputation (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.  Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen für die beiden Komponentenscores angegeben und Kaplan-Meier Kurven dargestellt.						
	Des Weiteren wird die mittlere Änderung zu Monat 12 gegenüber Baseline inklusive des standardisierten Effektmaßes Hedges' g berichtet. Hierfür wurden Modelle für wiederholte Messwerte mit gemischten Effekten gerechnet (feste Effekte: Behandlung, Studienzeitpunkt, Behandlung×Studienzeitpunkt, Region, Baseline-Wert; zufälliger Effekt: Patient). Fehlende Werte werden hierbei nicht ersetzt.						
	Als Sensitivitätsanalysen werden für die beiden Komponentenscores zusätzlich die Effektschätzer gezeigt, die sich ohne Imputation ergeben sowie die, die mittels LOCF-Imputation resultieren.						
	Der Endpunkt wird für die FAS Population analysiert.						

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 4-62: Bewertung des Verzerrungspotenzials für SF-36 nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber (in diesem Fall der Patient) keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

Tabelle 4-63: Ergebnisse für SF-36 nach 12 Monaten – Responderanalysen der Komponentenscores aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib +MTX			limumab -MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responde r/N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	)-Kombi: nur ei	n csDMARD in `	Vortherapie			
Körperl. Kompo-	47/79 (59,5)		47/75 (62,7)		1,0 [0,7-1,2], p=0,686 0,9 [0,5-1,7], p=0,687	
nenten- score $\geq 2,5$					-3,2% [-18,6%-12,2%], p=0,686	
		1,0 [1,0-3,0]		1,0 [1,0-3,0]	1,0 [0,7-1,5], p=0,854	
Psych.	32/79		36/75		0,8 [0,6-1,2], p=0,350	
Kompo-	(40,5)		(48,0)		0,7 [0,4-1,4], p=0,350	
nenten- score $\geq 2,5$					-7,5% [-23,2%-8,2%], p=0,348	
		3,0		1,0	0,8 [0,6-1,2], p=0,293	
		[1,0-3,0]		[1,0-3,0]		
Population d	)-Kombi: mehr	als ein csDMAR	D in Vortherapi	e		
Körperl.	59/100		62/103		1,0 [0,8-1,2], p=0,862	
Kompo-	(59,0)		(60,2)		1,0 [0,5-1,7], p=0,862	
nenten- score $\geq 2,5$					-1,2% [-14,7%-12,3%], p=0,862	
		1,0		1,0	1,2 [0,9-1,5], p=0,351	
		[NA-NA]		[1,0-NA]		
Psych.	45/100		43/103		1,1 [0,8-1,5], p=0,640	
Kompo-	(45,0)		(41,8)		1,1 [0,7-2,0], p=0,640	
nenten- score $\geq 2,5$					3,3% [-10,4%-16,8%], p=0,640	
		1,0		1,0	1,0 [0,7-1,4], p=0,960	
		[1,0-3,0]		[1,0-3,0]		
Quelle: (40)						
Abkürzunge		latives Risiko, Methotrexat, N			Hazard Ratio, 95%-KI =	

# Population c)-Kombi:

Die Responderraten für den Körperlichen Komponentenscore sind vergleichbar in beiden Studienarmen (59,5% bzw. 62,7%). Für den Psychischen Komponentenscore ist die Responderrate höher unter Adalimumab + MTX (48,0% vs. 40,5%). Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderrate für den Körperlichen Komponentenscore ist vergleichbar in beiden Studienarmen (59,0% bzw. 60,2%). Für den Psychischen Komponentenscore ist die Responderrate höher unter Tofacitinib + MTX (45,0% vs. 41,8%). Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

Abbildung 46: Kaplan-Meier Kurven für den Endpunkt SF-36 Körperlicher Komponentenscore ≥ 2,5 nach 12 Monaten (Population c)-Kombi)

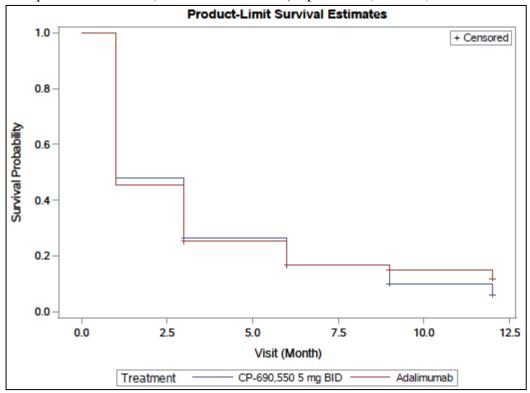


Abbildung 47: Kaplan-Meier Kurven für den Endpunkt SF-36 Körperlicher Komponentenscore ≥ 2,5 nach 12 Monaten (Population d)-Kombi)

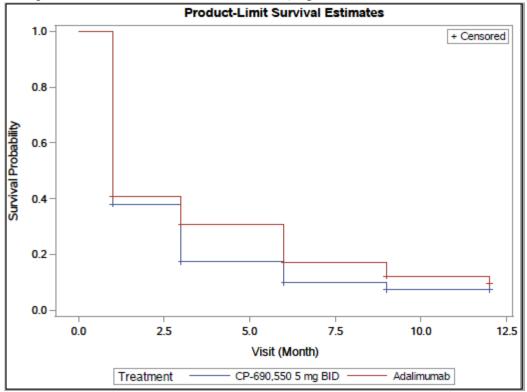


Abbildung 48: Kaplan-Meier Kurven für den Endpunkt SF-36 Psychischer Komponentenscore ≥ 2,5 nach 12 Monaten (Population c)-Kombi)

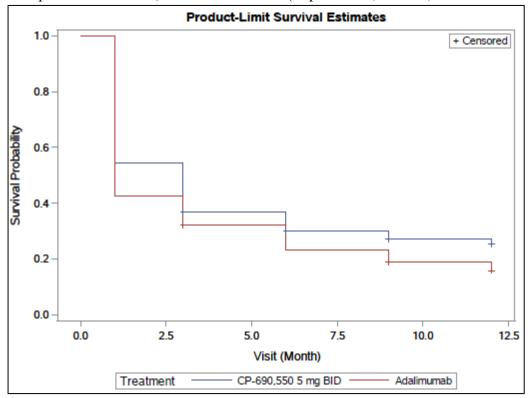


Abbildung 49: Kaplan-Meier Kurven für den Endpunkt SF-36 Psychischer Komponentenscore ≥ 2,5 nach 12 Monaten (Population d)-Kombi)

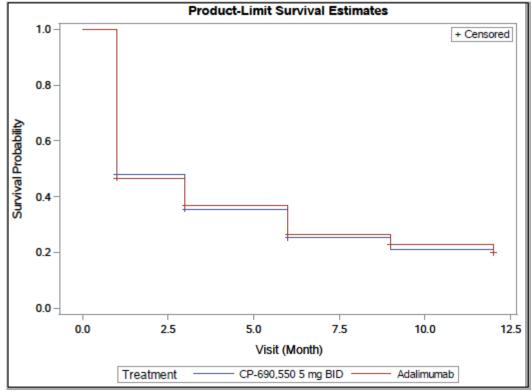


Tabelle 4-64: Sensitivitätsanalysen für SF-36 nach 12 Monaten - Komponentenscores aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064		Tofacitinib +MTX	Adalimumab +MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
		Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c	-Kombi: nur ein csDMARD ir	Vortherapie		
keine	Körperlicher Komponentenscore ≥ 2,5	45/59 (76,3)	47/61 (77,1)	1,0 [0,8-1,2], p=0,920 1,0 [0,4-2,2], p=0,920 -0,8% [-15,9%-14,4%], p=0,920
Imputation	Psychischer Komponentenscore ≥ 2,5	32/59 (54,2)	36/61 (59,0)	0,9 [0,7-1,3], p=0,598 0,8 [0,4-1,7], p=0,598 -4,8% [-22,5%-12,9%], p=0,597
LOCF	Körperlicher Komponentenscore $\geq 2,5$	65/79 (70,9)	52/75 (69,3)	1,0 [0,8-1,3], p=0,834 1,1 [0,5-2,2], p=0,833 1,6% [-12,9%-16,0%], p=0,833
	Psychischer Komponentenscore ≥ 2,5	40/79 (50,6)	43/75 (57,3)	0,9 [0,7-1,2], p=0,405 0,8 [0,4-1,4], p=0,405 -6,7% [-22,4%-9,0%], p=0,403
Population d	)-Kombi: mehr als ein csDMA	RD in Vortherapi	e	
keine	Körperlicher Komponentenscore ≥ 2,5	60/79 (76,0)	60/82 (73,2)	1,0 [0,9-1,2], p=0,686 1,2 [0,6-2,4], p=0,686 2,8% [-10,7%-16,2%], p=0,686
Imputation	Psychischer Komponentenscore ≥ 2,5	45/79 (57,0)	43/82 (52,4)	1,1 [0,8-1,4], p=0,565 1,2 [0,6-2,2], p=0,565 4,5% [-10,8%-19,9%], p=0,564
LOCF	Körperlicher Komponentenscore ≥ 2,5	71/100 (71,0)	71/103 (68,9)	1,0 [0,9-1,2], p=0,748 1,1 [0,6-2,0], p=0,748 2,1% [-10,5%-14,7%], p=0,748
LOCF	Psychischer Komponentenscore ≥ 2,5	51/100 (51,0)	47/103 (45,6)	1,1 [0,8-1,5], p=0,445 1,2 [0,7-2,2], p=0,444 5,4% [-8,4%-19,1%], p=0,443
Quelle: (38)				

Quelle: (38)

Abkürzungen: RR = Relatives Risiko, OR = Odds Ratio, 95%-KI = 95%-Konfidenzintervall, MTX =

Methotrexat, LOCF = Last Observation Carried Forward, NA = nicht angebbar

## Population c)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

## Population d)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

Tabelle 4-65: Ergebnisse für SF-36 nach 12 Monaten – Responderanalysen der Einzeldimensionen aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib +MTX	Adalimumab +MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)-Kombi: nur ein csDMAR	D in Vortherapie		
Vitalität ≥ 5	36/79 (45,6)	38/75 (50,7)	0,9 [0,7-1,3], p=0,527 0,8 [0,4-1,5], p=0,527 -5,1% [-20,9%-10,7%], p=0,526
Körperliche Funktion ≥ 5	40/79 (50,6)	38/75 (50,7)	1,0 [0,7-1,4], p=0,997 1,0 [0,5-1,9], p=0,997 -0,03% [-15,8%-15,8%], p=0,997
Körperliche Schmerzen ≥ 5	37/79 (46,8)	39/75 (52,0)	0,9 [0,7-1,2], p=0,522 0,8 [0,4-1,5], p=0,522 -5,2% [-20,9%-10,6%], p=0,521
Allgemeine Gesundheit ≥ 5	30/79 (38,0)	26/75 (34,7)	1,1 [0,7-1,7], p=0,670 1,2 [0,6-2,2], p=0,670 3,3% [-11,9%-18,5%], p=0,669
Körperliche Rollenfunktion ≥ 5	32/79 (40,5)	28/75 (37,3)	1,1 [0,7-1,6], p=0,687 1,1 [0,6-2,2], p=0,687 3,2% [-12,2%-18,6%], p=0,686
Emotionale Rollenfunktion $\geq 5$	28/79 (35,4)	26/75 (34,7)	1,0 [0,7-1,6], p=0,920 1,0 [0,5-2,0], p=0,920 0,8% [-14,3%-15,9%], p=0,920
Psychisches Wohlbefinden ≥ 5	29/79 (36,7)	30/75 (40,0)	0,9 [0,6-1,4], p=0,675 0,9 [0,5-1,7], p=0,675 -3,3% [-18,7%-12,1%], p=0,675
Soziale Funktion ≥ 5	38/79 (48,1)	43/75 (57,3)	0,8 [0,6-1,1], p=0,253 0,7 [0,4-1,3], p=0,252 -9,2% [-24,9%-6,5%], p=0,249
Population d)-Kombi: mehr als ein csD	MARD in Vortherapie	•	
Vitalität ≥ 5	46/100 (46,0)	45/103 (43,7)	1,1 [0,8-1,4], p=0,741 1,1 [0,6-1,9], p=0,741 2,3% [-11,4%-16,0%], p=0,741
Körperliche Funktion $\geq 5$	45/100 (45,0)	54/103 (52,4)	0,9 [0,7-1,1], p=0,292 0,7 [0,4-1,3], p=0,290 -7,4% [-21,1%-6,3%], p=0,289
Körperliche Schmerzen ≥ 5	49/100 (49,0)	56/103 (54,4)	0,9 [0,7-1,2], p=0,445 0,8 [0,5-1,4], p=0,444 -5,4% [-19,1%-8,4%], p=0,443

Allgemeine Gesundheit ≥ 5	40/100	44/103	0,9 [0,7-1,3], p=0,695
	(40,0)	(42,7)	0,9 [0,5-1,6], p=0,694
			-2,7% [-16,3%-10,8%], p=0,694
Körperliche Rollenfunktion $\geq 5$	42/100	44/103	1,0 [0,7-1,4], p=0,918
	(42,0)	(42,7)	1,0 [0,6-1,7], p=0,918
			-0,7% [-14,3%-12,9%], p=0,918
Emotionale Rollenfunktion $\geq 5$	38/100	37/103	1,1 [0,7-1,5], p=0,759
	(38,0)	(35,9)	1,1 [0,6-1,9], p=0,759
			2,1% [-11,2%-15,4%], p=0,759
Psychisches Wohlbefinden ≥ 5	41/100	40/103	1,1 [0,8-1,5], p=0,753
	(41,0)	(38,8)	1,1 [0,6-1,9], p=0,753
			2,2% [-11,3%-15,6%], p=0,753
Soziale Funktion $\geq 5$	47/100	50/103	1,0 [0,7-1,3], p=0,826
	(47,0)	(48,5)	1,0 [0,5-1,6], p=0,826
			-1,5% [-15,3%-12,2%], p=0,826
Quelle: (40)			

Abkürzungen: RR = Relatives Risiko, OR = Odds Ratio, HR = Hazard Ratio, 95%-KI = 95%-

Konfidenzintervall, MTX = Methotrexat, NA = nicht angebbar

## Population c)-Kombi:

Die Responderraten sind für die Domänen "Körperliche Funktion" und "Emotionale Rollenfunktion" in beiden Studienarmen vergleichbar. Für "Allgemeine Gesundheit" und "Körperliche Rollenfunktion" zeigt sich ein numerischer Vorteil im Tofacitinib-Arm, für "Vitalität", "Körperliche Schmerzen", "Psychisches Wohlbefinden" und "Soziale Funktion" im Adalimumab-Arm. Die beobachteten Unterschiede sind jedoch jeweils nicht statistisch signifikant.

#### Population d)-Kombi:

Die Responderraten sind für die meisten Domänen des SF-36 in beiden Studienarmen vergleichbar. Für "Körperliche Funktion" und "Körperliche Schmerzen" zeigt sich jeweils ein numerischer Vorteil im Adalimumab-Arm. Die beobachteten Unterschiede sind jedoch jeweils nicht statistisch signifikant.

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

Tabelle 4-66: Ergebnisse für SF-36 nach 12 Monaten – mittlere Änderungen aus RCT mit dem zu bewertenden Arzneimittel

Studie			Desk		Adjustierte Analyse <sup>a</sup> :		
A3921064	Baseline			Monat 12		Mittlere	Mittlere
					Mittlere Änderung zwischen Baseline und Monat 12	Änderung über 12 Monate Differenz	Änderung über 12 Monate Hedges'g
	N	$\begin{array}{c cccc} N & & Mittelwert \\ & \pm SD & & N & & \underline{+} SD \\ \end{array}$	Mittelwert <u>+</u> SD	Mittelwert ± SD	Tofacitinb – Adalimumab [95%-KI] p-Wert	Tofacitinb – Adalimumab [95%-KI] p-Wert	
Population c)-Kombi: nur ein csDM	ARD in V	ortherapie					
Körperl. Komponentenscore Tofacitinib + MTX Adalimumab + MTX	81 76	33,2±7,68 31,8±6,41	59 61	41,7±8,95 40,7±9,12	8,2±8,42 9,0±7,92	0,91 [-1,58;3,41] p=0,472	0,12 [-0,24;0,48] p=0,494
Psych. Komponentenscore Tofacitinib + MTX Adalimumab + MTX	81 76	39,6±12,72 40,0±11,61	59 61	44,9±9,17 43,7±11,01	4,3±9,02 3,6±11,19	0,81 [-2,22;3,84] p=0,597	0,09 [-0,27;0,45] p=0,617
Vitalität Tofacitinib + MTX Adalimumab + MTX	81 76	39,3±8,86 39,2±9,99	60 61	47,1±8,80 46,1±9,45	6,8±8,43 7,2±10,30	0,53 [-2,12;3,18] p=0,695	0,07 [-0,29;0,43] p=0,709
Körperliche Funktion Tofacitinib + MTX Adalimumab + MTX	81 76	31,1±9,16 30,9±9,42	59 61	39,7±9,65 39,0±10,54	7,9±9,99 8,0±10,34	1,06 [-1,90;4,01] p=0,483	0,12 [-0,24;0,48] p=0,504
Körperliche Schmerzen Tofacitinib + MTX Adalimumab + MTX	81 76	34,5±7,77 31,9±7,12	60 61	44,0±9,72 42,4±10,55	9,2±10,50 11,0±10,21	0,91 [-2,14;3,95] p=0,558	0,10 [-0,26;0,46] p=0,576
Allgemeine Gesundheit Tofacitinib + MTX Adalimumab + MTX	81 76	35,9±8,87 34,6±8,01	60 61	42,2±10,60 40,3±9,74	6,0±8,07 6,0±7,78	1,14 [-1,20;3,49] p=0,338	0,16 [-0,19;0,52] p=0,364

Tofacitinib (XELJANZ®)

	1			1			
Körperliche Rollenfunktion Tofacitinib + MTX Adalimumab + MTX	81 76	33,6±9,76 34,6±9,51	60 61	40,7±9,07 40,6±9,50	6,9±7,50 5,9±9,78	1,08 [-1,76;3,92] p=0,454	0,13 [-0,23;0,49] p=0,475
Emotionale Rollenfunktion Tofacitinib + MTX Adalimumab + MTX	81 76	33,9±13,56 35,9±11,43	60 61	39,9±10,27 39,0±12,45	5,1±10,54 2,6±12,74	1,63 [-1,99;5,25] p=0,376	0,15 [-0,21;0,51] p=0,401
Psychisches Wohlbefinden Tofacitinib + MTX Adalimumab + MTX	81 76	38,5±12,10 39,3±11,79	60 61	44,5±9,72 43,6±11,31	4,4±9,72 5,0±10,50	0,10 [-2,84;3,05] p=0,945	0,01 [-0,35;0,37] p=0,948
Soziale Funktion  Tofacitinib + MTX  Adalimumab + MTX	81 76	37,3±11,09 33,8±11,53	60 61	43,5±10,67 42,4±11,21	6,0±11,38 7,7±11,34	-0,13 [-3,27;3,01] p=0,933	-0,01 [-0,37;0,34] p=0,936
Population d)-Kombi: mehr als ein c	sDMARD	in Vortherapie					
Körperl. Komponentenscore Tofacitinib + MTX Adalimumab + MTX	102 104	33,5±7,88 33,2±6,75	79 82	41,3±8,95 40,5±8,82	8,1±8,02 7,6±7,65	0,63 [-1,45;2,71] p=0,551	0,08 [-0,23;0,39] p=0,594
Psych. Komponentenscore Tofacitinib + MTX Adalimumab + MTX	102 104	40,4±10,36 40,9±11,99	79 82	44,7±10,15 44,7±10,23	4,7±10,76 4,2±10,66	0,20 [-2,35;2,75] p=0,878	0,02 [-0,29;0,33] p=0,890
Vitalität  Tofacitinib + MTX  Adalimumab + MTX	102 104	42,1±9,24 40,3±9,40	79 83	48,2±9,16 46,3±9,90	6,2±8,73 6,2±9,12	0,38 [-2,03;2,79] p=0,756	0,04 [-0,27;0,35] p=0,781
Körperliche Funktion Tofacitinib + MTX Adalimumab + MTX	102 104	33,2±9,55 32,1±8,66	79 84	40,4±10,88 38,9±11,14	7,9±9,51 6,9±8,79	1,19 [-1,23;3,61] p=0,334	0,14 [-0,17;0,44] p=0,386
Körperliche Schmerzen Tofacitinib + MTX Adalimumab + MTX	102 104	33,1±7,51 33,6±7,33	79 84	43,1±9,71 42,7±9,21	10,5±10,11 9,8±9,32	0,55 [-1,95;3,06] p=0,665	0,06 [-0,25;0,37] p=0,700

Stand: 21.04.2017

Seite 186 von 442

Tofacitinib (XELJANZ®)

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

Allgemeine Gesundheit Tofacitinib + MTX Adalimumab + MTX	102 104	35,0±8,85 35,5±7,89	79 84	39,7±9,14 40,4±9,12	5,2±8,24 5,3±8,28	-0,32 [-2,47;1,82] p=0,767	-0,04 [-0,35;0,27] p=0,790
Körperliche Rollenfunktion Tofacitinib + MTX Adalimumab + MTX	102 104	34,6±8,79 34,9±8,29	79 84	41,6±9,60 41,1±9,22	6,9±9,43 7,3±8,82	-0,07 [-2,40;2,27] p=0,956	-0,01 [-0,32;0,30] p=0,961
Emotionale Rollenfunktion Tofacitinib + MTX Adalimumab + MTX	102 104	34,5±11,61 35,0±12,33	79 83	39,8±11,94 40,2±11,06	5,9±13,90 5,8±12,22	-0,25 [-3,28;2,79] p=0,874	-0,02 [-0,33;0,29] p=0,887
Psychisches Wohlbefinden Tofacitinib + MTX Adalimumab + MTX	102 104	39,7±10,64 40,0±11,33	79 83	44,9±9,92 44,3±10,55	5,7±9,53 4,7±10,11	0,83 [-1,70;3,36] p=0,520	0,09 [-0,22;0,40] p=0,562
Soziale Funktion Tofacitinib + MTX Adalimumab + MTX	102 104	36,5±10,43 37,9±11,03	79 84	42,5±10,28 43,3±10,70	6,1±11,48 5,4±10,07	0,28 [-2,30;2,86] p=0,832	0,03 [-0,28;0,34] p=0,848

Stand: 21.04.2017

Quelle: (40)

 $Abk\"{u}rzungen: SD = Standardabweichung, MTX = Methotrexat$ 

Auch die Analyse der mittleren Änderungen liefert für die jeweiligen Komponentenscores und Einzeldomänen für beide Populationen c) und d) keine statistisch signifikanten Unterschiede zwischen den beiden Therapien.

Tofacitinib (XELJANZ®)

Seite 187 von 442

<sup>&</sup>lt;sup>a</sup> aus einem Modell für wiederholte Messwerte mit gemischten Effekten (feste Effekte: Behandlung, Studienzeitpunkt, Behandlung×Studienzeitpunkt, Region, Baseline-Wert; zufälliger Effekt: Patient)

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

#### 4.3.1.3.13 FACIT-F nach 12 Monaten – RCT

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

Tabelle 4-67: Operationalisierung von FACIT-F nach 12 Monaten

Studie	Operationalisierung
A3921064	Der FACIT-F ist ein Instrument zur Erfassung der Erschöpfung (Fatigue).
	Die Fragebögen wurden von den Patienten jeweils zu Beginn jeden Visits ausgefüllt. Der FACIT-F hat einen Wertebereich von 0 bis 52; je höher der Wert, desto größer die Erschöpfung.
	Es werden Responderanalysen zu Monat 12 dargestellt. Ein Patient gilt demnach als Responder, falls sein Score sich nach 12 Monaten Studiendauer um mind. 4 Punkte verbessert hat. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der Definition.
	Als Sensitivitätsanalysen werden zusätzlich die Effektschätzer gezeigt, die sich ohne Imputation ergeben sowie die, die mittels LOCF-Imputation resultieren.
	Der Endpunkt wird für die FAS Population analysiert.

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

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

Tabelle 4-68: Bewertung des Verzerrungspotenzials für FACIT-F nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber (in diesem Fall der Patient) keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-69: Ergebnisse für FACIT- $F \ge 4$  nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib + MTX			mumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	n csDMARD in	Vortherapie		
A3921064	37/79		41/75		0,9 [0,6-1,2], p=0,332
	(46,8)		(54,7)		0,7 [0,4-1,4], p=0,332
					-7,8% [-23,6%-7,9%], p=0,330
		3,0		1,0	0,8 [0,5-1,1], p=0,139
		[1,0-6,0]		[1,0-3,0]	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	53/100		51/103		1,1 [0,8-1,4], p=0,620
	(53,0)		(49,5)		1,2 [0,7-2,0], p=0,620
					3,5% [-10,3%-17,2%], p=0,619
		3,0		1,0	0,9 [0,7-1,3], p=0,565
		[1,0-3,0]		[1,0-3,0]	-
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

## Population c)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (54,7% vs. 46,8%). Der Unterschied ist nicht statistisch signifikant.

## Population d)-Kombi:

Die Responderrate ist höher unter Tofacitinib + MTX (53,0% vs. 49,5%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 50: Kaplan-Meier Kurven für den Endpunkt FACIT-F  $\geq$  4 nach 12 Monaten (Population c)-Kombi)

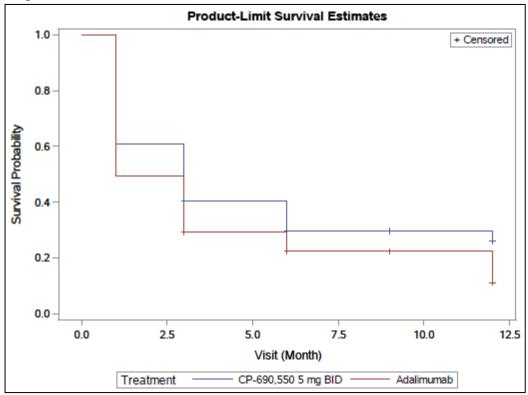


Abbildung 51: Kaplan-Meier Kurven für den Endpunkt FACIT- $F \ge 4$  nach 12 Monaten (Population d)-Kombi)

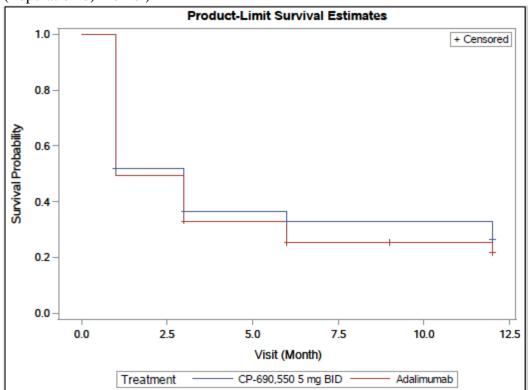


Tabelle 4-70: Sensitivitätsanalysen für FACIT-F ≥ 4 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

der/N Anza		RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
MARD in Vortherapie		
	44.64	
	41/61 (67,2)	0,9 [0,7-1,2], p=0,412 0,7 [0,4-1,5], p=0,410 -7,2% [-24,3%-9,9%], p=0,408
	47/75 (62,7)	0,9 [0,7-1,2], p=0,471 0,8 [0,4-1,5], p=0,471 -5,7% [-21,2%-9,76%], p=0,470
n csDMARD in Vorthe	rapie	
	49/83 (59,0)	1,1 [0,9-1,4], p=0,373 1,3 [0,7-2,5], p=0,373 6,8% [-8,1%-21,7%], p=0,371
	62/103 (60,2)	1,0 [0,8-1,3], p=0,978 1,0 [0,6-1,7], p=0,978 -0,2% [-13,7%-13,3%], p=0,978

Abkürzungen: RR = Relatives Risiko, OR = Odds Ratio, 95%-KI = 95%-Konfidenzintervall, MTX = Methotrexat, LOCF = Last Observation Carried Forward, NA = nicht angebbar

## Population c)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

## Population d)-Kombi:

Die Ergebnisse der Sensitivitätsanalyse ohne Imputation unterstützen die Ergebnisse der primären Analyse mittels NRI. Mittels LOCF sind die Raten vergleichbar in beiden Studienarmen. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## **4.3.1.3.14 HAQ-DI nach 12 Monaten – RCT**

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

Tabelle 4-71: Operationalisierung von HAQ-DI nach 12 Monaten

Studie	Operationalisierung					
A3921064	Der HAQ-DI ( <i>Health Assessment Questionnaire – Disability Index</i> ) ist ein Instrument zur Erfassung der Behinderung durch RA. Es werden acht alltägliche Funktionsbereiche (Anziehen, Aufstehen, Essen, Gehen, Körperpflege, Gegenstände reichen, Greifen, weitere Tätigkeiten) abgefragt und in einem Score zusammengefasst. Für jeden der acht Funktionsbereiche werden die Werte 0 (keine Einschränkung), 1 (leichte Einschränkung), 2 (starke Einschränkung) oder 3 (unfähig, die Tätigkeit auszuführen) vergeben. Der HAQ-DI Score berechnet sich dann als Mittelwert aus diesen acht Werten.					
	Die Fragebögen wurden von den Patienten jeweils zu Beginn jeden Visits ausgefüllt. Der HADI hat einen Wertebereich von 0 bis 3; je höher der Wert, desto größer die Behinderung.					
	Es werden Responderanalysen zu Monat 12 mit unterschiedlichen Schwellenwerten dargestellt. Ein Patient gilt demnach als Responder, falls sein Score sich nach 12 Monaten Studiendauer um mind. 0,22 oder um 0,3 oder um 0,5 Punkte verbessert hat. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.					
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der jeweils zutreffenden Definition.					
	Als Sensitivitätsanalysen werden zusätzlich die Effektschätzer gezeigt, die sich ohne Imputation ergeben sowie die, die mittels LOCF-Imputation resultieren.					
	Der Endpunkt wird für die FAS Population analysiert.					

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

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

Tabelle 4-72: Bewertung des Verzerrungspotenzials für HAQ-DI nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber (in diesem Fall der Patient) keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-73: Ergebnisse für HAQ-DI ≥ 0,22 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib + MTX			mumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N	Median (Monate) [95%-KI]	Anzahl Responder/ N	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	in csDMARD in	Vortherapie		
A3921064	49/79		49/75		1,0 [0,8-1,2], p=0,670
	(62,0)		(65,3)		0,9 [0,5-1,7], p=0,670
					-3,3% [-18,5%-11,9%], p=0,669
		1,0		1,0	1,0 [0,7-1,4], p=0,952
		[1,0-3,0]		[1,0-3,0]	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	56/100		65/103		0,9 [0,7-1,1], p=0,304
	(56,0)		(63,1)		0,7 [0,4-1,3], p=0,303
					-7,1% [-20,6%-6,4%], p=0,301
		1,0		1,0	0,9 [0,7-1,3], p=0,620
		[1,0-3,0]		[NA-NA]	
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

## Population c)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (65,3% vs. 62,0%). Der Unterschied ist nicht statistisch signifikant.

## Population d)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (63,1% vs. 56,0%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 52: Kaplan-Meier Kurven für den Endpunkt HAQ-DI  $\geq$  0,22 nach 12 Monaten (Population c)-Kombi)

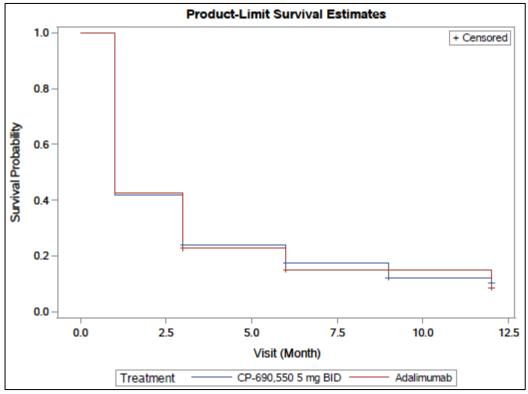


Abbildung 53: Kaplan-Meier Kurven für den Endpunkt HAQ-DI  $\geq$  0,22 nach 12 Monaten (Population d)-Kombi)

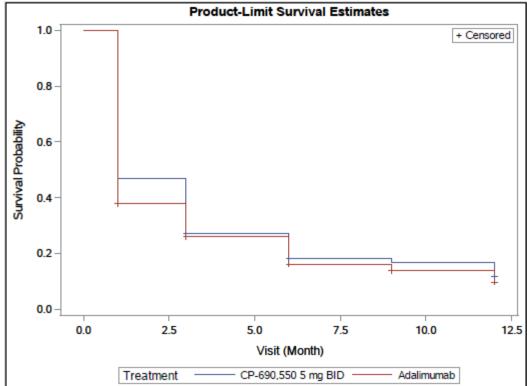


Tabelle 4-74: Ergebnisse für HAQ-DI ≥ 0,3 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib + MTX			mumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	in csDMARD in	Vortherapie		
A3921064	40/79 (50,6)		46/75 (61,3)		0,8 [0,6-1,1], p=0,183 0,7 [0,3-1,2], p=0,182 -10,7% [-26,3%-4,9%], p=0,179
		3,0 [1,0-3,0]		3,0 [1,0-3,0]	1,0 [0,7-1,5], p=0,834
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	50/100 (50,0)		58/103 (56,3)		0,9 [0,7-1,2], p=0,369 0,8 [0,5-1,4], p=0,368 -6,3% [-20,0%-7,4%], p=0,367
		3,0 [1,0-3,0]		1,0 [1,0-3,0]	0,9 [0,7-1,2], p=0,538
Ū	en: RR = Re		o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

# Population c)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (61,3% vs. 50,6%). Der Unterschied ist nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (56,3% vs. 50,0%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 54: Kaplan-Meier Kurven für den Endpunkt HAQ-DI  $\geq 0.3$  nach 12 Monaten (Population c)-Kombi)

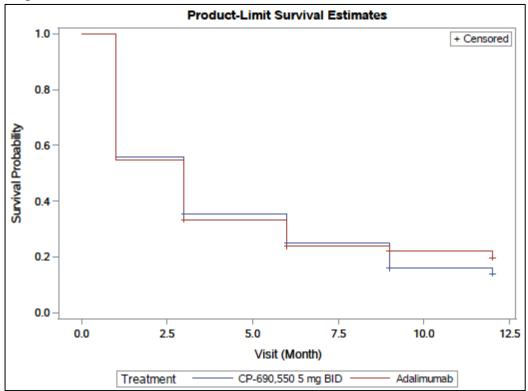


Abbildung 55: Kaplan-Meier Kurven für den Endpunkt HAQ-DI  $\geq 0.3$  nach 12 Monaten (Population d)-Kombi)

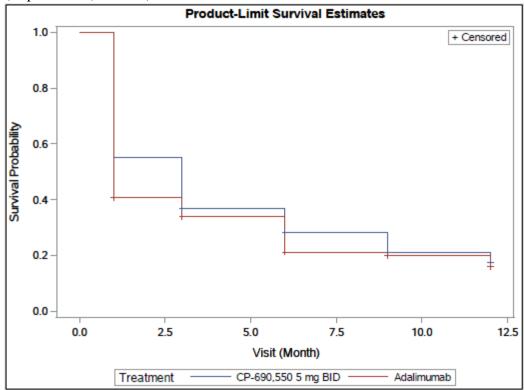


Tabelle 4-75: Ergebnisse für HAQ-DI ≥ 0,5 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie		citinib ATX		mumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab	
	Anzahl Responder N	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
					HR [95%-KI], p-Wert	
Population c	-Kombi: nur e	in csDMARD in	Vortherapie			
A3921064	33/79		39/75		0,8 [0,6-1,1], p=0,206	
	(41,8)		(52,0)		0,7 [0,4-1,3], p=0,204	
					-10,2% [-25,9%-5,5%], p=0,201	
		3,0		3,0	1,0 [0,7-1,5], p=0,953	
		[3,0-6,0]		[3,0-6,0]		
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e		
A3921064	45/100		52/103		0,9 [0,7-1,2], p=0,435	
	(45,0)		(50,5)		0,8 [0,5-1,4], p=0,434	
					-5,5% [-19,2%-8,2%], p=0,433	
		3,0		3,0	0,9 [0,6-1,2], p=0,343	
		[3,0-6,0]		[1,0-6,0]	-	
Quelle: (40)						
Ū			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-	

# Population c)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (52,0% vs. 41,8%). Der Unterschied ist nicht statistisch signifikant.

# Population d)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (50,5% vs. 45,0%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 56: Kaplan-Meier Kurven für den Endpunkt HAQ-DI ≥ 0,5 nach 12 Monaten (Population c)-Kombi)

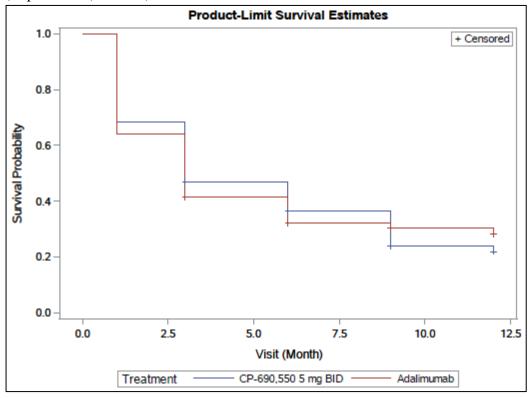


Abbildung 57: Kaplan-Meier Kurven für den Endpunkt HAQ-DI  $\geq 0,\!5$ nach 12 Monaten (Population d)-Kombi)

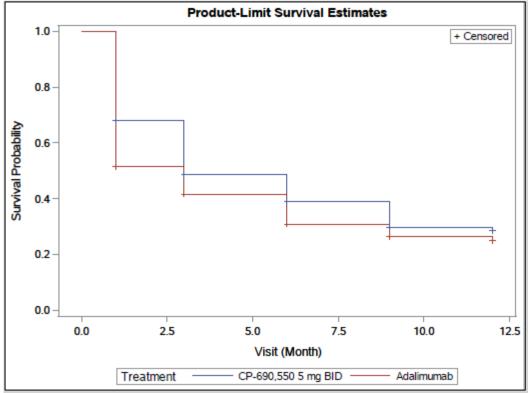


Tabelle 4-76: Sensitivitätsanalysen für HAQ-DI ≥ 0,22 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)-K	Kombi: nur ein csDMARD in V	ortherapie	
keine Imputation	49/59 (83,1)	49/61 (80,3)	1,0 [0,9-1,2], p=0,700 1,2 [0,5-3,0], p=0,700 2,7% [-11,1%-16,6%], p=0,700
LOCF	60/79 (76,0)	55/75 (73,3)	1,0 [0,9-1,3], p=0,710 1,2 [0,6-2,4], p=0,709 2,6% [-11,1%-16,4%], p=0,709
Population d)-F	Kombi: mehr als ein csDMARI	) in Vortherapie	
keine Imputation	55/79 (69,6)	63/83 (75,9)	0,9 [0,8-1,1], p=0,372 0,7 [0,4-1,5], p=0,370 -6,3% [-20,0%-7,4%], p=0,369
LOCF	68/100 (68,0)	74/103 (71,8)	1,0 [0,8-1,1], p=0,551 0,8 [0,5-1,5], p=0,551

Abkürzungen: RR = Relatives Risiko, OR = Odds Ratio, 95%-KI = 95%-Konfidenzintervall, MTX = Methotrexat, LOCF = Last Observation Carried Forward, NA = nicht angebbar

## Population c)-Kombi:

Im Gegensatz zur primären Analyse mittels NRI liefern die Ergebnisse der Sensitivitätsanalysen höhere Responderraten unter Tofacitinib + MTX. Der Unterschied ist jeweils nicht statistisch signifikant.

## Population d)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

Tabelle 4-77: Sensitivitätsanalysen für HAQ-DI ≥ 0,3 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)-l	Kombi: nur ein csDMARD in V	ortherapie ortherapie	
keine Imputation	40/59 (67,8)	46/61 (75,4)	0,9 [0,7-1,1], p=0,358 0,7 [0,3-1,5], p=0,356 -7,6% [-23,7%-8,5%], p=0,354
LOCF	51/79 (64,6)	51/75 (68,0)	1,0 [0,8-1,2], p=0,651 0,9 [0,4-1,7], p=0,652 -3,4% [-18,4%-11,5%], p=0,651
Population d)-	Kombi: mehr als ein csDMARI	) in Vortherapie	
keine Imputation	49/79 (62,0)	56/83 (67,5)	0,9 [0,7-1,2], p=0,470 0,8 [0,4-1,5], p=0,469 -5,4% [-20,2%-9,3%], p=0,468
LOCF	62/100 (62,0)	66/103 (64,1)	1,0 [0,8-1,2], p=0,759 0,9 [0,5-1,6], p=0,759 -2,1% [-15,4%-11,2%], p=0,759
_		OR = Odds Ratio, 95%-KI rried Forward, NA = nicht ange	= 95%-Konfidenzintervall, MTX = ebbar

## Population c)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

## Population d)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

Tabelle 4-78: Sensitivitätsanalysen für HAQ-DI ≥ 0,5 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)-l	Kombi: nur ein csDMARD in V	ortherapie ortherapie	
keine Imputation	33/59 (55,9)	39/61 (63,9)	0,9 [0,7-1,2], p=0,374 0,7 [0,3-1,5], p=0,372 -8,0% [-25,5%-9,5%], p=0,370
LOCF	43/79 (54,4)	42/75 (56,0)	1,0 [0,7-1,3], p=0,845 0,9 [0,5-1,8], p=0,845 -1,6% [-17,3%-14,1%], p=0,845
Population d)-	Kombi: mehr als ein csDMARI	) in Vortherapie	
keine Imputation	45/79 (57,0)	50/83 (60,2)	1,0 [0,7-1,2], p=0,672 0,9 [0,5-1,6], p=0,672 -3,3% [-18,5%-11,9%], p=0,672
LOCF	56/100 (56,0)	59/103 (57,3)	1,0 [0,8-1,2], p=0,854 1,0 [0,5-1,7], p=0,854 -1,3% [-14,9%-12,4%], p=0,854
_		OR = Odds Ratio, 95%-KI ried Forward, NA = nicht ange	= 95%-Konfidenzintervall, MTX = ebbar

## Population c)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

## Population d)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## **4.3.1.3.15 WLQ nach 12 Monaten – RCT**

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

Tabelle 4-79: Operationalisierung von WLQ nach 12 Monaten

Studie	Operationalisierung						
A3921064	Der WLQ ( <i>Work Limitations Questionnaire</i> ) ist ein Instrument zur Erfassung der durch Krankheit verursachten Einschränkungen im Arbeitsleben. Der WLQ besteht aus vier Einzeldimensionen und einem Indexscore zur Erfassung des allgemeinen Arbeitsverlusts:						
	• Indexscore						
	• Zeitmanagement						
	Körperliche Anforderungen						
	Mentale/Interpersonelle Anforderungen						
	Ergebnisorientierte Anforderungen						
	Die Fragebögen wurden von den Patienten jeweils zu Baseline und zu Beginn von Monat 3, 6 und 12 ausgefüllt. Die einzelnen Dimensionen haben jeweils einen Wertebereich von 0 bis 100, der Indexscore von 0 bis 28,6; je höher der Wert, desto höher die Beeinträchtigung der Arbeitsleistung.						
	Es werden jeweils Responderanalysen zu Monat 12 dargestellt mit den Schwellenwerten von mind. 5 (Indexscore) bzw. 13 (pro Dimension) Punkten Verbesserung. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.						
	Des Weiteren wird die mittlere Änderung zu Monat 12 gegenüber Baseline inklusive des standardisierten Effektmaßes Hedges' g berichtet (feste Effekte: Behandlung, Studienzeitpunkt, Behandlung×Studienzeitpunkt, Region, Baseline-Wert; zufälliger Effekt: Patient). Fehlende Werte werden hierbei nicht ersetzt.						
	Der Endpunkt wird für die FAS Population analysiert.						
	Die Fragebögen des WLQ wurden nur von 30 – 40% der Patienten ausgefüllt, die zum Zeitpunkt der Studiendurchführung berufstätig waren. Daher sind die Patientenzahlen, die in die Analyse dieses Endpunkts eingehen, deutlich niedriger als diejenigen des FAS Kollektivs.						
	Der Anteil der fehlenden Werte ist für diese berufstätigen Patienten durch die niedrige Rücklaufquote nochmals deutlich reduziert: ca. 40-60 % für Teilpopulation c)-Kombi und ca. 65-70 % für Teilpopulation d)-Kombi in den jeweiligen Studienarmen. Die Ergebnisse für den WLQ sind daher als hoch verzerrt einzustufen und können daher nicht sinnvoll interpretiert werden.						

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

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

Tabelle 4-80: Bewertung des Verzerrungspotenzials für WLQ nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	unklar	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber (in diesem Fall der Patient) keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Der WLQ kann nur von Patienten ausgefüllt werden, die berufstätig sind. Insofern ist von vornherein eine geringere Anzahl von Patienten mit Daten zu diesem Endpunkt zu erwarten, als es dem ITT-Prinzip entsprechen würde. Alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt, gehen in die Auswertung des WLQ ein. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig.

Es ist schwer zu beurteilen, ob die Verteilung berufstätiger Studienpatienten auf die Studienarme zu einer Verzerrung beitragen könnte (bezüglich dieses Faktors wurde nicht stratifiziert randomisiert). Jedoch ist diese Unsicherheit im Zusammenhang mit der insgesamt niedrigen Anzahl von Daten dazu angetan, das Verzerrungspotenzial des WLQ auf Endpunktebene als hoch einzustufen. Auch der hohe Anteil von fehlenden Werten (ca. 40-60% für Teilpopulation c)-Kombi und ca. 30-35% für Teilpopulation d)-Kombi in den jeweiligen Studienarmen) führt dazu, dass eine sinnvolle Interpretation der Ergebnisse des WLQ nicht möglich ist.

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

Tabelle 4-81: Ergebnisse für WLQ nach 12 Monaten – Responderanalysen aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)-Kombi: nur ein csDMARD in	Vortherapie		
Indexscore ≥ 5	7/25 (28,0)	8/32 (25,0)	1,1 [0,5-2,7], p=0,798 1,2 [0,4-3,8], p=0,799 3,0% [-20,1%-26,1%], p=0,799
Zeitmanagement ≥ 13	7/23 (30,4)	18/28 (64,3)	0,5 [0,2-0,9], p=0,030 0,2 [0,1-0,8], p=0,019 -33,9% [-59,7%;-8,0%], p=0,010
Körperliche Anforderungen ≥ 13	5/25 (20,0)	9/31 (29,0)	0,7 [0,3-1,8], p=0,446 0,6 [0,2-2,1], p=0,440 -9,0% [-31,4%-13,4%], p=0,429
Mentale/Interpersonelle Anforderungen ≥ 13	8/24 (33,3)	9/31 (29,0)	1,2 [0,5-2,5], p=0,732 1,2 [0,4-3,9], p=0,732 4,3% [-20,4%-29,0%], p=0,733
Ergebnisorientierte Anforderungen ≥ 13	7/23 (30,4)	14/31 (45,2)	0,7 [0,3-1,4], p=0,289 0,5 [0,2-1,7], p=0,275 -14,7% [-40,4%-11,0%], p=0,261
Population d)-Kombi: mehr als ein csDMAF	RD in Vortherapie	2	
Indexscore ≥ 5	10/38 (26,3)	3/34 (8,8)	2,9 [0,9-10,0], p=0,075 3,7 [0,9-14,8], p=0,065 17,5% [0,6%-34,4%], p=0,043
Zeitmanagement ≥ 13	21/37 (56,8)	11/32 (34,4)	1,7 [0,95-2,9], p=0,077 2,5 [0,9-6,7], p=0,066 22,4% [-0,6%-45,3%], p=0,056
Körperliche Anforderungen ≥ 13	6/37 (16,2)	10/32 (31,3)	0,5 [0,2-1,3], p=0,151 0,4 [0,1-1,4], p=0,150 -15,0% [-35,0%-4,9%], p=0,140
Mentale/Interpersonelle Anforderungen ≥ 13	17/37 (46,0)	5/32 (15,6)	2,9 [1,2-7,1], p=0,016 4,6 [1,5-14,5], p=0,010 30,3% [9,9%-50,7%], p=0,004
Ergebnisorientierte Anforderungen ≥ 13	15/36 (41,7)	13/32 (40,6)	1,0 [0,6-1,8], p=0,931 1,0 [0,4-2,8], p=0,931

1,0% [-22,4%-24,5%], p=0,931

Stand: 21.04.2017

Quelle: (40)

Abkürzungen: RR = Relatives Risiko, OR = Odds Ratio, HR = Hazard Ratio, 95%-KI = 95%-Konfidenzintervall, MTX = Methotrexat, NA = nicht angebbar

### Population c)-Kombi:

Für die Dimension "Zeitmanagement" liegt die Responderrate unter Adalimumab + MTX statistisch signifikant höher als unter Tofacininib + MTX (64,3% vs. 30,4%; p=0,030). Auch für "Körperliche Anforderungen" und "Ergebnisorientierte Anforderungen" sind die Raten jeweils höher im Adalimumab-Arm (29,0% vs. 20,0% bzw. 45,2% vs. 30,4%). Diese Unterschiede sind jedoch nicht statistisch signifikant. Für den Indexscore sowie "Mentale/Interpersonelle Anforderungen" sind die Responderraten höher unter Tofacitinb + MTX (28,0% vs. 25,0% bzw. 33,3% vs. 29,0%). Diese Unterschiede sind jedoch ebenfalls jeweils nicht statistisch signifikant.

### Population d)-Kombi:

Für die Dimensionen "Mentale/Interpersonelle Anforderungen" liegt die Responderrate unter Tofacitinib + MTX statistisch signifikant höher als unter Adalimumab + MTX (46,0% vs. 15,6%; p=0,016). Auch für den Indexscore und "Zeitmanagement" sind die Raten jeweils höher im Tofacitinib-Arm (26,3% vs. 8,8% bzw. 56,8% vs. 34,4%). Diese Unterschiede sind jedoch nicht statistisch signifikant. Für "Körperliche Anforderungen" liegt die Responderrate höher im Adalimumab-Arm (31,3% vs. 16,2%), wobei der Unterschied jedoch nicht statistisch signifikant ist. Die Responderraten für "Ergebnisorientierte Anforderungen" sind vergleichbar in beiden Studienarmen.

Durch die oben beschriebene Einstufung diese Endpunkts als hochverzerrt können diese Ergebnisse für beide Teilpopulationen jedoch nicht sinnvoll interpretiert werden.

Stand: 21.04.2017

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

Tabelle 4-82: Ergebnisse für WLQ nach 12 Monaten – mittlere Änderungen aus RCT mit dem zu bewertenden Arzneimittel

Studie			Desk	riptive Analyse		Adjustierte	Analyse <sup>a</sup> :
A3921064		Baseline		Monat 12	_	Mittlere	Mittlere
					Mittlere Änderung zwischen Baseline und Monat 12	Änderung über 12 Monate Differenz	Änderung über 12 Monate Hedges'g
	N	Mittelwert <u>+</u> SD	N	Mittelwert <u>+</u> SD	Mittelwert ± SD	Tofacitinb – Adalimumab [95%-KI] p-Wert	Tofacitinb – Adalimumab [95%-KI] p-Wert
Population c)-Kombi: nur ein csDM	IARD in Vo	ortherapie					
Indexscore  Tofacitinib + MTX  Adalimumab + MTX	33 36	9,9±5,35 10,6±5,19	15 21	6,1±6,21 6,4±4,97	-2,7±5,63 -3,2±4,52	-1,47 [-4,64;1,70] p=0,361	-0,27 [-0,95;0,40] p=0,420
<b>Zeitmanagement</b> Tofacitinib + MTX Adalimumab + MTX	31 33	37,6±26,39 50,6±25,49	13 19	30,0±31,42 22,5±226,52	-5,0±24,92 -26,1±31,53	8,65 [-9,71;27,00] p=0,352	0,30 [-0,42;1,02] p=0,402
Körperliche Anforderungen Tofacitinib + MTX Adalimumab + MTX	33 35	51,3±24,79 43,2±19,96	13 21	39,7±37,62 53,3±33,81	-3,3±39,62 10,7±36,40	-12,00 [-30,03;6,03] p=0,190	-0,41 [-1,12;0,30] p=0,245
Mentale/Interpersonelle Anforderungen Tofacitinib + MTX Adalimumab + MTX	31 34	25,7±21,41 30,9±26,72	14 21	16,0±21,62 13,4±16,34	-5,7±19,63 -11,5±21,68	-2,39 [-16,72;11,93] p=0,741	-0,10 [-0,79;0,58] p=0,763
Ergebnisorientierte Anforderungen Tofacitinib + MTX Adalimumab + MTX	31 35	40,2±28,69 39,1±20,91	15 21	20,8±26,45 20,7±22,15	-14,3±32,22 -13,0±24,16	-6,34 [-21,01;8,32] p=0,392	-0,25 [-0,92;0,43] p=0,458
Population d)-Kombi: mehr als ein	csDMARD	in Vortherapie				•	•
Indexscore  Tofacitinib + MTX  Adalimumab + MTX	42 38	9,2±5,15 9,1±4,84	29 28	6,0±3,28 6,7±5,44	-2,89±5,49 -1,51±4,50	-0,66 [-2,72;1,40]	-0,14 [-0,67;0,38]

Tofacitinib (XELJANZ®)

Seite 211 von 442

# Dossier zur Nutzenbewertung – Modul 4 A

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

						p=0,528	p=0,582
Zeitmanagement Tofacitinib + MTX Adalimumab + MTX	40 37	42,4±21,88 41,9±26,36	28 25	19,4±19,41 28,0±26,15	-19,8±24,61 -9,2±31,77	-8,92 [-20,63;2,79] p=0,134	-0,37 [-0,92;0,18] p=0,185
Körperliche Anforderungen Tofacitinib + MTX Adalimumab + MTX	42 36	48,5±22,35 44,5±21,69	29 26	66,2±27,67 47,4±34,47	18,7±36,00 1,7±36,46	14,97 [0,36;29,58] p=0,045	0,49 [-0,05;1,03] p=0,076
Mentale/Interpersonelle							
Anforderungen						-9,49	-0,53
Tofacitinib + MTX	41	$24,4\pm22,01$	27	9,3±10,74	-11,6±16,16	[-17,98;-1,00]	[-1,08;0,02]
Adalimumab + MTX	37	$25,2\pm21,74$	26	19,2±20,56	-2,33±16,52	p=0,029	p=0,060
Ergebnisorientierte							
Anforderungen						-1,96	-0,09
Tofacitinib + MTX	40	$31,6\pm24,82$	28	17,2±17,71	-14,5±32,30	[-12,51;8,59]	[-0,63;0,46]
Adalimumab + MTX	37	$33,8\pm24,58$	25	21,9±25,01	-10,1±24,27	p=0,714	p=0,749

Stand: 21.04.2017

Quelle: (40)

Abkürzungen: SD = Standardabweichung, MTX = Methotrexat

<sup>&</sup>lt;sup>a</sup> aus einem Modell für wiederholte Messwerte mit gemischten Effekten (feste Effekte: Behandlung, Studienzeitpunkt, Behandlung×Studienzeitpunkt, Region, Baseline-Wert; zufälliger Effekt: Patient)

### Population c)-Kombi:

Die Analyse der mittleren Änderungen zeigt einen numerischen Vorteil von Adalimumab + MTX gegenüber Tofacitinib + MTX in der Dimension "Zeitmanagement". Im Gegensatz zur obigen Responderanalyse ist der Unterschied nicht statistisch signifikant. Für den Indexscore und die übrigen Dimensionen liegt jeweils ein Vorteil im Tofacitinib-Arm vor. Die beobachteten Unterschiede sind jedoch nicht statistisch signifikant.

#### Population d)-Kombi:

Die Analyse der mittleren Änderungen zeigt einen statistisch signifikanten Vorteil von Adalimumab + MTX gegenüber Tofacitinib + MTX in der Dimension "Körperliche Anforderungen" (p=0,045). Für die Dimension "Mentale/Interpersonelle Anforderungen" liegt ein statistisch signifikanter Vorteil im Tofacitinib-Arm vor (p=0,029). Der standardisierte Effektschätzer (Hedges' g) ist jedoch für diese beiden Dimensionen statistisch nicht signifikant. Für den Indexscore, für "Zeitmanagement" sowie für "Ergebnisorientierte Anforderungen" sind die mittleren Änderungen jeweils numerisch etwas ausgeprägter im Tofacitinib-Arm. Diese Unterschiede sind jedoch nicht statistisch signifikant.

Bei der Interpretation der Ergebnisse ist zu beachten, dass die Fallzahlen der Patienten, die den WLQ-Fragebogen ausgefüllt haben, im Vergleich zur Gesamtpopulation deutlich niedriger sind. Dieser Aspekt wie auch der hohe Anteil von fehlenden Werten führen dazu, dass eine sinnvolle Interpretation der Ergebnisse des WLQ nicht möglich ist.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

#### 4.3.1.3.16 MOS Schlafscore nach 12 Monaten – RCT

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

Tabelle 4-83: Operationalisierung von MOS Schlafscore nach 12 Monaten

Studie	Operationalisierung
A3921064	Der MOS Schlafscore ist ein Instrument zur Erfassung von Schlafproblemen. Der Score besteht aus sieben Subskalen und einem Summenscore, der die Schlafprobleme quantifiziert. In diesem Dossier werden die Ergebnisse des Summenscores dargestellt.
	Die Fragebögen wurden von den Patienten jeweils zu Beginn jeden Visits ausgefüllt. Der MOS Schlafscore hat einen Wertebereich von 0 bis 100; je höher der Wert, desto größer die Schlafprobleme.
	Es werden Responderanalysen zu Monat 12 dargestellt. Ein Patient gilt demnach als Responder, falls sein Score sich nach 12 Monaten Studiendauer um mind. 6 Punkte verbessert hat. Fehlende Werte werden gemäß der sog. <i>Nonresponder Imputation</i> (NRI) ausgewertet, bei welcher für Patienten mit fehlenden Werten zu Monat 12 LOCF angewandt wird und Studienabbrecher generell als Nonresponder gezählt werden. Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Zusätzlich werden Hazard Ratios für <i>Time-to-Event</i> Analysen angegeben und Kaplan-Meier Kurven dargestellt. Das Ereignis ( <i>event</i> ) ist dabei eine eingetretene Response entsprechend der Definition.
	Als Sensitivitätsanalysen werden zusätzlich die Effektschätzer gezeigt, die sich ohne Imputation ergeben sowie die, die mittels LOCF-Imputation resultieren.
	Der Endpunkt wird für die FAS Population analysiert.

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

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

Tabelle 4-84: Bewertung des Verzerrungspotenzials für MOS Schlafscore nach 12 Monaten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber (in diesem Fall der Patient) keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten des *Full Analysis Set* (FAS) berücksichtigt, d.h. alle randomisierten Patienten, für die mind. eine Erhebung nach Baseline vorliegt. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-85: Ergebnisse für MOS Schlafscore ≥ 6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tofacitinib + MTX			mumab MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder N (%)	Median (Monate) [95%-KI]	Anzahl Responder/ N (%)	Median (Monate) [95%-KI]	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
					HR [95%-KI], p-Wert
Population c	)-Kombi: nur ei	n csDMARD in	Vortherapie		
A3921064	29/78		33/74		0,8 [0,6-1,2], p=0,354
	(37,2)		(44,6)		0,7 [0,4-1,4], p=0,353
					-7,4% [-23,0%-8,2%], p=0,352
		1,0		3,0	0,9 [0,6-1,3], p=0,687
		[1,0-3,0]		[1,0-3,0]	
Population d	)-Kombi: mehr	als ein csDMA	RD in Vortherapi	e	
A3921064	33/99		40/103		0,9 [0,6-1,2], p=0,417
	(33,3)		(38,8)		0,8 [0,4-1,4], p=0,416
					-5,5% [-18,7%-7,7%], p=0,415
		3,0		3,0	0,9 [0,7-1,3], p=0,693
		[1,0-6,0]		[3,0-6,0]	-
Quelle: (40)					
_			o, OR = Odds NA = nicht ange		Hazard Ratio, 95%-KI = 95%-

## Population c)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (44,6% vs. 37,2%). Der Unterschied ist nicht statistisch signifikant.

## Population d)-Kombi:

Die Responderrate ist höher unter Adalimumab + MTX (38,8% vs. 33,3%). Der Unterschied ist nicht statistisch signifikant.

Abbildung 58: Kaplan-Meier Kurven für den Endpunkt MOS Schlafscore ≥ 6 nach 12 Monaten (Population c)-Kombi)

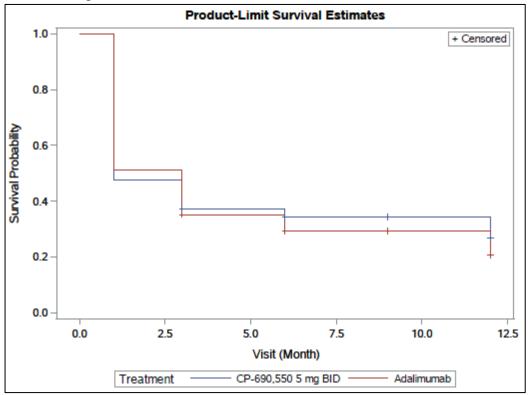


Abbildung 59: Kaplan-Meier Kurven für den Endpunkt MOS Schlafscore ≥ 6 nach 12 Monaten (Population d)-Kombi)

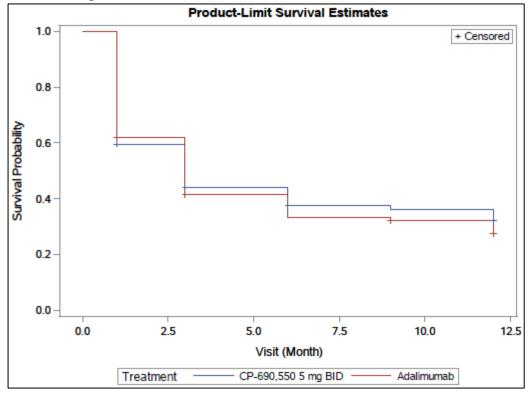


Tabelle 4-86: Sensitivitätsanalysen für MOS Schlafscore ≥ 6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI], p-Wert OR [95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
Population c)-l	Kombi: nur ein csDMARD in V	ortherapie ortherapie	
keine Imputation	28/59 (47,5)	33/61 (54,1)	0,9 [0,6-1,3], p=0,469 0,8 [0,4-1,6], p=0,467 -6,6% [-24,5%-11,2%], p=0,466
LOCF	37/78 (47,4)	38/74 (51,4)	0,9 [0,7-1,3], p=0,629 0,9 [0,5-1,6], p=0,630 -3,9% [-19,8%-12,0%], p=0,629
Population d)-	Kombi: mehr als ein csDMARI	) in Vortherapie	
keine Imputation	33/78 (42,3)	39/83 (47,0)	0,9 [0,6-1,3], p=0,552 0,8 [0,4-1,5], p=0,551 -4,7% [-20,0%-10,7%], p=0,550
LOCF	40/99 (40,4)	48/103 (46,6)	0,9 [0,6-1,2], p=0,376 0,8 [0,4-1,4], p=0,375 -6,2% [-19,9%-7,5%], p=0,373
_		OR = Odds Ratio, 95%-KI ried Forward, NA = nicht ange	= 95%-Konfidenzintervall, MTX = ebbar

## Population c)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

## Population d)-Kombi:

Die Ergebnisse der Sensitivitätsanalysen unterstützen weitgehend die Ergebnisse der primären Analyse mittels NRI. Auch hier sind keine statistisch signifikanten Unterschiede zwischen den Studienarmen zu erkennen.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.17 Unerwünschte Ereignisse bis 12 Monate – RCT

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

Tabelle 4-87: Operationalisierung von unerwünschten Ereignissen bis 12 Monate

Studie	Operationalisierung
A3921064	Als unerwünschte Ereignisse ("adverse events") werden jegliche während der Studientherapie auftretende ("all causality"), unerwünschte (medizinische) Ereignisse definiert. Sie müssen nicht unbedingt einen kausalen Zusammenhang mit der Studienmedikation aufweisen.
	Der Berichtszeitraum für unerwünschte Ereignisse umfasst den gesamten Behandlungszeitraum der Studientherapie einschließlich einer 28-tägigen Periode nach der letzten Einnahme der jeweiligen Studienmedikation. Die unerwünschten Ereignisse wurden nach MedDRA, Version 13.1, kodiert.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-88: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse bis 12 Monate in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* 

berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-89: Ergebnisse für unerwünschte Ereignisse bis 12 Monate aus RCT mit dem zu bewertenden Arzneimittel

Studie		Patienten in Analyse		atienten mit is, N (%)	RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab
Population c)-Kombi: nur ein csDMARD in Vortherapie					
A3921064	83	78	61 (73,5)	53 (67,9)	1,1 [0,9-1,3], p=0,442 1,3 [0,7-2,6], p=0,440 5,5% [-8,5%-19,6%], p=0,439
Population d)	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2	
A3921064	103	104	75 (72,8)	78 (75,0)	1,0 [0,8-1,1], p=0,721 0,9 [0,5-1,7], p=0,721 -2,2% [-14,1%-9,8%], p=0,439
Quelle: (41)					
Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio					

## Population c)-Kombi:

Die Rate der Patienten mit unerwünschten Ereignissen ist höher unter Tofacitinib + MTX (73,5% vs. 67,9%). Der Unterschied ist nicht statistisch signifikant.

## Population d)-Kombi:

Die Rate der Patienten mit unerwünschten Ereignissen ist höher unter Adalimumab + MTX (75,0% vs. 72,8%). Der Unterschied ist nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als

Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.18 Schwerwiegende unerwünschte Ereignisse bis 12 Monate – RCT

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

Tabelle 4-90: Operationalisierung von schwerwiegenden unerwünschten Ereignissen nach 12 Monaten

Studie	Operationalisierung				
A3921064	Als unerwünschte Ereignisse ("adverse events") werden jegliche während der Studientherapie auftretende ("all causality"), unerwünschte (medizinische) Ereignisse definiert. Sie müssen nicht unbedingt einen kausalen Zusammenhang mit der Studienmedikation aufweisen.				
	Der Berichtszeitraum für unerwünschte Ereignisse umfasst den gesamten Behandlungszeitraum der Studientherapie einschließlich einer 28-tägigen Periode nach der letzten Einnahme der jeweiligen Studienmedikation. Die unerwünschten Ereignisse wurden nach MedDRA Version 13.1 kodiert.  Als schwerwiegende unerwünschte Ereignisse (SUE: "serious adverse event") wurden die folgenden Ereignisse bezeichnet, wenn sie sich in Zusammenhang mit der Therapie ereigneten:				
	<ul> <li>Tod</li> <li>Lebensbedrohliches Ereignis</li> <li>Stationäre Behandlung oder eine Verlängerung einer stationären Behandlung</li> <li>Eintritt anhaltender Erwerbsunfähigkeit oder Behinderung</li> <li>Kongenitale Anomalie und Geburtsfehler</li> </ul> Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer. Fehlende Werte werden nicht ersetzt. Die Analyse erfolgte für die Safety Analysis Population.				

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

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

Tabelle 4-91: Bewertung des Verzerrungspotenzials für schwerwiegende unerwünschte Ereignisse bis 12 Monate in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-92: Ergebnisse für schwerwiegende unerwünschte Ereignisse bis 12 Monate aus RCT mit dem zu bewertenden Arzneimittel

Studie		Anzahl der Patienten in der Analyse		atienten mit is, N (%)	RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab
Population c)-Kombi: nur ein csDMARD in Vortherapie					
A3921064	83	78	13 (15,7)	4 (5,1)	3,1 [1,0-9,0], p=0,042 3,4 [1,1-11,0], p=0,038 10,5% [1,3%-19,8%], p=0,025
Population d)	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2	
A3921064	103	104	17 (16,5)	13 (12,5)	1,3 [0,7-2,6], p=0,415 1,4 [0,6-3,0], p=0,414 4,0% [-5,6%-13,6%], p=0,413
Quelle: (41)					
Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio					

## Population c)-Kombi:

Die Rate der Patienten mit schwerwiegenden unerwünschten Ereignissen ist höher unter Tofacitinib + MTX (15,7% vs. 5,1%). Der Unterschied ist statistisch signifikant (p=0,042).

## Population d)-Kombi:

Die Rate der Patienten mit schwerwiegenden unerwünschten Ereignissen ist höher unter Tofacitinib + MTX (16,5% vs. 12,5%). Der Unterschied ist nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.19 Unerwünschte Ereignisse, die bis Monat 12 zum Therapieabbruch geführt hatten– RCT

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

Tabelle 4-93: Operationalisierung von unerwünschten Ereignissen, die bis Monat 12 zum Therapieabbruch geführt hatten

Studie	Operationalisierung
A3921064	Als unerwünschte Ereignisse ("adverse events") werden jegliche während der Studientherapie auftretende ("all causality"), unerwünschte (medizinische) Ereignisse definiert. Sie müssen nicht unbedingt einen kausalen Zusammenhang mit der Studienmedikation aufweisen.
	Der Berichtszeitraum für unerwünschte Ereignisse umfasst den gesamten Behandlungszeitraum der Studientherapie einschließlich einer 28-tägigen Periode nach der letzten Einnahme der jeweiligen Studienmedikation. Die unerwünschten Ereignisse wurden nach MedDRA Version 13.1 kodiert.
	In diesem Abschnitt werden die Ergebnisse zu unerwünschten Ereignissen, die zum Therapieabbruch geführt hatten, berichtet.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-94: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse, die bis Monat 12 zum Therapieabbruch geführt hatten in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-95: Ergebnisse für unerwünschte Ereignisse, die bis Monat 12 zum Therapieabbruch geführt hatten aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse		Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab
Population c)-Kombi: nur ein csDMARD in Vortherapie					
A3921064	83	78	8 (9,6)	5 (6,4)	1,5 [0,5-4,4], p=0,457 1,6 [0,5-5,0], p=0,455 3,2% [-5,1%-11,6%], p=0,449
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapio	2	
A3921064	103	104	13 (12,6)	16 (15,4)	0,8 [0,4-1,6], p=0,568 0,8 [0,4-1,7], p=0,567 -2,8% [-12,2%-6,7%], p=0,566
Quelle: (41)					
Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio					

#### Population c)-Kombi:

Die Rate der Patienten mit unerwünschten Ereignissen, die zum Therapieabbruch geführt hatten, ist höher unter Tofacitinib + MTX (9,6% vs. 6,4%). Der Unterschied ist nicht statistisch signifikant.

## Population d)-Kombi:

Die Rate der Patienten mit unerwünschten Ereignissen, die zum Therapieabbruch geführt hatten, ist höher unter Adalimumab + MTX (15,4% vs. 12,6%). Der Unterschied ist nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.20 Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen bis 12 Monate – RCT

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

Tabelle 4-96: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Schwere Infektionen" werden alle unerwünschten Ereignisse zusammengefasst, die mit der folgenden MedDRA (Version 13.1) <i>System Organ Class</i> kodiert wurden:
	Infections and Infestations,
	und die zudem vom Prüfarzt als schwere Ereignisse eingestuft wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-97: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird

das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-98: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse		Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab	
Population c)-Kombi: nur ein csDMARD in Vortherapie						
A3921064	83	78	2 (2,4)	0 (0,0)	4,7 [0,2-96,4], p=0,315 4,8 [0,2-101,9], p=0,313 2,4% [-1,6%-6,4%], p=0,242	
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2		
A3921064	103	104	4 (3,9)	1 (1,0)	4,0 [0,5-35,5], p=0,208 4,2 [0,5-37,9], p=0,206 2,9% [-1,3%-7,1%], p=0,107	
Quelle: (41)  Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio						

## Population c)-Kombi:

Es wurden nur wenige schwere Infektionen berichtet: zwei Ereignisse im Tofacitinib-Arm (2,4%), keines im Adalimumab-Arm. Es liegt kein statistisch signifikanter Unterschied vor.

## Population d)-Kombi:

Es wurden nur wenige schwere Infektionen berichtet: vier Ereignisse im Tofacitinib-Arm (3,9%), eines im Adalimumab-Arm (1,0%). Es liegt kein statistisch signifikanter Unterschied vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

# 4.3.1.3.21 Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung bis 12 Monate – RCT

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

Tabelle 4-99: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Jedwede Infektion und parasitäre Erkrankung" werden alle unerwünschten Ereignisse zusammengefasst, die mit der folgenden MedDRA (Version 13.1) <i>System Organ Class</i> kodiert wurden:
	Infections and Infestations.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-100: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird

das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-101: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse		Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert	
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab	
Population c)-Kombi: nur ein csDMARD in Vortherapie						
A3921064	83	78	35 (42,2)	25 (32,1)	1,3 [0,9-2,0], p=0,189 1,5 [0,8-2,9], p=0,186 10,1% [-4,7%-25,0%], p=0,181	
Population d)	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2	-	
A3921064	103	104	37 (35,9)	40 (38,5)	0,9 [0,7-1,3], p=0,706 0,9 [0,5-1,6], p=0,706 -2,5% [-15,7%-10,6%], p=0,705	
Quelle: (41)						
Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio						

#### Population c)-Kombi:

Die Rate der Patienten mit Infektionen oder parasitären Erkrankungen ist höher unter Tofacitinib + MTX (42,2% vs. 32,1%). Der Unterschied ist nicht statistisch signifikant.

## Population d)-Kombi:

Die Rate der Patienten mit Infektionen oder parasitären Erkrankungen ist höher unter Adalimumab + MTX (38,5% vs. 35,9%). Der Unterschied ist nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen

quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.22 Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster bis 12 Monate – RCT

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

Tabelle 4-102: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Herpes Zoster" werden alle unerwünschten Ereignisse zusammengefasst, die mit der folgenden MedDRA (Version 13.1) <i>System Organ Class</i> kodiert wurden:
	Infections and Infestations,
	und die zudem als Herpes Zoster adjudiziert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-103: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird

das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-104: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse		Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab
Population c)-Kombi: nur ein csDMARD in Vortherapie					
A3921064	83	78	3 (3,6)	0 (0,0)	6,6 [0,3-125,4], p=0,210 6,8 [0,3-134,3], p=0,206 3,6% [-1,0%-8,2%], p=0,125
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	e	
A3921064	103	104	3 (2,9)	4 (3,8)	0,8 [0,2-3,3], p=0,711 0,8 [0,2-3,4], p=0,711 -0,9% [-5,9%-4,0%], p=0,710
Quelle: (41) Abkürzunge	n: N = Anzahl	, KI = Konfiden	zintervall, RR =	Relatives Risiko,	OR = Odds Ratio

## Population c)-Kombi:

Es wurden nur wenige Herpes Zoster Fälle berichtet: drei Ereignisse im Tofacitinib-Arm (3,6%), keines im Adalimumab-Arm. Es liegt kein statistisch signifikanter Unterschied vor.

## Population d)-Kombi:

Es wurden nur wenige Herpes Zoster Fälle berichtet: drei Ereignisse im Tofacitinib-Arm (2,9%), vier im Adalimumab-Arm (3,8%). Es liegt kein statistisch signifikanter Unterschied vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.23 Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose bis 12 Monate – RCT

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

Tabelle 4-105: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Tuberkulose" werden alle unerwünschten Ereignisse zusammengefasst, die mit der folgenden MedDRA (Version 13.1) <i>System Organ Class</i> kodiert wurden:
	Infections and Infestations,
	und die zudem als Tuberkulose adjudiziert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-106: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird

das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-107: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse		Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert		
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab		
Population c)-Kombi: nur ein csDMARD in Vortherapie							
A3921064	83	78	0 (0,0)	0 (0,0)	nicht definiert nicht definiert 0,0% [-2,4%-2,4%], p=1,000		
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2			
A3921064	103	104	0 (0,0)	0 (0,0)	nicht definiert nicht definiert 0,0% [-1,9%-1,9%], p=1,000		
Quelle: (41)							
Abkürzunge	Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio						

## Population c)-Kombi:

Es wurden keine Tuberkulose-Ereignisse berichtet.

## Population d)-Kombi:

Es wurden keine Tuberkulose-Ereignisse berichtet.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.24 Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs bis 12 Monate – RCT

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

Tabelle 4-108: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Nicht-Melanom Hautkrebs" werden alle unerwünschten Ereignisse zusammengefasst, die mit der MedDRA (Version 13.1) <i>System Organ Class</i> :
	Neoplasms Benign, Malignant and Unspecified (incl. Cysts and Polyps),
	und die mit dem High Level Group Term:
	Skin Neoplasms Malignant and Unspecified,
	und die mit einem der folgenden Preferred Terms:
	Non Melanoma Skin Cancer-Other oder Basal Cell Carcinoma oder Squamous Cell Carcinoma kodiert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-109: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-110: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse		Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert				
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab				
Population c)	-Kombi: nur e	in csDMARD in	Vortherapie						
A3921064	83	78	0 (0,0)	1 (1,3)	0,3 [0,0-7,6], p=0,475 3,1 [0,0-7,7], p=0,475 -1,3% [-4,7%-2,1%], p=0,464				
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapio	e					
A3921064	103	104	0 (0,0)	1 (1,0)	0,3 [0,0-8,2], p=0,503 0,3 [0,0-8,3], p=0,503 -1,0% [-3,6%-1,7%], p=0,474				
Quelle: (41) Abkürzunge	7 2 7 171								

## Population c)-Kombi:

Es wurde lediglich ein Nicht-Melanom Hautkrebs berichtet, welcher im Adalimumab-Arm aufgetreten war. Es liegt kein statistisch signifikanter Unterschied vor.

#### Population d)-Kombi:

Es wurde lediglich ein Nicht-Melanom Hautkrebs berichtet, welcher im Adalimumab-Arm aufgetreten war. Es liegt kein statistisch signifikanter Unterschied vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.25 Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation bis 12 Monate – RCT

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

Tabelle 4-111: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Gastrointestinale Perforation" werden alle unerwünschten Ereignisse zusammengefasst, die mit der folgenden MedDRA (Version 13.1) <i>System Organ Class</i> kodiert wurden:
	Gastrointestinal Disorders,
	und die zudem als Gastrointestinales Ereignisse adjudiziert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-112: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird

das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-113: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie		Anzahl der Patienten in der Analyse		atienten mit is, N (%)	RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert		
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab		
Population c)	-Kombi: nur e	in csDMARD in	Vortherapie				
A3921064	83	78	0 (0,0)	0 (0,0)	nicht definiert nicht definiert 0,0% [-2,4%-2,4%], p=1,000		
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	e			
A3921064	103	104	0 (0,0)	0 (0,0)	nicht definiert nicht definiert 0,0% [-1,9%-1,9%], p=1,000		
Quelle: (41)							
Abkürzunge	n: N = Anzahl	, KI = Konfiden	zintervall, RR =	Relatives Risiko,	OR = Odds Ratio		

## Population c)-Kombi:

Es wurden keine Gastrointestinalen Perforationen berichtet.

## Population d)-Kombi:

Es wurden keine Gastrointestinalen Perforationen bereichtet.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

# 4.3.1.3.26 Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs) bis 12 Monate – RCT

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

Tabelle 4-114: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs) bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Neoplasien (ohne Nicht-Melanom Hautkrebs)" werden alle als unerwünschten Ereignisse gemeldeten Neoplasien zusammengefasst, die NICHT mit einem der folgenden <i>Preferred Terms</i> nach MedDRA (Version 13.1) kodiert wurden:
	Non Melanoma Skin Cancer-Other oder Basal Cell Carcinoma oder Squamous Cell Carcinoma.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-115: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs) bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird

das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-116: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs) bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse		Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert			
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab			
Population c)	-Kombi: nur ei	in csDMARD in	Vortherapie					
A3921064	83	78	1 (1,2)	0 (0,0)	2,8 [0,1-68,2], p=0,523 2,9 [0,1-71,1], p=0,523 1,2% [-2,1%-4,5%], p=0,478			
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2				
A3921064	103	104	1 (1,0)	2 (1,9)	0,5 [0,1-5,5], p=0,574 0,5 [0,04-5,6], p=0,574 -1,0% [-4,2%-2,3%], p=0,566			
Quelle: (41)								
Abkürzunge	Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio							

## Population c)-Kombi:

Es wurde lediglich eine Malignität (ohne Nicht-Melanom Hautkrebs) berichtet, welche im Tofacitinib-Arm aufgetreten war. Es liegt kein statistisch signifikanter Unterschied vor.

## Population d)-Kombi:

Es wurden nur wenige Malignitäten (ohne Nicht-Melanom Hautkrebs) berichtet: ein Ereignis im Tofacitinib-Arm (1,0%), zwei im Adalimumab-Arm (1,9%). Es liegt kein statistisch signifikanter Unterschied vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

## 4.3.1.3.27 Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte bis 12 Monate – RCT

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

Tabelle 4-117: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Erhöhte Leberwerte" werden alle unerwünschten Ereignisse zusammengefasst, die mit der MedDRA (Version 13.1) <i>System Organ Class</i> :
	Investigations,
	und die mit dem High Level Group Term:
	Hepatobiliary Investigations,
	und die mit dem High Level Term:
	Liver Function Analyses,
	und die mit einem der folgenden Preferred Terms:
	Alanine Aminotransferase Increased oder Aspartate Aminotransferase Increased oder Blood Alkaline Phosphatase Increased oder Liver Function Test Abnormal kodiert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-118: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-119: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse		Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert		
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab		
Population c)	Population c)-Kombi: nur ein csDMARD in Vortherapie						
A3921064	83	78	4 (4,8)	1 (1,3)	3,8 [0,4-32,9], p=0,232 3,9 [0,4-35,7], p=0,228 3,5% [-1,7%-8,8%], p=0,186		
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2			
A3921064	103	104	1 (1,0)	2 (1,9)	0,5 [0,1-5,5], p=0,574 0,5 [0,04-5,6], p=0,574 -1,0% [-4,2%-2,3%], p=0,566		
Quelle: (41)							
Abkürzunge	n: N = Anzahl,	, KI = Konfiden	zintervall, RR =	Relatives Risiko,	OR = Odds Ratio		

## Population c)-Kombi:

Die Rate der Patienten mit erhöhten Leberwerten ist höher unter Tofacitinib + MTX (4,8% vs. 1,3%). Der Unterschied ist nicht statistisch signifikant.

#### Population d)-Kombi:

Die Rate der Patienten mit erhöhten Leberwerten ist vergleichbar in beiden Studienarmen (1,0% bzw. 1,9%). Der Unterschied ist nicht statistisch signifikant.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

# 4.3.1.3.28 Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut bis 12 Monate – RCT

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

Tabelle 4-120: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Erhöhte Kreatininwerte im Blut" werden alle unerwünschten Ereignisse zusammengefasst, die mit der MedDRA (Version 13.1) System Organ Class:
	Investigations,
	und die mit dem High Level Group Term:
	Renal and Urinary Tract Investigations and Urinanalyses,
	und die mit dem High Level Term:
	Renal Function Analyses,
	und dem folgenden Preferred Term:
	Blood Creatinine Increased
	kodiert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-121: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-122: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl der Patienten in der Analyse			atienten mit is, N (%)	RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert			
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab			
Population c)-Kombi: nur ein csDMARD in Vortherapie								
A3921064	83	78	0 (0,0)	0 (0,0)	nicht definiert nicht definiert 0,0% [-2,4%-2,4%], p=1,000			
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2				
A3921064	103	104	1 (1,0)	1 (1,0)	1,0 [0,1-15,9], p=0,995 1,0 [0,1-16,4], p=0,995 0,0% [-2,7%-2,7%], p=0,995			
Quelle: (41)								
Abkürzunge	Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio							

#### Population c)-Kombi:

Es wurden keine Fälle von erhöhten Kreatininwerten berichtet.

### Population d)-Kombi:

Es wurden in beiden Studienarmen jeweils nur ein Fall von erhöhten Kreatininwerten berichtet (jeweils 1,0%). Es liegt kein statistisch signifikanter Unterschied vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

# 4.3.1.3.29 Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis 12 Monate – RCT

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

Tabelle 4-123: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Neutropenie" werden alle unerwünschten Ereignisse zusammengefasst, die mit der MedDRA (Version 13.1) <i>System Organ Class</i> :
	Investigations,
	und die mit dem High Level Group Term:
	Haematology Investigations (incl. Blood Groups),
	und die mit dem High Level Term:
	White Blood Cell Analyses,
	und die mit einem der folgenden Preferred Terms:
	Neutrophil Count Decreased oder Granulocyte Count Decreased kodiert wurden,
	bzw.
	die mit der System Organ Class:
	Blood and Lymphatic System Disorders,
	und die mit dem High Level Group Term:
	White Blood Cell Disorders,
	und die mit dem High Level Term:
	Neutropenias kodiert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-124: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-125: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Neutropenie bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie		Anzahl der Patienten in der Analyse		atienten mit is, N (%)	RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert				
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab				
Population c)-Kombi: nur ein csDMARD in Vortherapie									
A3921064	83	78	0 (0,0)	0 (0,0)	nicht definiert nicht definiert 0,0% [-2,4%-2,4%], p=1,000				
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2					
A3921064	103	104	1 (1,0)	0 (0,0)	3,0 [0,1-73,5], p=0,496 3,1 [0,1-76,0], p=0,495 1,0% [-1,7%-3,6%], p=0,472				
Quelle: (41)									
Abkürzunge	Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio								

#### Population c)-Kombi:

Es wurden keine Neutropenien berichtet.

#### <u>Population d)-Kombi:</u>

Es wurden lediglich ein Fall von Neutropenie im Tofacitinib-Arm berichtet (1,0%). Es liegt kein statistisch signifikanter Unterschied vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

# 4.3.1.3.30 Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie bis 12 Monate – RCT

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

Tabelle 4-126: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Lymphopenie" werden alle unerwünschten Ereignisse zusammengefasst, die mit der MedDRA (Version 13.1) <i>System Organ Class</i> :
	Blood and Lymphatic System Disorders kodiert wurden,
	bzw.
	die mit der System Organ Class:
	Investigations,
	und die mit einem der folgenden Preferred Terms:
	B-Lymphocyte Count Decreased oder Lymphocyte Count Decreased oder Lymphopenia oder T-Lymphocyte Count Decreased oder Lymphocytopenia Neonatal kodiert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-127: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-128: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie		Patienten in analyse	Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert			
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab			
Population c)-Kombi: nur ein csDMARD in Vortherapie								
A3921064	83	78	0 (0,0)	0 (0,0)	nicht definiert nicht definiert 0,0% [-2,4%-2,4%], p=1,000			
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapie	2				
A3921064	103	104	0 (0,0)	0 (0,0)	nicht definiert nicht definiert 0,0% [-1,9%-1,9%], p=1,000			
Quelle: (41)								
Abkürzunge	Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio							

### Population c)-Kombi:

Es wurden keine Lymphopenien berichtet.

#### Population d)-Kombi:

Es wurden keine Lymphopenien berichtet.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

# 4.3.1.3.31 Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12 Monate – RCT

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

Tabelle 4-129: Operationalisierung von "Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12 Monate"

Studie	Operationalisierung
A3921064	Unter "Anämie" werden alle unerwünschten Ereignisse zusammengefasst, die mit der MedDRA (Version 13.1) <i>System Organ Class</i> :
	Blood and Lymphatic System Disorders kodiert wurden,
	und die mit der standardisierten Abfrage für Anämie (MedDRA standardized query Haematopoietic Erythropenia [20000029]) kodiert wurden.
	Dargestellt werden die Anzahl der Ereignisse, das relative Risiko, das Odds Ratio sowie die Risikodifferenz für den auf diese Absolutzahlen bezogenen Effektschätzer.
	Fehlende Werte werden nicht ersetzt.
	Die Analyse erfolgte für die Safety Analysis Population.

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

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

Tabelle 4-130: Bewertung des Verzerrungspotenzials für "Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12 Monate" in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
A3921064	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Auf Studienebene wurde das Verzerrungpotenzial als niedrig eingestuft. Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Es werden alle Patienten der *Safety Population* 

berücksichtigt, d.h. alle randomisierten Patienten, die mind. eine Dosis der Studienmedikation erhalten haben. Die Auswertung des Endpunkts folgt den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch keine sonstigen Aspekte, die das Verzerrungspotenzial bezüglich dieses Endpunkts beeinflussen könnten.

Das Verzerrungspotenzial auf Endpunktebene wird folglich als niedrig eingestuft.

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

Tabelle 4-131: Ergebnisse für "Unerwünschte Ereignisse von besonderem Interesse: Anämie bis 12 Monate" aus RCT mit dem zu bewertenden Arzneimittel

Studie		Patienten in Analyse	Anzahl Patienten mit Ereignis, N (%)		RR [95%-KI], p-Wert OR ]95%-KI], p-Wert Risikodifferenz [95%-KI], p-Wert			
	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib + MTX	Adalimumab + MTX	Tofacitinib vs. Adalimumab			
Population c)-Kombi: nur ein csDMARD in Vortherapie								
A3921064	83	78	1 (1,2)	0 (0,0)	2,8 [1,2-68,2], p=0,523 2,9 [0,1-71,1], p=0,523 1,2% [-2,1%-4,5%], p=0,478			
Population d	-Kombi: mehr	als ein csDMAR	D in Vortherapio	e				
A3921064	103	104	1 (1,0)	0 (0,0)	3,0 [0,1-73,5], p=0,496 3,1 [0,1-76,0], p=0,495 1,0% [-1,7%-3,6%], p=0,472			
Quelle: (41)								
Abkürzunge	Abkürzungen: N = Anzahl, KI = Konfidenzintervall, RR = Relatives Risiko, OR = Odds Ratio							

#### Population c)-Kombi:

Es wurden lediglich ein Fall von Neutropenie im Tofacitinib-Arm berichtet (1,2%). Es liegt kein statistisch signifikanter Unterschied vor.

#### Population d)-Kombi:

Es wurden keine Fälle von Anämie berichtet.

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.

Es wurde keine Meta-Analyse durchgeführt, da nur eine Studie eingeschlossen wurde.

#### 4.3.1.3.32 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.

Im Anschluss werden die Ergebnisse der Subgruppenanalysen dargestellt. Folgende Subgruppen wurden für alle patientenrelevanten Endpunkte analysiert:

- Alter ( $\leq 65 / > 65$  Jahre)
- Geschlecht (männlich / weiblich)
- Region (Europa / USA und Kanada / Lateinamerika / Sonstige)
- Schweregrad nach DAS28-4(ESR) ( $\leq 5,1/>5,1$ )

In diesem Abschnitt werden pro Endpunkt jeweils die Ergebnisse dargestellt, falls der zugehörige Interaktionstest einen Hinweis  $(0.05 \le p < 0.2)$  oder einen Beleg (p < 0.05) auf Effektmodifikation zeigt. Die gesamten Subgruppenanalysen sind in den Zusatzanalysen (40, 41) hinterlegt und können dort im Detail eingesehen werden.

Vorausgeschickt werden soll jedoch, dass die Kontrolle der Irrtumswahrscheinlichkeit 1. Art angesichts der großen Zahl an auf dem nominellen Niveau von 5% (Belege für Effektmodifikation) bzw. 20% (Belege und Hinweise für Effektmodifikation) durchgeführten Interaktionstests nicht mehr gegeben ist. Die Wahrscheinlichkeit für falsch positive Ergebnisse ist dadurch stark erhöht, weshalb zusätzliche Kriterien (biologische Plausibilität, Richtung von Effektmodifikationen, Bekanntheit bestimmer Effekte) heranzuziehen sind, um falsch positive Testergebnisse nicht in inadäquater Weise zu interpretieren.

Insgesamt wurden für 53 dafür relevante Endpunkte jeweils 4 Subgruppenfaktoren getestet, d.h. insgesamt wurden 212 Interaktionstests für jede der Teilpopulationen c) bzw. d) durchgeführt. Damit ergibt sich die erwartete Anzahl von rund 32 falsch positiven Hinweisen auf Effektmodifikation (0,05  $\geq$  p-Wert < 0,20 trotz Gültigkeit der Nullhypothese, d.h. Nicht-Bestehen einer Effektmodifikation) und von rund 11 falsch positiven Belegen auf Effektmodifikation (p-Wert < 0,05; 212 \* 0,15 = 31,8 bzw. 212 \* 0,05 = 10,6). Diese Verhältnisse gelten wie erwähnt für jede der separat zu bewertenden Subpopulation der Gesamtstudie ORAL STANDARD und sind für die Interpretation der "Subgruppeneffekte" zu beachten.

# 4.3.1.3.32.1 Subgruppenanalysen für Endpunkt DAS28-4(ESR) nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-132: p-Werte der Interaktionstests für DAS28-4(ESR) nach 12 Monaten

Subgruppenfaktor	Interaktions-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie			
	$DAS28-4(ESR) < 2,6$ $DAS28-4(ESR) \le 3,2$		
Geschlecht	NA	p=0,306	
Alter	p=0,318	p=0,822	
Region	NA	NA	
DAS28-4(ESR)	p=0,930	p=0,578	
Population d)-Kombi: mehr als ein	csDMARD in Vortherapie		
	DAS28-4(ESR) < 2,6	$DAS28-4(ESR) \le 3,2$	
Geschlecht	p=0,488	p=0,445	
Alter	NA	p=0,951	
Region	NA	p=0,211	
DAS28-4(ESR)	p=0,029	p=0,300	

Quelle: (40)

#### **Subgruppen:** Geschlecht

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegt ein Beleg für Effektmodifikation für den Schwellenwert DAS28-4(ESR) < 2,6 in der Population d)-Kombi vor (p=0,029).

Tabelle 4-133: Subgruppenergebnisse für DAS28-4(ESR) < 2,6 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mehr als ei	n csDMARD in V	Vortherapie		
Schweregrad: DAS28-4(ESR)	3/4	5/10	1,5 [0,7-3,5]	
≤ 5,1	(75,0)	(50,0)	p=0,344	Dalas
Schweregrad: DAS28-4(ESR)	6/87	15/83	0,4 [0,2-0,9]	- Beleg
> 5,1	(6,9)	(18,1)	p=0,035	
Quelle: (40) Abkürzungen: RR = Relative Methotrexat	es Risiko, 95º	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In der sehr kleinen Gruppe mit niedrigerem Schweregrad ist die Responderrate numerisch höher im Tofacitinib-Arm. Dieser Unterschied ist jedoch nicht statistisch signifikant und wegen der Gruppengröße kaum interpretierbar. Für die Patienten mit höherem Schweregrad liegt die Responderrate statistisch signifikant höher im Adalimumab-Arm (p=0,035).

Eine sinnvolle Interpretation dieses Endpunkts ist jedoch hier auf Subgruppenebene wie auf Ebene der Gesamtpopulation nicht möglich aufgrund der niedrigen Patientenzahl des FAS, welche ein hohes Verzerrungspotenzial zur Folge hat (siehe 4.3.1.3.1).

# 4.3.1.3.32.2 Subgruppenanalysen für Endpunkt EULAR-Response (ESR) nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-134: p-Werte der Interaktionstests für EULAR-Response (ESR) nach 12 Monaten

Subgruppenfaktor	Interaktions-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie			
	gute Verbesserung	gute oder moderate Verbesserung	
Geschlecht	p=0,306	p=0,366	
Alter	p=0,822	p=0,884	
Region	NA	p=0,582	
DAS28-4(ESR)	p=0,578	p=0,325	
Population d)-Kombi: mehr als ein	csDMARD in Vortherapie		
	gute Verbesserung	gute oder moderate Verbesserung	
Geschlecht	p=0,445	p=0,216	
Alter	p=0,951	p=0,687	
Region	p=0,211	p=0,249	
DAS28-4(ESR)	p=0,300	p=0,899	

Quelle: (40)

#### **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

### Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Schweregrad nach DAS28-4(ESR)

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

Eine sinnvolle Interpretation dieses Endpunkts ist jedoch hier auf Subgruppenebene wie auf Ebene der Gesamtpopulation nicht möglich aufgrund der niedrigen Patientenzahl des FAS, welche ein hohes Verzerrungspotenzial zur Folge hat (siehe 4.3.1.3.3).

# 4.3.1.3.32.3 Subgruppenanalysen für Endpunkt ACR/EULAR-Remission und ACR Response nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diese Endpunkte.

Tabelle 4-135: p-Werte der Interaktionstests für ACR/EULAR-Remission und ACR Response nach 12 Monaten

Subgruppenfaktor		Interaktions-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie				
	ACR70 Response	ACR50 Response	ACR/EULAR- Remission	
Geschlecht	NA	p=0,362	p=0,913	
Alter	NA	p=0,907	p=0,431	
Region	NA	p=0,793	NA	
DAS28-4(ESR)	NA	p=0,890	p=0,837	
Population d)-Kombi: meh	r als ein csDMARD in Vortherapi	e		
	ACR70 Response	ACR50 Response	ACR/EULAR- Remission	
Geschlecht	p=0,776	p=0,642	p=0,505	
Alter	p=0,974	p=0,978	NA	
Region	p=0,162	p=0,575	NA	
DAS28-4(ESR)	p=0,744	p=0,623	p=0,564	

Quelle: (40)

#### **Subgruppen:** Geschlecht

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegt ein Hinweis auf Effektmodifikation für ACR70 in der Population d)-Kombi vor (p=0,162).

Tabelle 4-136: Subgruppenergebnisse für ACR70 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mehr a	als ein csDMARD in V	Vortherapie		
Europa	11/63 (17,5)	14/60 (23,3)	0,8 [0,4-1,5] p=0,421	
USA/Kanada	5/9 (55,6)	1/11 (9,1)	6,1 [0,9-43,3] p=0,070	- -
Lateinamerika	2/6 (33,3)	2/11 (18,2)	1,8 [0,3-9,9] p=0,482	– Hinweis
Sonstige	6/22 (27,3)	3/21 (14,3)	1,9 [0,6-6,7] p=0,311	_
Quelle: (40) Abkürzungen: RR = Rel Methotrexat	latives Risiko, 95	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In Europa ist eine höhere Responderrate im Adalimumab-Arm zu sehen, in USA/Kanada, Lateinamerika und den sonstigen Regionen im Tofacitinib-Arm. Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

#### **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### 4.3.1.3.32.4 Subgruppenanalysen für Endpunkt SDAI nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-137: p-Werte der Interaktionstests für SDAI nach 12 Monaten

Subgruppenfaktor	Interaktions-p-Wert			
Population c)-Kombi: nur ein csDN	Population c)-Kombi: nur ein csDMARD in Vortherapie			
	<b>SDAI ≤ 3,3</b>	SDAI ≤ 11		
Geschlecht	p=0,571	p=0,916		
Alter	p=0,367	p=0,953		
Region	NA	p=0,624		
DAS28-4(ESR)	NA	p=0,891		
Population d)-Kombi: mehr als ein	csDMARD in Vortherapie			
	<b>SDAI ≤ 3,3</b>	SDAI ≤ 11		
Geschlecht	p=0,295	p=0,284		
Alter	p=0,530	p=0,387		
Region	NA	p=0,379		
DAS28-4(ESR)	p=0,711	p=0,232		

Quelle: (40)

### **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Schweregrad nach DAS28-4(ESR)

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### 4.3.1.3.32.5 Subgruppenanalysen für Endpunkt CDAI nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-138: p-Werte der Interaktionstests für CDAI nach 12 Monaten

Subgruppenfaktor	Interaktions-p-Wert			
Population c)-Kombi: nur ein csDN	Population c)-Kombi: nur ein csDMARD in Vortherapie			
	<b>CDAI ≤ 2,8</b>	CDAI ≤ 10		
Geschlecht	NA	p=0,848		
Alter	p=0,614	p=0,767		
Region	NA	p=0,515		
DAS28-4(ESR)	NA	p=0,989		
Population d)-Kombi: mehr als ein	csDMARD in Vortherapie			
	<b>CDAI ≤ 2,8</b>	CDAI ≤ 10		
Geschlecht	p=0,298	p=0,244		
Alter	p=0,650	p=0,477		
Region	NA	p=0,116		
DAS28-4(ESR)	p=0,561	p=0,476		

Quelle: (40)

### **Subgruppen:** Geschlecht

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegt ein Hinweis auf Effektmodifikation für den Schwellenwert CDAI  $\leq$  10 in der Population d)-Kombi vor (p=0,116).

Tabelle 4-139: Subgruppenergebnisse für CDAI ≤ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: me	hr als ein csDMARD in \	Vortherapie		
Europa	24/63 (38,1)	27/60 (45,0)	0,9 [0,6-1,3] p=0,438	
USA/Kanada	6/9 (66,7)	3/11 (27,3)	2,4 [0,8-7,1] p=0,102	-
Lateinamerika	3/6 (50,0)	3/11 (27,3)	1,8 [0,5-6,4] p=0,343	- Hinweis
Sonstige	9/22 (40,9)	4/21 (19,0)	2,2 [0,8-5,9] p=0,140	_
Quelle: (40) Abkürzungen: RR = Methotrexat	Relatives Risiko, 95	%-KI = 95%-Ko	onfidenzintervall, MTX =	

In Europa ist eine höher Responderrate im Adalimumab-Arm zu sehen, in USA/Kanada, Lateinamerika und den sonstigen Regionen im Tofacitinib-Arm. Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

### **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.6 Subgruppenanalysen für Endpunkt Schmerzbeurteilung durch den Patienten nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-140: p-Werte der Interaktionstests für Schmerzbeurteilung durch den Patienten ≥ 10 nach 12 Monaten

Subgruppenfaktor	Interaktions-p-Wert	
Population c)-Kombi: nur ein csDMARD in Vortherapie		
Geschlecht	p=0,383	
Alter	p=0,732	
Region	p=0,564	
DAS28-4(ESR)	p=0,322	
Population d)-Kombi: mehr als ein csDMARD in Vortherapie		
Geschlecht	p=0,252	
Alter	p=0,183	
Region	p=0,147	
DAS28-4(ESR)	NA	

Quelle: (40)

# Subgruppen: Geschlecht

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegt ein Hinweis auf Effektmodifikation in der Population d)-Kombi vor (p=0,183).

Tabelle 4-141: Subgruppenergebnisse für Schmerzbeurteilung durch den Patienten ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mehr als e	ein csDMARD in V	Vortherapie		
Alter ≤ 65 Jahre	56/87	57/92	1,0 [0,8-1,3]	
	(64,4)	(62,0)	p=0,738	- Hinweis
Alter > 65 Jahre	6/13	8/11	0,6 [0,3-1,3]	Hillweis
	(46,2)	(72,7)	p=0,196	
Quelle: (40) Abkürzungen: RR = Relativ Methotrexat	ves Risiko, 95º	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In der Gruppe der jüngeren Patienten sind die Responderraten vergleichbar in beiden Studienarmen. Bei den älteren Patienten ist die Rate im Adalimumab-Arm höher als im Tofacitinib-Arm. Es liegen keine statistisch signifikanten Unterschiede vor.

#### **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegt ein Hinweis auf Effektmodifikation in der Population d)-Kombi vor (p=0,147).

Tabelle 4-142: Subgruppenergebnisse für Schmerzbeurteilung durch den Patienten ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: me	hr als ein csDMARD in '	Vortherapie		
Europa	37/63 (58,7)	44/60 (73,3)	0,8 [0,6-1,04] p=0,091	
USA/Kanada	6/9 (66,7)	5/11 (45,5)	1,5 [0,7-3,3] p=0,345	_
Lateinamerika	4/6 (66,7)	5/11 (45,5)	1,5 [0,6-3,5] p=0,383	- Hinweis
Sonstige	15/22 (68,2)	11/21 (52,4)	1,3 [0,8-2,1] p=0,299	_
Quelle: (40) Abkürzungen: RR = Methotrexat	Relatives Risiko, 95	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In Europa ist eine höher Responderrate im Adalimumab-Arm zu sehen, in USA/Kanada, Lateinamerika und den sonstigen Regionen im Tofacitinib-Arm. Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

#### **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.7 Subgruppenanalysen für Endpunkt Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA) nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-143: p-Werte der Interaktionstests für PatGA ≥ 10 nach 12 Monaten

Subgruppenfaktor	Interaktions-p-Wert	
Population c)-Kombi: nur ein csDMARD in Vortherapie		
Geschlecht	p=0,691	
Alter	p=0,281	
Region	p=0,677	
DAS28-4(ESR)	p=0,287	
Population d)-Kombi: mehr als ein csDMARD in Vortherapie		
Geschlecht	p=0,132	
Alter	p=0,710	
Region	p=0,149	
DAS28-4(ESR)	p=0,910	

Quelle: (40)

#### **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegt ein Hinweis auf Effektmodifikation in der Population d)-Kombi vor (p=0,132).

Tabelle 4-144: Subgruppenergebnisse für PatGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation	
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie					
männlich	11/13 (84,6)	15/26 (57,7)	1,5 [0,98-2,2] p=0,062		
weiblich	53/87 (60,9)	46/77 (59,7)	1,0 [0,8-1,3] p=0,878	- Hinweis	
Quelle: (40) Abkürzungen: RR = Re Methotrexat	latives Risiko, 95º	%-KI = 95%-Ko	nfidenzintervall, MTX =		

Bei den Männern ist die Responderrate höher im Tofacitinib-Arm, bei den Frauen sind die Raten vergleichbar in beiden Studienarmen. Die Unterschiede sind jeweils nicht statistisch signifikant.

#### Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegt ein Hinweis auf Effektmodifikation in der Population d)-Kombi vor (p=0,149).

Tabelle 4-145: Subgruppenergebnisse für PatGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mehr	als ein csDMARD in	Vortherapie		
Europa	40/63 (63,5)	43/60 (71,7)	0,9 [0,7-1,1] p=0,334	
USA/Kanada	7/9 (77,8)	4/11 (36,4)	2,1 [0,9-5,0] p=0,082	- · · ·
Lateinamerika	3/6 (50,0)	4/11 (36,4)	1,4 [0,5-4,2] p=0,577	- Hinweis
Sonstige	14/22 (63,6)	10/21 (47,6)	1,3 [0,8-2,3] p=0,300	_
Quelle: (40) Abkürzungen: RR = Re Methotrexat	elatives Risiko, 95	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In Europa sind höhere Responderraten im Adalimumab-Arm zu sehen, in USA/Kanada, Lateinamerika und den sonstigen Regionen im Tofacitinib-Arm. Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

### **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.8 Subgruppenanalysen für Endpunkt Gesamtbeurteilung der Krankheitsaktivität durch den Arzt (PhysGA) nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-146: p-Werte der Interaktionstests für PhysGA ≥ 10 nach 12 Monaten

Subgruppenfaktor	Interaktions-p-Wert			
Population c)-Kombi: nur ein csDMARD in Vortherapie				
Geschlecht	p=0,479			
Alter	p=0,612			
Region	p=0,096			
DAS28-4(ESR)	p=0,745			
Population d)-Kombi: mehr als ein csDMARD in Vortherapie				
Geschlecht	p=0,040			
Alter	p=0,686			
Region	p=0,351			
DAS28-4(ESR)	p=0,916			

Quelle: (40)

#### **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegt ein Beleg für Effektmodifikation in der Population d)-Kombi vor (p=0,040).

Tabelle 4-147: Subgruppenergebnisse für PhysGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation	
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie					
männlich	11/13 (84,6)	16/26	1,4 [0,9-2,0] p=0,102		
weiblich	59/86 (68,6)	(61,5) 60/77 (77,9)	0,9 [0,7-1,1] p=0,180	- Beleg	
Quelle: (40) Abkürzungen: RR = Re Methotrexat	· · · · · · · · · · · · · · · · · · ·		•		

Bei den Männern ist die Responderrate höher im Tofacitinib-Arm, bei den Frauen im Adalimumab-Arm. Die Unterschiede sind jeweils nicht statistisch signifikant.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegt ein Hinweis auf Effektmodifikation in der Population c)-Kombi vor (p=0,096).

Tabelle 4-148: Subgruppenergebnisse für PhysGA ≥ 10 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population c)-Kombi: nur	ein csDMARD in Vorth	erapie		
Europa	36/41 (87,8)	32/41 (78,0)	1,1 [0,9-1,4] p=0,245	
USA/Kanada	11/18 (61,1)	6/12 (50,0)	1,2 [0,6-2,4] p=0,560	_
Lateinamerika	4/10 (40,0)	9/10 (90,0)	0,4 [0,2-0,98] p=0,043	- Hinweis
Sonstige	7/9 (77,8)	11/12 (91,7)	0,9 [0,6-1,3] p=0,407	_
Quelle: (40) Abkürzungen: RR = Methotrexat	Relatives Risiko, 95	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In Europa und in USA/Kanada sind höhere Responderraten im Tofacitinib-Arm zu sehen, in den sonstigen Regionen im Adalimumab-Arm. Diese Unterschiede sind jeweils nicht statistisch signifikant. Eine statistisch signifikant höhere Rate im Adalimumab-Arm ist in Lateinamerika zu beobachten (p=0,043).

#### **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.9 Subgruppenanalysen für Endpunkt Anzahl empfindlicher/druckschmerzhafter Gelenke nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-149: p-Werte der Interaktionstests für Anzahl empfindlicher/druckschmerzhafter Gelenke  $\leq 1$  nach 12 Monaten

Subgruppenfaktor	Interaktions-p-Wert			
Population c)-Kombi: nur ein csDMARD in Vortherapie				
	28 untersuchte Gelenke	68 untersuchte Gelenke		
Geschlecht	p=0,829	p=0,574		
Alter	p=0,751	p=0,719		
Region	p=0,318	p=0,309		
DAS28-4(ESR)	p=0,768	p=0,917		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie				
	28 untersuchte Gelenke	68 untersuchte Gelenke		
Geschlecht	p=0,655	p=0,809		
Alter	p=0,989	p=0,361		
Region	p=0,217	p=0,272		
DAS28-4(ESR)	p=0,704	p=0,933		

Quelle: (40)

#### **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Schweregrad nach DAS28-4(ESR)

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.10 Subgruppenanalysen für Endpunkt Anzahl geschwollener Gelenke nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-150: p-Werte der Interaktionstests für Anzahl geschwollener Gelenke ≤ 1 nach 12 Monaten

Subgruppenfaktor	Interakti	ons-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie				
	28 untersuchte Gelenke 66 untersuchte Gelen			
Geschlecht	p=0,823	p=0,982		
Alter	p=0,744	p=0,137		
Region	p=0,749	p=0,446		
DAS28-4(ESR)	p=0,803	p=0,712		
Population d)-Kombi: mehi	als ein csDMARD in Vortherapie			
	28 untersuchte Gelenke	66 untersuchte Gelenke		
Geschlecht	p=0,565	p=0,768		
Alter	p=0,748	p=0,338		
Region	p=0,100	p=0,260		
DAS28-4(ESR)	p=0,635	p=0,808		

Quelle: (40)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegt ein Hinweis auf Effektmodifikation für die Operatinalisierung der 66 untersuchten Gelenke in der Population c)-Kombi vor (p=0,137).

Tabelle 4-151: Subgruppenergebnisse für Anzahl geschwollener Gelenke (66) ≤ 1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population c)-Kombi: nur ein c	sDMARD in Vorth	erapie		
Alter ≤ 65 Jahre	24/60	26/63	1,0 [0,6-1,5]	
	(40,0)	(41,3)	p=0,886	- Hinweis
Alter > 65 Jahre	9/19	2/12	2,8 [0,7-11,0]	Hillweis
	(47,4)	(16,7)	p=0,130	
Quelle: (40) Abkürzungen: RR = Rela Methotrexat	tives Risiko, 95º	%-KI = 95%-Ko	nfidenzintervall, MTX =	

Bei den jüngeren Patienten sind die Responderraten vergleichbar in beiden Studienarmen, bei den Älteren ist die Rate höher unter Tofacitinib + MTX. Es liegen jedoch keine statistisch signifikanten Unterschiede vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegt ein Hinweis auf Effektmodifikation für die Operationalisierung der 28 untersuchten Gelenke in der Population d)-Kombi vor (p=0,100).

Tabelle 4-152: Subgruppenergebnisse für Anzahl geschwollener Gelenke (28)  $\leq$  1 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mel	nr als ein csDMARD in '	Vortherapie		
Europa	24/63 (38,1)	27/60 (45,0)	0,9 [0,6-1,3] p=0,438	
USA/Kanada	6/9 (66,7)	1/11 (9,1)	7,3 [1,1-50,3] p=0,043	
Lateinamerika	3/6 (50,0)	4/11 (36,4)	1,4 [0,5-4,2] p=0,577	- Hinweis
Sonstige	9/22 (40,9)	5/21 (23,8)	1,7 [0,7-4,3] p=0,246	_
Quelle: (40) Abkürzungen: RR = 1 Methotrexat	Relatives Risiko, 95	%-KI = 95%-Ko	onfidenzintervall, MTX =	

In Europa ist eine höhere Responderrate im Adalimumab-Arm zu sehen, in USA/Kanada, Lateinamerika und den sonstigen Regionen im Tofacitinib-Arm. Der Unterschied ist für USA/Kanada statistisch signifikant zugunsten von Tofacitinib + MTX (p=0,043), die anderen Unterschiede sind jeweils nicht statistisch signifikant.

#### **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.11 Subgruppenanalysen für Endpunkt SF-36 nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für die Komponentenscores des SF-36.

Tabelle 4-153: p-Werte der Interaktionstests für SF-36 Responderanalysen der Komponentenscores nach 12 Monaten (Verbesserung jeweils ≥ 2,5)

Subgruppenfaktor	Interaktions-p-Wert			
Population c)-Kombi: nur ein csDMARD in Vortherapie				
	Körperlicher Psychischer Komponentenscore Komponentenscore			
Geschlecht	p=0,930	p=0,495		
Alter	p=0,623	p=0,419		
Region	p=0,582	NA		
DAS28-4(ESR)	p=0,309	p=0,569		
Population d)-Kombi: mehr als ei	n csDMARD in Vortherapie			
	Körperlicher Komponentenscore	Psychischer Komponentenscore		
Geschlecht	p=0,724	p=0,515		
Alter	p=0,187	p=0,379		
Region	p=0,295	p=0,324		
DAS28-4(ESR)	p=0,585	p=0,575		

Quelle: (40)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegt ein Hinweis auf Effektmodifikation für den Körperlichen Komponentenscore in der Population d)-Kombi vor (p=0,187).

Tabelle 4-154: Subgruppenergebnisse für SF-36 Körperlicher Komponentenscore Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mehr	als ein csDMARD in V	Vortherapie		
Alter ≤ 65 Jahre	53/87	54/92	1,0 [0,8-1,3]	
	(60,9)	(58,7)	p=0,762	TT::-
Alter > 65 Jahre	6/13	8/11	0,6 [0,3-1,3]	- Hinweis
	(46,2)	(72,7)	p=0,196	
Quelle: (40)	5	0.507.47		
Abkürzungen: RR = R Methotrexat	elatīves Risiko, 95	%-KI = 95%-Ko	onfidenzintervall, MTX =	

Für die jüngeren Patienten sind die Responderraten vergleichbar in beiden Studienarmen. In der Gruppe der Älteren ist die Rate unter Adalimumab + MTX höher. Es liegt kein statistisch signifikanter Unterschied vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Schweregrad nach DAS28-4(ESR)

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für die Einzeldimensionen des SF-36.

Stand: 21.04.2017

Tabelle 4-155: p-Werte der Interaktionstests für SF-36 Responderanalysen der Einzeldimensionen nach 12 Monaten (Verbesserung jeweils ≥ 5)

Subgruppenfaktor		Interaktions-p-Wert			
Population c)-Kombi: nur ein cs	DMARD in Vortherapie				
	Vitalität	Körperliche Funktion	Körperliche Schmerzen	Allgemeine Gesundheit	
Geschlecht	p=0,706	p=0,832	p=0,620	p=0,795	
Alter	p=0,669	p=0,281	p=0,235	p=0,735	
Region	p=0,342	p=0,292	p=0,401	p=0,448	
DAS28-4(ESR)	p=0,673	p=0,626	NA	p=0,342	
Population d)-Kombi: mehr als	ein csDMARD in Vortherapie				
	Vitalität	Körperliche Funktion	Körperliche Schmerzen	Allgemeine Gesundheit	
Geschlecht	p=0,768	p=0,615	p=0,556	p=0,943	
Alter	p=1,000	p=0,096	p=0,294	p=0,513	
Region	NA	p=0,726	p=0,891	p=0,006	
DAS28-4(ESR)	p=0,470	p=0,551	p=0,600	p=0,431	

Tofacitinib (XELJANZ®)

Seite 294 von 442

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

Subgruppenfaktor		Interaktions-p-Wert				
Population c)-Kombi: nur ein csDMARD in Vortherapie						
	Körperliche Rollenfunktion	Emotionale Rollenfunktion	Psychisches Wohlbefinden	Soziale Funktion		
Geschlecht	p=0,719	p=0,844	p=0,196	p=0,803		
Alter	p=0,213	p=0,741	p=0,332	p=0,438		
Region	p=0,607	NA	NA	p=0,247		
DAS28-4(ESR)	p=0,736	p=0,695	p=0,892	p=0,112		
Population d)-Kombi: meh	r als ein csDMARD in Vortherap	oie		•		
	Körperliche Rollenfunktion	Emotionale Rollenfunktion	Psychisches Wohlbefinden	Soziale Funktion		
Geschlecht	p=0,625	p=0,440	p=0,682	p=0,146		
Alter	p=0,243	p=0,103	p=0,399	p=0,063		
Region	p=0,408	p=0,397	p=0,195	p=0,417		
DAS28-4(ESR)	p=0,413	p=0,250	p=0,587	p=0,427		

Stand: 21.04.2017

Quelle: (40)

Tofacitinib (XELJANZ®) Seite 295 von 442

# **Subgruppen:** Geschlecht

Bezüglich des Subgruppenmerkmals "Geschlecht" liegt ein Hinweis auf Effektmodifikation für "Psychisches Wohlbefinden" in der Population c)-Kombi vor (p=0,196). In der Population d)-Kombi liegt für "Soziale Funktion" ein Hinweis auf Effektmodifikation vor (p=0,146).

Tabelle 4-156: Subgruppenergebnisse für SF-36 Einzeldimensionen Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht

Studie A3921064		Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
		Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population c)-Ko	mbi: nur ein csDMA	RD in Vortherapie			
Psychisches	männlich	3/13 (23,1)	6/12 (50,0)	0,5 [0,2-1,5] p=0,185	TI'.
Wohlbefinden	weiblich	26/66 (39,4)	24/63 (38,1)	1,0 [0,7-1,6] p=0,880	- Hinweis
Population d)-Ko	mbi: mehr als ein csI	OMARD in Vortherapie	<b>;</b>		
Soziale	männlich	7/13 (53,8)	9/26 (34,6)	1,6 [0,8-3,2] p=0,235	Himmin
Funktion	weiblich	40/87 (46,0)	41/77 (53,2)	0,9 [0,6-1,2] p=0,352	- Hinweis
Quelle: (40) Abkürzungen: R	R = Relatives Risiko	o, 95%-KI = 95%-Kon	fidenzintervall, M	TX = Methotrexat	

# Population c)-Kombi:

Bei den Männern ist die Responderrate für "Psychisches Wohlbefinden" höher im Adalimumab-Arm, bei den Frauen sind die Raten vergleichbar. Es liegen keine statistisch signifikanten Unterschiede vor.

#### Population d)-Kombi:

Bei den Männern ist die Responderrate für "Soziale Funktion" höher im Tofacitinib-Arm, bei den Frauen ist die Rate im Adalimumab-Arm höher. Es liegen keine statistisch signifkanten Unterschiede vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegt in der Population d)-Kombi für "Körperliche Funktion", "Emotionale Rollenfunktion" und "Soziale Funktion" jeweils ein Hinweis auf Effektmodifikation vor (p=0,096 bzw. p=0,103 bzw. p=0,063).

Tabelle 4-157: Subgruppenergebnisse für SF-36 Einzeldimensionen Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter

Studie A3921064		Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
		Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Ko	mbi: mehr als ein csDM	ARD in Vortherapio	2		
Körperliche	Alter ≤ 65 Jahre	41/87 (47,1)	46/92 (50,0)	0,9 [0,7-1,3] p=0,701	TIT.
Funktion	Alter > 65 Jahre	4/13 (30,8)	8/11 (72,7)	0,4 [0,2-1,03] p=0,059	- Hinweis
Emotionale	Alter ≤ 65 Jahre	37/87 (42,5)	33/92 (35,9)	1,2 [0,8-1,7] p=0,363	TIT.
Rollenfunktion	Alter > 65 Jahre	1/13 (7,7)	4/11 (36,4)	0,2 [0,03-1,6] p=0,135	- Hinweis
Soziale Funktion	Alter ≤ 65 Jahre	45/87 (51,7)	44/92 (47,8)	1,1 [0,8-1,5] p=0,602	
	Alter > 65 Jahre	2/13 (15,4)	6/11 (54,5)	0,3 [0,1-1,1] p=0,073	- Hinweis
Quelle: (40) Abkürzungen: R	R = Relatives Risiko, 9.				

Die Responderraten für "Körperliche Funktion" sind für die jüngeren Patienten vergleichbar, für die älteren höher im Adalimumab-Arm. Es liegen keine statistisch signifikanten Unterschiede vor.

Für "Emotionale Rollenfunktion" und "Soziale Funktion" liegt die Responderrate bei den Jüngeren jeweils höher im Tofacitinib-Arm, bei den Älteren liegt sie jeweils höher im Adalimumab-Arm. Es liegen keine statistisch signifikanten Unterschiede vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegt ein Beleg für Effektmodifikation für "Allgemeine Gesundheit" und ein Hinweis auf Effektmodifikation für "Psychisches Wohlbefinden" in der Population d)-Kombi vor (p=0,006 bzw. p=0,195).

Tabelle 4-158: Subgruppenergebnisse für SF-36 Einzeldimensionen Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region

Studie A3921064		Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
		Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Ko	mbi: mehr als ein csDM	ARD in Vortherapie	2		
	Europa	22/63	33/60	0,6 [0,4-0,95]	
		(34,9)	(55,0)	p=0,029	
	USA/Kanada	5/9	1/11	6,1 [0,9-43,3]	_
Allgemeine		(55,6)	(9,1)	p=0,070	– Beleg
Gesundheit	Lateinamerika	2/6	7/11	0,5 [0,2-1,8]	
		(33,3)	(63,6)	p=0,298	
	Sonstige	11/22	3/21	3,5 [1,1-10,8]	
		(50,0)	(14,3)	p=0,030	
	Europa	25/63	29/60	0,8 [0,6-1,2]	
		(39,7)	(48,3)	p=0,336	
	USA/Kanada	5/9	3/11	2,0 [0,7-6,3]	_
Psychisches		(55,6)	(27,3)	p=0,216	***
Wohlbefinden	Lateinamerika	1/6	3/11	0,6 [0,1-4,7]	- Hinweis
		(16,7)	(27,3)	p=0,635	
	Sonstige	10/22	5/21	1,9 [0,8-4,7]	_
		(45,5)	(23,8)	p=0,155	

Abkürzungen: RR = Relatives Risiko, 95%-KI = 95%-Konfidenzintervall, MTX = Methotrexat

In Europa und Lateinamerika liegt die Responderrate für "Allgemeine Gesundheit" jeweils höher im Adalimumab-Arm. Für Europa ist dieser Unterschied zudem statistisch signifikant (p=0,029). In USA/Kanada und den "sonstigen Regionen" ist die Rate jeweils höher im Tofacitinib-Arm. Für die "sonstigen Regionen" ist dieser Unterschied statistisch signifikant (p=0,030).

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

In Europa und Lateinamerika liegt die Responderrate für "Psychisches Wohlbefinden" jeweils höher im Adalimumab-Arm. In USA/Kanada und den "sonstigen Regionen" ist die Rate jeweils höher im Tofacitinib-Arm. Keiner dieser beobachteten Unterschiede ist statistisch signifikant.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegt ein Hinweis auf Effektmodifikation für "Soziale Funktion" in der Population c)-Kombi vor (p=0,112).

Tabelle 4-159: Subgruppenergebnisse für SF-36 Einzeldimensionen Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad

Studie A3921064		Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
		Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population c)-K	ombi: nur ein csDMARD in	Vortherapie			
Soziale	Schweregrad: DAS28-4(ESR) $\leq$ 5,1	4/6 (66,7)	1/5 (20,0)	3,3 [0,5-21,0] p=0,200	Himmi
Funktion	Schweregrad: DAS28-4(ESR) > 5,1	31/69 (44,9)	39/66 (59,1)	0,8 [0,6-1,1] p=0,103	- Hinweis
Quelle: (40) Abkürzungen: I	RR = Relatives Risiko, 95%	o-KI = 95%-Kon	fidenzintervall, M	TTX = Methotrexat	

Bei Patienten mit niedrigerem Schweregrad ist die Responderrate für "Soziale Funktion" unter Tofacitinib + MTX höher als unter Adalimumab + MTX. Bei der Interpretation dieses Ergebnisses sollte jedoch die geringe Fallzahl dieser Subgruppe berücksichtigt werden. Für die Gruppe der Patienten mit höherem Schweregrad ist die Responderrate im Adalimumab-Arm höher. Es liegen keine statistisch signifikanten Unterschiede vor.

# 4.3.1.3.32.12Subgruppenanalysen für Endpunkt FACIT-F nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-160: p-Werte der Interaktionstests für FACIT-F Responderanalyse nach 12 Monaten (Verbesserung  $\geq$  4)

Subgruppenfaktor	Interaktions-p-Wert
Population c)-Kombi: nur ein csDN	AARD in Vortherapie
Geschlecht	p=0,871
Alter	p=0,834
Region	p=0,699
DAS28-4(ESR)	p=0,268
Population d)-Kombi: mehr als ein	csDMARD in Vortherapie
Geschlecht	p=0,405
Alter	p=0,537
Region	p=0,674
DAS28-4(ESR)	p=0,874

Quelle: (40)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.13 Subgruppenanalysen für Endpunkt HAQ-DI nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-161: p-Werte der Interaktionstests für HAQ-DI nach 12 Monaten

Subgruppenfaktor		Interaktions-p-Wert				
Population c)-Kombi: nur ein csDMARD in Vortherapie						
	<b>HAQ-DI</b> ≥ 0,22	$HAQ-DI \ge 0.3$	$HAQ-DI \ge 0.5$			
Geschlecht	p=0,338	p=0,310	p=0,864			
Alter	p=0,409	p=0,498	p=0,203			
Region	p=0,242	p=0,379	p=0,876			
DAS28-4(ESR)	p=0,196	p=0,484	p=0,990			
Population d)-Kombi: meh	r als ein csDMARD in Vortherap	pie				
	<b>HAQ-DI</b> ≥ 0,22	$HAQ-DI \ge 0.3$	$HAQ-DI \ge 0.5$			
Geschlecht	p=0,632	p=0,876	p=0,940			
Alter	p=0,322	p=0,169	p=0,576			
Region	p=0,677	p=0,830	p=0,794			
DAS28-4(ESR)	p=0,636	p=0,410	p=0,439			

Quelle: (40)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegt jeweils ein Hinweis auf Effektmodifikation für den Schwellenwert  $HAQ-DI \ge 0.3$  in der Population d)-Kombi vor (p=0,169).

Tabelle 4-162: Subgruppenergebnisse für HAQ-DI ≥ 0,3 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mehr	als ein csDMARD in \	Vortherapie		
Alter ≤ 65 Jahre	46/87 (52,9)	51/92 (55,4)	1,0 [0,7-1,3] p=0,731	
Alter > 65 Jahre	4/13 (30,8)	7/11 (63,6)	0,5 [0,2-1,2] p=0,126	- Hinweis
Quelle: (40)				
Abkürzungen: RR = Re Methotrexat	elatives Risiko, 95	%-KI = 95%-Ko	onfidenzintervall, MTX =	

In der Gruppe der jüngeren Patienten sind die Responderraten vergleichbar in beiden Studienarmen. Bei den älteren Patienten ist die Rate im Adalimumab-Arm höher. Es liegen keine statistisch signifikanten Unterschiede vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegt ein Hinweis auf Effektmodifikation für den Schwellenwert  $HAQ-DI \ge 0,22$  in der Population c)-Kombi vor (p=0,196).

Tabelle 4-163: Subgruppenergebnisse für HAQ-DI ≥ 0,22 nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population c)-Kombi: nur ein csD	MARD in Vorth	erapie		
Schweregrad: DAS28-4(ESR) ≤ 5,1	1/6 (16,7)	3/5 (60,0)	0,3 [0,0-1,9] p=0,193	
Schweregrad: DAS28-4(ESR) > 5,1	44/69 (63,8)	42/66 (63,6)	1,0 [0,8-1,3] p=0,987	- Hinweis
Quelle: (40)				
Abkürzungen: RR = Relativ Methotrexat	es Risiko, 959	%-KI = 95%-Ko	nfidenzintervall, MTX =	

Bei Patienten mit niedrigerem Schweregrad ist die Responderrate höher im Adalimumab-Arm. Bei der Interpretation dieses Ergebnisses sollte jedoch die geringe Fallzahl dieser Subgruppe berücksichtigt werden. Für die Gruppe der Patienten mit höherem Schweregrad sind die Responderraten vergleichbar in beiden Studienarmen. Es liegen keine statistisch signifikanten Unterschiede vor.

# 4.3.1.3.32.14 Subgruppenanalysen für Endpunkt WLQ nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

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

Tabelle 4-164: p-Werte der Interaktionstests für WLQ nach 12 Monaten

Subgruppenfaktor		Interaktions-p-Wert					
Population c)-Kombi: nur ein csD	MARD in Vorthera	oie					
	Indexscore ≥ 5	Zeitmanagement ≥ 13	Körperliche Anforderungen ≥ 13	Mentale/Interpersonelle Anforderungen ≥ 13	Ergebnisorientierte Anforderungen ≥ 13		
Geschlecht	p=0,935	p=0,550	p=0,396	p=0,581	p=0,901		
Alter	NA	NA	NA	NA	NA		
Region	NA	NA	NA	NA	NA		
DAS28-4(ESR)	NA	p=0,956	NA	NA	p=0,733		
Population d)-Kombi: mehr als ei	n csDMARD in Vort	herapie					
	Indexscore ≥ 5	Zeitmanagement ≥ 13	Körperliche Anforderungen ≥ 13	Mentale/Interpersonelle Anforderungen ≥ 13	Ergebnisorientierte Anforderungen ≥ 13		
Geschlecht	p=0,681	p=0,478	NA	p=0,582	p=0,908		
Alter	NA	NA	NA	NA	NA		
Region	NA	NA	NA	NA	NA		
DAS28-4(ESR)	NA	p=0,203	p=0,058	p=0,626	p=0,390		

Stand: 21.04.2017

Quelle: (40)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

#### **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegt ein Hinweis auf Effektmodifikation für "Körperliche Anforderungen" in der Population d)-Kombi vor (p=0,058).

Tabelle 4-165: Subgruppenergebnisse für WLQ Responderanalysen nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad

Studie A3921064		Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
		Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kor	mbi: mehr als ein csDMAR	D in Vortherapie			
Körperliche	Schweregrad: DAS28-4(ESR) $\leq$ 5,1	2/3 (66,7)	1/5 (20,0)	3,3 [0,5-22,9] p=0,221	II'a ata
Anforderungen	Schweregrad: DAS28-4(ESR) > 5,1	4/33 (12,1)	8/26 (30,8)	0,4 [0,1-1,2] p=0,092	- Hinweis
Quelle: (40) Abkürzungen: RI	R = Relatives Risiko, 95%	o-KI = 95%-Kon	fidenzintervall, M	TTX = Methotrexat	

Bei Patienten mit niedrigerem Schweregrad ist die Responderrate unter Tofacitinib + MTX höher als unter Adalimumab + MTX. Bei der Interpretation dieses Ergebnisses sollte jedoch die geringe Fallzahl dieser Subgruppe berücksichtigt werden. Für die Gruppe der Patienten mit höherem Schweregrad ist die Responderrate im Adalimumab-Arm höher. Es liegen keine statistisch signifikanten Unterschiede vor.

Eine sinnvolle Interpretation dieses Endpunkts ist jedoch hier auf Subgruppenebene wie auf Ebene der Gesamtpopulation nicht möglich aufgrund der niedrigen Patientenzahl des FAS, welche ein hohes Verzerrungspotenzial zur Folge hat (siehe 4.3.1.3.15).

# 4.3.1.3.32.15 Subgruppenanalysen für Endpunkt MOS Schlafscore nach 12 Monaten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-166: p-Werte der Interaktionstests für MOS Schlafscore Responderanalyse (Verbesserung  $\geq 6$ )

Subgruppenfaktor Interaktions-p-Wer			
Population c)-Kombi: nur ein csDM	AARD in Vortherapie		
Geschlecht	p=0,548		
Alter	p=0,763		
Region	p=0,195		
DAS28-4(ESR)	p=0,588		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie			
Geschlecht	p=0,580		
Alter	p=0,091		
Region	p=0,328		
DAS28-4(ESR)	p=0,055		

Quelle: (40)

# **Subgruppen:** Geschlecht

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegt ein Hinweis auf Effektmodifikation in der Population d)-Kombi vor (p=0,091).

Tabelle 4-167: Subgruppenergebnisse für MOS Schlafscore Responderanalyse nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mehr als	ein csDMARD in V	Vortherapie		
Alter ≤ 65 Jahre	31/86	34/92	1,0 [0,7-1,4]	
	(36,0)	(37,0)	p=0,900	- Hinweis
Alter > 65 Jahre	2/13	6/11	0,3 [0,1-1,1]	Hinweis
	(15,4)	(54,5)	p=0,073	
Quelle: (40) Abkürzungen: RR = Relati Methotrexat	ves Risiko, 95º	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In der Gruppe der jüngeren Patienten sind die Raten in beiden Studienarmen vergleichbar. Bei den Älteren ist die Rate unter Adalimumab + MTX höher. Es liegen keine statistisch signifikanten Unterschiede vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegt ein Hinweis auf Effektmodifikation in der Population c)-Kombi vor (p=0,195).

Tabelle 4-168: Subgruppenergebnisse für MOS Schlafscore Responderanalyse nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Region

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population c)-Kombi: nur ei	n csDMARD in Vorth	erapie		
Europa	15/41 (36,6)	19/40 (47,5)	0,8 [0,5-1,3] p=0,323	
USA/Kanada	9/18 (50,0)	4/12 (33,3)	1,5 [0,6-3,8] p=0,390	TI'm ata
Lateinamerika	2/10 (20,0)	7/10 (70,0)	0,3 [0,1-1,1] p=0,060	- Hinweis
Sonstige	3/9 (33,3)	3/12 (25,0)	1,3 [0,4-5,1] p=0,676	_
Quelle: (40) Abkürzungen: RR = Re Methotrexat	elatives Risiko, 95	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In Europa und Lateinamerika sind höhere Responderraten im Adalimumab-Arm zu sehen, in USA/Kanada und den sonstigen Regionen im Tofacitinib-Arm. Die beobachteten Unterschiede sind jeweils nicht statistisch signifikant.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegt ein Hinweis auf Effektmodifikation in der Population d)-Kombi vor (p=0,055).

Tabelle 4-169: Subgruppenergebnisse für MOS Schlafscore Responderanalyse nach 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Schweregrad

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Responder/N (%)	Anzahl Responder/N (%)	RR [95%-KI] p-Wert	
Population d)-Kombi: mehr als ei	n csDMARD in \	Vortherapie		
Schweregrad: DAS28-4(ESR) ≤ 5,1	3/4 (75,0)	3/10 (30,0)	2,5 [0,8-7,5] p=0,104	TT.
Schweregrad: DAS28-4(ESR) > 5,1	29/94 (30,9)	35/90 (38,9)	0,8 [0,5-1,2] p=0,255	- Hinweis
Quelle: (40) Abkürzungen: RR = Relativ Methotrexat	es Risiko, 95	%-KI = 95%-Ko	nfidenzintervall, MTX =	

In der Gruppe mit niedrigerem Schweregrad liegt die Responderrate höher im Tofacitinib-Arm. Dieser Unterschied ist jedoch nicht statistisch signifikant, und bei der Interpretation muss die niedrige Fallzahl berücksichtigt werden. Bei den Patienten mit höherem Schweregrad ist die Rate höher im Adalimumab-Arm. Auch dieser Unterschied ist nicht statistisch signifikant.

# 4.3.1.3.32.16 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse bis 12 Monate – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-170: p-Werte der Interaktionstests für unerwünschte Ereignisse bis 12 Monate

Subgruppenfaktor	Interaktions-p-Wert		
Population c)-Kombi: nur ein csDM	AARD in Vortherapie		
Geschlecht	p=0,012		
Alter	p=0,274		
Region	p=0,908		
DAS28-4(ESR)	p=0,879		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie			
Geschlecht	p=0,626		
Alter	p=0,712		
Region	p=0,460		
DAS28-4(ESR)	p=0,274		

Quelle: (41)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegt ein Beleg für Effektmodifikation in der Population c)-Kombi vor (p=0,012).

Tabelle 4-171: Subgruppenergebnisse für unerwünschte Ereignisse bis 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Patienten mit Ereignis/N	Anzahl Patienten mit Ereignis/N	RR [95%-KI] p-Wert	
Population c)-Kombi: nur ein	(%)	(%)	-	
1 opulation c)-ixollibi. Ilui en	i CSDIVIARD III VOI UI	et apie		
männlich	5/13	10/12	0,5 [0,2-0,96]	
	(38,5)	(83,3)	p=0,039	- Palag
weiblich	56/70	43/66	1,2 [0,99-1,5]	- Beleg
	(80,0)	(65,2)	p=0,057	
Quelle: (41) Abkürzungen: RR = Re Methotrexat	latives Risiko, 95	%-KI = 95%-Ko	nfidenzintervall, MTX =	

Bei den Männern ist die Rate statistisch signifikant niedriger unter Tofacitinib + MTX (p=0,039). Bei den Frauen ist die Rate im Adalimumab-Arm niedriger. Dieser Unterschied ist noch nicht statistisch signifikant.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.17 Subgruppenanalysen für Endpunkt schwerwiegende unerwünschte Ereignisse bis 12 Monate – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-172: p-Werte der Interaktionstests für schwerwiegende unerwünschte Ereignisse bis 12 Monate

Subgruppenfaktor	Interaktions-p-Wert			
Population c)-Kombi: nur ein csDMARD in Vortherapie				
Geschlecht	p=0,272			
Alter	p=0,032			
Region	p=0,781			
DAS28-4(ESR)	p=0,382			
Population d)-Kombi: mehr als ein	csDMARD in Vortherapie			
Geschlecht	p=0,679			
Alter	p=0,604			
Region	p=0,333			
DAS28-4(ESR)	NA			

Quelle: (41)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegt ein Beleg für Effektmodifikation in der Population c)-Kombi vor (p=0,032).

Tabelle 4-173: Subgruppenergebnisse für schwerwiegende unerwünschte Ereignisse bis 12 Monaten aus RCT mit dem zu bewertenden Arzneimittel: Alter

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Patienten mit Ereignis/N (%)	Anzahl Patienten mit Ereignis/N (%)	RR [95%-KI] p-Wert	
Population c)-Kombi: nur ein csDMARD in Vortherapie				
Alter ≤ 65 Jahre	12/64	2/66	6,2 [1,4-26,6]	
	(18,8)	(3,0)	p=0,014	Dalas
Alter > 65 Jahre	1/19	2/12	0,3 [0,0-3,1]	- Beleg
	(5,3)	(16,7)	p=0,324	
Quelle: (41) Abkürzungen: RR = Relati Methotrexat	ves Risiko, 95	%-KI = 95%-Ko	nfidenzintervall, MTX =	

Für die jüngeren Patienten ist die Rate schwerwiegender unerwünschter Ereignisse im Tofactinib-Arm statistisch signifikant höher (p=0,014). Bei den Älteren ist die Rate unter Adalimumab + MTX höher. Dieser Unterschied ist jedoch nicht statistisch signifikant.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Schweregrad nach DAS28-4(ESR)

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.18 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse, die bis Monat 12 zum Therapieabbruch geführt hatten – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-174: p-Werte der Interaktionstests für unerwünschte Ereignisse, die bis Monat 12 zum Therapieabbruch geführt hatten

Subgruppenfaktor	Interaktions-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie			
Geschlecht	NA		
Alter	NA		
Region	NA		
DAS28-4(ESR)	NA		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie			
Geschlecht	p=0,439		
Alter	p=0,548		
Region	NA		
DAS28-4(ESR)	NA		

Quelle: (41)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.19 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Schwere Infektionen – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-175: p-Werte der Interaktionstests für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Schwere Infektionen

Subgruppenfaktor	Interaktions-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie			
Geschlecht	NA		
Alter	NA		
Region	NA		
DAS28-4(ESR)	NA		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie			
Geschlecht	NA		
Alter	NA		
Region	NA		
DAS28-4(ESR)	NA		

Quelle: (41)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.20 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Jedwede Infektion und parasitäre Erkrankung – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-176: p-Werte der Interaktionstests für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Jedwede Infektion und parasitäre Erkrankung

Subgruppenfaktor	Interaktions-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie			
Geschlecht	p=0,101		
Alter	p=0,316		
Region	p=0,773		
DAS28-4(ESR)	p=0,719		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie			
Geschlecht	p=0,462		
Alter	p=0,544		
Region	p=0,355		
DAS28-4(ESR)	NA		

Quelle: (41)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegt ein Hinweis auf Effektmodifikation in der Population c)-Kombi vor (p=0,101).

Tabelle 4-177: Subgruppenergebnisse für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Jedwede Infektion und parasitäre Erkrankung aus RCT mit dem zu bewertenden Arzneimittel: Geschlecht

Studie A3921064	Tofacitinib + MTX	Adalimumab + MTX	Behandlungseffekt (Tofacitinib vs. Adalimumab)	Effekt- modifikation
	Anzahl Patienten mit Ereignis/N (%)	Anzahl Patienten mit Ereignis/N (%)	RR [95%-KI] p-Wert	
Population c)-Kombi: nur ein csl	OMARD in Vorth	erapie		
männlich	4/13	6/12	0,6 [0,2-1,7]	
	(30,8)	(50,0)	p=0,338	- Hinweis
weiblich	31/70	19/66	1,5 [0,97-2,4]	Hillweis
	(44,3)	(28,8)	p=0,067	
Quelle: (41)				
Abkürzungen: RR = Relativ Methotrexat	ves Risiko, 95º	%-KI = 95%-Ko	nfidenzintervall, MTX =	

Bei den Männern ist die Rate im Adalimumab-Arm höher, bei den Frauen hingegen im Tofacitinib-Arm. Es liegen jedoch keine statistisch signifikanten Unterschiede vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.21 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Herpes Zoster – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-178: p-Werte der Interaktionstests für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Herpes Zoster

Subgruppenfaktor	Interaktions-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie			
Geschlecht	NA		
Alter	NA		
Region	NA		
DAS28-4(ESR)	NA		
Population d)-Kombi: mehr als ein csDMARD in Vortherapie			
Geschlecht	NA		
Alter	NA		
Region	NA		
DAS28-4(ESR)	NA		

Quelle: (41)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.22 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Tuberkulose – RCT

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt, da keine Ereignisse berichtet wurden.

# 4.3.1.3.32.23 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Nicht-Melanom Hautkrebs – RCT

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt, da sowohl für Population c)-Kombi als auch für Population d)-Kombi jeweils nur ein Ereignis berichtet wurde.

# 4.3.1.3.32.24 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Gastrointestinale Perforation – RCT

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt, da keine Ereignisse berichtet wurden.

# 4.3.1.3.32.25 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Malignität (ohne Nicht-Melanom Hautkrebs) – RCT

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt, da insgesamt nur sehr wenige Ereignisse berichtet wurden.

# 4.3.1.3.32.26 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Erhöhte Leberwerte – RCT

Folgende Tabelle gibt einen Überblick über die p-Werte der Interaktionstests für diesen Endpunkt.

Tabelle 4-179: p-Werte der Interaktionstests für unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Erhöhte Leberwerte

Subgruppenfaktor	Interaktions-p-Wert		
Population c)-Kombi: nur ein csDMARD in Vortherapie			
Geschlecht	NA		
Alter	NA		
Region	NA		
DAS28-4(ESR)	NA		
Population d)-Kombi: mehr als ein	csDMARD in Vortherapie		
Geschlecht	NA		
Alter	NA		
Region	NA		
DAS28-4(ESR)	NA		

Quelle: (41)

# **Subgruppen: Geschlecht**

Bezüglich des Subgruppenmerkmals "Geschlecht" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# Subgruppen: Alter

Bezüglich des Subgruppenmerkmals "Alter" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Region**

Bezüglich des Subgruppenmerkmals "Region" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# **Subgruppen: Schweregrad nach DAS28-4(ESR)**

Bezüglich des Subgruppenmerkmals "Schweregrad" liegen jeweils keine Hinweise oder Belege auf Effektmodifikation vor.

# 4.3.1.3.32.27 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Erhöhte Kreatininwerte im Blut – RCT

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt, da insgesamt nur sehr wenige Ereignisse berichtet wurden.

# 4.3.1.3.32.28 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Neutropenie – RCT

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt, da insgesamt nur sehr wenige Ereignisse berichtet wurden.

# 4.3.1.3.32.29 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Lymphopenie – RCT

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt, da keine Ereignisse berichtet wurden.

# 4.3.1.3.32.30 Subgruppenanalysen für Endpunkt unerwünschte Ereignisse von besonderem Interesse bis 12 Monate: Anämie – RCT

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt, da insgesamt nur sehr wenige Ereignisse berichtet wurden.

#### 4.3.1.3.33 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.

Nachfolgend sind die Ergebnisse der in den obigen Abschnitten dargestellten Effekte der eingeschlossenen Studie ORAL STANDARD zusammengefasst. Tabelle 4-191 zeigt die Ergebnisse für die Teilpopulation c)-Kombi, Tabelle 4-192 die Ergebnisse der Teilpopulation d)-Kombi.

Subgruppenergebnisse werden in den Tabellen jeweils aufgeführt, falls im konkreten Fall mindestens ein Hinweis auf Effektmodifikation vorliegt (d.h. Interaktions-p-Wert < 0,2). Eine Interpretation des Zusatznutzens auf Ebene der Subgruppen ist erforderlich und wird in Abschnitt 4.4.2 vorgenommen, wenn ein Beleg für eine Effektmodifikation vorliegt (Interaktions-p-Wert < 0,05).

Tabelle 4-180: Übersicht über die Ergebnisse aus RCT auf Endpunktebene nach 12 Monaten – Population c)-Kombi

Studie: A3921064 (ORAL STANDARD)	RR [95% KI], p-Wert	Interaktionstest p-Wert (falls < 0,2)
DAS28-4(ESR) < 2,6	0,8 [0,3-1,7], p=0,484	
$DAS28-4(ESR) \le 3,2$	0,8 [0,4-1,4], p=0,370	
DAS28-4(CRP) < 2,6	1,1 [0,7-1,8], p=0,622	
$DAS28-4(CRP) \le 3,2$	1,1 [0,8-1,6], p=0,512	
EULAR-Response (ESR): gute Verbesserung	0,8 [0,4-1,4], p=0,370	
EULAR-Response (ESR): gute oder moderate Verbesserung	1,1 [0,9-1,4], p=0,366	
EULAR-Response (CRP): gute Verbesserung	1,1 [0,8-1,6], p=0,506	
EULAR-Response (CRP): gute oder moderate Verbesserung	1,0 [0,8-1,2], p=0,765	
ACR70	1,7 [0,9-3,3], p=0,084	
ACR50	1,4 [0,9-2,0], p=0,130	
ACR/EULAR-Remission	1,7 [0,5-5,5], p=0,402	
SDAI ≤ 3,3	1,9 [0,8-4,8], p=0,176	
SDAI ≤ 11	1,2 [0,8-1,6], p=0,407	
$CDAI \le 2.8$	1,6 [0,7-3,9], p=0,276	
<b>CDAI ≤ 10</b>	1,3 [0,9-1,8], p=0,190	
Schmerzbeurteilung durch den Patienten (Verbesserung $\geq 10$ )	0,9 [0,7-1,2], p=0,573	
Gesamtbeurteilung der Krankheitsaktivität durch den Patienten [PatGA] (Verbesserung $\geq 10$ )	0,9 [0,7-1,2], p=0,463	

Stand: 21.04.2017

Tofacitinib (XELJANZ®)

Seite 334 von 442

C	. 10)	10000101 0667	
Gesamtbeurteilung der Krankheitsaktivität durch den Arzt [PhysGA] (Verbesserung ≥	•	1,0 [0,8-1,2], p=0,667	
	Europa USA/Kanada	1,1 [0,9-1,4], p=0,245 1,2 [0,6-2,4], p=0,560	0,096
	Lateinamerika	0,4 [0,2-0,98], p=0,043	0,096
	Sonstige	0,9 [0,6-1,3], p=0,407	
	Sollstige	_	
Anzahl empfindlicher/druckschmerzhafter Gelenke $\leq 1-28$ untersuchte Gelenke		1,0 [0,6-1,7], p=0,887	
Anzahl empfindlicher/druckschmerzhafter Gelenke ≤ 1 – 68 untersuchte Gelenke		0,9 [0,5-1,7], p=0,717	
Anzahl geschwollener Gelenke $\leq 1-28$ untersuchte Gelenke		1,0 [0,7-1,4], p=0,977	
Anzahl geschwollener Gelenke ≤ 1 – 66 untersuchte Gelenke		1,1 [0,8-1,7], p=0,574	
	Alter $\leq$ 65 Jahre	1,0 [0,6-1,5], p=0,886	0,137
	Alter > 65 Jahre	2,8 [0,7-11,0], p=0,130	
SF-36			
- Körperlicher Komponentenscore (Verbesserung $\geq 2,5$ )		1,0 [0,7-1,2], p=0,686	
- Psychischer Komponentenscore (Verbesserung ≥ 2,5)		0,8 [0,6-1,2], p=0,350	
- Vitalität (Verbesserung ≥ 5)		0,9 [0,7-1,3], p=0,527	
- Körperliche Funktion (Verbesserung ≥ 5)		1,0 [0,7-1,4], p=0,997	
- Körperliche Schmerzen (Verbesserung $\geq 5$ )		0,9 [0,7-1,2], p=0,522	
- Allgemeine Gesundheit (Verbesserung ≥ 5)		1,1 [0,7-1,7], p=0,670	
- Körperliche Rollenfunktion (Verbesserung $\geq 5$ )		1,1 [0,7-1,6], p=0,687	
- Emotionale Rollenfunktion (Verbesserung $\geq 5$ )		1,0 [0,7-1,6], p=0,920	
- Psychisches Wohlbefinden (Verbesserung ≥ 5)		0,9 [0,6-1,4], p=0,675	
	männlich	0,5 [0,2-1,5], p=0,185	0,196
	weiblich	1,0 [0,7-1,6], p=0,880	
- Soziale Funktion (Verbesserung ≥ 5)		0,8 [0,6-1,1], p=0,253	
	$DAS28-4(ESR) \le 5,1$	3,3 [0,5-21,0], p=0,200	0,112
	DAS28-4(ESR) > 5,1	0,8 [0,6-1,1], p=0,103	
FACIT-F (Verbesserung ≥ 4)		0,9 [0,6-1,2], p=0,332	

Stand: 21.04.2017

Tofacitinib (XELJANZ®) Seite 335 von 442

			T
$HAQ-DI \ge 0,22$		1,0 [0,8-1,2], p=0,670	
DAS28-4(E		0,3 [0,0-1,9], p=0,193	0,196
DAS28-4(E	(SR) > 5,1	1,0 [0,8-1,3], p=0,987	
$HAQ-DI \ge 0.3$	(	0,8 [0,6-1,1], p=0,183	
$HAQ-DI \ge 0.5$	(	0,8 [0,6-1,1], p=0,206	
WLQ			
- Indexscore (Verbesserung $\geq 5$ )		1,1 [0,5-2,7], p=0,798	
- Zeitmanagement (Verbesserung ≥ 13)	(	0,5 [0,2-0,9], p=0,030	
- Körperliche Anforderungen (Verbesserung ≥ 13)	(	0,7 [0,3-1,8], p=0,446	
- Mentale/Interpersonelle Anforderungen (Verbesserung $\geq 13$ )		1,2 [0,5-2,5], p=0,732	
- Ergebnisorientierte Anforderungen (Verbesserung $\geq 13$ )	(	0,7 [0,3-1,4], p=0,289	
MOS Schlafscore (Verbesserung $\geq 6$ )		0,8 [0,6-1,2], p=0,354	
	Europa	0,8 [0,5-1,3], p=0,323	
USA	A/Kanada	1,5 [0,6-3,8], p=0,390	0,195
Latei	inamerika	0,3 [0,1-1,1], p=0,060	
	Sonstige	1,3 [0,4-5,1], p=0,676	
Unerwünschte Ereignisse		1,1 [0,9-1,3], p=0,442	
	männlich	0,5 [0,2-0,96], p=0,039	0,012
	weiblich	1,2 [0,99-1,5], p=0,057	
Schwerwiegende unerwünschte Ereignisse	,	3,1 [1,04-9,0], p=0,042	
	≤ 65 Jahre	6,2 [1,4-26,6], p=0,014	0,032
Alter >	> 65 Jahre	0,3 [0,0-3,1], p=0,324	
Unerwünschte Ereignisse, die zum Therapieabbruch geführt hatten		1,5 [0,5-4,4], p=0,457	
Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen		4,7 [0,2-96,4], p=0,315	
Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung	g	1,3 [0,9-2,0], p=0,189	
	männlich	0,6 [0,2-1,7], p=0,338	0,101
	weiblich	1,5 [0,97-2,4], p=0,067	

Stand: 21.04.2017

Tofacitinib (XELJANZ®) Seite 336 von 442

# Dossier zur Nutzenbewertung – Modul 4 A

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

Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster	6,6 [0,3-125,4], p=0,210
Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose	nicht definiert
Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs	0,3 [0,0-7,6], p=0,475
Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation	nicht definiert
Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs)	2,8 [0,1-68,2], p=0,523
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte	3,8 [0,4-32,9], p=0,232
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut	nicht definiert
Unerwünschte Ereignisse von besonderem Interesse: Neutropenie	nicht definiert
Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie	nicht definiert
Unerwünschte Ereignisse von besonderem Interesse: Anämie	2,8 [1,2-68,2], p=0,523

Stand: 21.04.2017

Tofacitinib (XELJANZ®) Seite 337 von 442

Tabelle 4-181: Übersicht über die Ergebnisse aus RCT auf Endpunktebene nach 12 Monaten – Population d)-Kombi

Studie: A3921064 (ORAL STANDARD)		RR [95% KI], p-Wert	Interaktionstest p-Wert (falls < 0,2)
DAS28-4(ESR) < 2,6		0,5 [0,2-0,96], p=0,038	
	$DAS28-4(ESR) \le 5,1$	1,5 [0,7-3,5], p=0,344	0,029
	DAS28-4(ESR) > 5,1	0,4 [0,2-0,9], p=0,035	
$DAS28-4(ESR) \le 3,2$		0,7 [0,4-1,3], p=0,250	
DAS28-4(CRP) < 2,6		1,0 [0,6-1,5], p=0,849	
$DAS28-4(CRP) \le 3,2$		1,0 [0,8-1,4], p=0,854	
EULAR-Response (ESR): gute Verbesserung		0,7 [0,4-1,3], p=0,250	
EULAR-Response (ESR): gute oder moderate Verbesserung		1,0 [0,8-1,2], p=0,828	
EULAR-Response (CRP): gute Verbesserung		1,1 [0,8-1,5], p=0,705	
EULAR-Response (CRP): gute oder moderate Verbesserung		0,9 [0,8-1,1], p=0,336	
ACR70		1,2 [0,7-2,1], p=0,430	
	Europa	0,8 [0,4-1,5], p=0,421	
	USA/Kanada	6,1 [0,9-43,3], p=0,070	0,162
	Lateinamerika	1,8 [0,3-9,9], p=0,482	
	Sonstige	1,9 [0,6-6,7], p=0,311	
ACR50		1,0 [0,7-1,4], p=0,971	
ACR/EULAR-Remission		1,1 [0,5-2,6], p=0,763	
SDAI ≤ 3,3		0,9 [0,4-1,6], p=0,621	
SDAI ≤ 11		1,3 [0,9-1,8], p=0,190	
CDAI ≤ 2,8		1,0 [0,5-2,1], p=0,933	

Stand: 21.04.2017

Tofacitinib (XELJANZ®)

Seite 338 von 442

CDAI ≤ 10	1 2 [0 0 1 7] - 0 276	
· · · · · · · · · · · · · · · · · · ·	1,2 [0,8-1,7], p=0,376	
Europa	0,9 [0,6-1,3], p=0,438	
USA/Kanada	2,4 [0,8-7,1], p=0,102	0,116
Lateinamerika	1,8 [0,5-6,4], p=0,343	
Sonstige	2,2 [0,8-5,9], p=0,140	
Schmerzbeurteilung durch den Patienten (Verbesserung ≥ 10)	0,1 [0,8-1,2], p=0,871	
Alter ≤ 65 Jahre	1,0 [0,8-1,3], p=0,738	0,183
Alter > 65 Jahre	0,6 [0,3-1,3], p=0,196	
Europa	0,8 [0,6-1,04], p=0,091	
USA/Kanada	1,5 [0,7-3,3], p=0,345	0,147
Lateinamerika	1,5 [0,6-3,5], p=0,383	
Sonstige	1,3 [0,8-2,1], p=0,299	
Gesamtbeurteilung der Krankheitsaktivität durch den Patienten [PatGA] (Verbesserung≥10)	1,1 [0,9-1,3], p=0,485	
männlich	1,5 [0,98-2,2], p=0,062	0,132
weiblich	1,0 [0,8-1,3], p=0,878	
Europa	0,9 [0,7-1,1], p=0,334	
USA/Kanada	2,1 [0,9-5,0], p=0,082	0,149
Lateinamerika	1,4 [0,5-4,2], p=0,577	
Sonstige	1,3 [0,8-2,3], p=0,300	
Gesamtbeurteilung der Krankheitsaktivität durch den Arzt [PhysGA] (Verbesserung≥10)	1,0 [0,8-1,1], p=0,626	
männlich	1,4 [0,9-2,0], p=0,102	0,040
weiblich	0,9 [0,7-1,1], p=0,180	
Anzahl empfindlicher/druckschmerzhafter Gelenke $\leq 1-28$ untersuchte Gelenke	0,8 [0,6-1,3], p=0,433	
Anzahl empfindlicher/druckschmerzhafter Gelenke $\leq 1-68$ untersuchte Gelenke	0,7 [0,4-1,2], p=0,215	
Anzahl geschwollener Gelenke ≤ 1 – 28 untersuchte Gelenke	1,2 [0,8-1,7], p=0,376	
Europa	0,9 [0,6-1,3], p=0,438	
USA/Kanada	7,3 [1,1-50,3], p=0,043	0,100
Lateinamerika	1,4 [0,5-4,2], p=0,577	
Sonstige	1,7 [0,7-4,3], p=0,246	

Stand: 21.04.2017

Tofacitinib (XELJANZ®) Seite 339 von 442

Anzahl geschwollener Gelenke $\leq 1-66$ untersuchte Gelenke		1,1 [0,7-1,6], p=0,655	
SF-36			
- Körperlicher Komponentenscore (Verbesserung ≥ 2,5)		1,0 [0,8-1,2], p=0,862	
	Alter ≤ 65 Jahre	1,0 [0,8-1,3], p=0,762	0,187
	Alter > 65 Jahre	0,6 [0,3-1,3], p=0,196	
- Psychischer Komponentenscore (Verbesserung $\geq 2,5$ )		1,1 [0,8-1,5], p=0,640	
- Vitalität (Verbesserung $\geq 5$ )		1,1 [0,8-1,4], p=0,741	
- Körperliche Funktion (Verbesserung $\geq 5$ )		0,9 [0,7-1,1], p=0,292	
	Alter ≤ 65 Jahre	0,9 [0,7-1,3], p=0,701	0,096
	Alter > 65 Jahre	0,4 [0,2-1,03], p=0,059	
- Körperliche Schmerzen (Verbesserung $\geq 5$ )		0,9 [0,7-1,2], p=0,445	
- Allgemeine Gesundheit (Verbesserung ≥ 5)		0,9 [0,7-1,3], p=0,695	
	Europa	0,6 [0,4-0,95], p=0,029	
	USA/Kanada	6,1 [0,9-43,3], p=0,070	0,006
	Lateinamerika	0,5 [0,2-1,8], p=0,298	
	Sonstige	3,5 [1,1-10,8], p=0,030	
- Körperliche Rollenfunktion (Verbesserung $\geq 5$ )		1,0 [0,7-1,4], p=0,918	
- Emotionale Rollenfunktion (Verbesserung $\geq 5$ )		1,1 [0,7-1,5], p=0,759	
	Alter ≤ 65 Jahre	1,2 [0,8-1,7], p=0,363	0,103
	Alter > 65 Jahre	0,2 [0,03-1,6], p=0,135	
- Psychisches Wohlbefinden (Verbesserung ≥ 5)		1,1 [0,8-1,5], p=0,753	
	Europa	0,8 [0,6-1,2], p=0,336	
	USA/Kanada	2,0 [0,7-6,3], p=0,216	0,195
	Lateinamerika	0,6 [0,1-4,7], p=0,635	
	Sonstige	1,9 [0,8-4,7], p=0,155	
- Soziale Funktion (Verbesserung $\geq 5$ )		1,0 [0,7-1,3], p=0,826	
	männlich	1,6 [0,8-3,2], p=0,235	0,146
	weiblich	0,9 [0,6-1,2], p=0,352	

Tofacitinib (XELJANZ®)

	Alter ≤ 65 Jahre	1,1 [0,8-1,5], p=0,602	0,063
	Alter > 65 Jahre	0,3 [0,1-1,1], p=0,073	
FACIT-F (Verbesserung $\geq 4$ )		1,1 [0,8-1,4], p=0,620	
$HAQ-DI \ge 0.22$		0,9 [0,7-1,1], p=0,304	
$HAQ-DI \ge 0.3$		0,9 [0,7-1,2], p=0,369	
	Alter $\leq$ 65 Jahre	1,0 [0,7-1,3], p=0,731	0,169
	Alter > 65 Jahre	0,5 [0,2-1,2], p=0,126	
$HAQ-DI \ge 0.5$		0,9 [0,7-1,2], p=0,435	
WLQ			
- Indexscore (Verbesserung $\geq 5$ )		2,9 [0,9-10,0], p=0,075	
- Zeitmanagement (Verbesserung ≥ 13)		1,7 [0,95-2,9], p=0,077	
- Körperliche Anforderungen (Verbesserung ≥ 13)		0,5 [0,2-1,3], p=0,151	
	$DAS28-4(ESR) \le 5,1$	3,3 [0,5-22,9], p=0,221	0,058
	DAS28-4(ESR) $> 5,1$	0,4 [0,1-1,2], p=0,092	
- Mentale/Interpersonelle Anforderungen (Verbesserung $\geq 13$ )		2,9 [1,2-7,1], p=0,016	
- Ergebnisorientierte Anforderungen (Verbesserung $\geq 13$ )		1,0 [0,6-1,8], p=0,931	
MOS Schlafscore (Verbesserung ≥ 6)		0,9 [0,6-1,2], p=0,417	
	Alter ≤ 65 Jahre	1,0 [0,7-1,4], p=0,900	0,091
	Alter > 65 Jahre	0,3 [0,1-1,1], p=0,073	
	$DAS28-4(ESR) \le 5,1$	2,5 [0,8-7,5], p=0,104	0,055
	DAS28-4(ESR) $> 5,1$	0,8 [0,5-1,2], p=0,255	
Unerwünschte Ereignisse		1,0 [0,8-1,1], p=0,721	
Schwerwiegende unerwünschte Ereignisse		1,3 [0,7-2,6], p=0,415	
Unerwünschte Ereignisse, die zum Therapieabbruch geführt hatten		0,8 [0,4-1,6], p=0,568	
Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen		4,0 [0,5-35,5], p=0,208	
Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und paras	sitäre Erkrankung	0,9 [0,7-1,3], p=0,706	
Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster		0,8 [0,2-3,3], p=0,711	

Stand: 21.04.2017

Tofacitinib (XELJANZ®)

Seite 341 von 442

# Dossier zur Nutzenbewertung – Modul 4 A

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

Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose	nicht definiert
Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs	0,3 [0,0-8,2], p=0,503
Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation	nicht definiert
Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs)	0,5 [0,1-5,5], p=0,574
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte	0,5 [0,1-5,5], p=0,574
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut	1,0 [0,1-15,9], p=0,995
Unerwünschte Ereignisse von besonderem Interesse: Neutropenie	3,0 [0,1-73,5], p=0,496
Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie	nicht definiert
Unerwünschte Ereignisse von besonderem Interesse: Anämie	3,0 [0,1-73,5], p=0,496

Stand: 21.04.2017

Tofacitinib (XELJANZ®)

Seite 342 von 442

#### 4.3.2 Weitere Unterlagen

# 4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

### 4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Es wurde kein indirekter Vergleich durchgeführt.

#### 4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht relevant.

# 4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-182: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<mortalität></mortalität>	<gesundheits- bezogene Lebensqualität&gt;</gesundheits- 	<endpunkt></endpunkt>	<endpunkt></endpunkt>	<endpunkt></endpunkt>
<studie 1=""></studie>	nein	ja	ja	ja	nein

# 4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

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

Tabelle 4-183: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Intervention	<vergleichs- therapie 1&gt;</vergleichs- 	<vergleichs- therapie 2&gt;</vergleichs- 	<vergleichs- therapie 3&gt;</vergleichs- 
1	<studie 1=""></studie>	•		•	•
2	<studie 2=""></studie>	•		•	
	<studie 3=""></studie>	•		•	
1	<studie 4=""></studie>		•	•	•
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht relevant.

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

Tabelle 4-184: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<studie 1=""></studie>	

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 4-185: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial	Verblindung	Adäquate Umsetzung	Ergebnisunabhängige	Keine sonstigen	Verzerrungspotenzial
	auf Studienebene	Endpunkterheber	des ITT-Prinzips	Berichterstattung	Aspekte	Endpunkt
<studie 1=""></studie>	<hoch <br="">niedrig&gt;</hoch>	<ja <br="" nein="">unklar&gt;</ja>	<ja <br="" nein="">unklar&gt;</ja>	<ja <br="" nein="">unklar&gt;</ja>	<ja nein=""></ja>	<hoch <br="">niedrig&gt;</hoch>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht relevant.

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

Tabelle 4-186: Ergebnisse für < Endpunkt xxx> aus RCT für indirekte Vergleiche

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

#### Nicht relevant.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.
- Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.
- Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.

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

Nicht relevant.

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

# 4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.32.

Nicht relevant.

#### 4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

# 4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Es wurden keine nicht randomisierten vergleichenden Studien zum Nachweis des Zusatznutzens herangezogen.

#### 4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-187: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

	lität der	eit der adäquate mg von elevanten	Verbli	ndung	ngige g	Aspekte
Studie	Zeitliche Parallelität Gruppen	Vergleichbarkeit Gruppen bzw. adi Berücksichtigung prognostisch relev Faktoren	Patient	Behandelnde Personen	Ergebnisunabhängige Berichterstaffung	Keine sonstigen
<studie 1=""></studie>	<ja <br="" nein="">unklar&gt;</ja>	<ja <br="" nein="">unklar&gt;</ja>	<ja <br="" nein="">unklar&gt;</ja>	<ja <br="" nein="">unklar&gt;</ja>	<ja <br="" nein="">unklar&gt;</ja>	<ja nein=""></ja>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht relevant.

#### 4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

#### 4.3.2.2.3.1 **Endpunkt** xxx> – nicht randomisierte vergleichende Studien

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

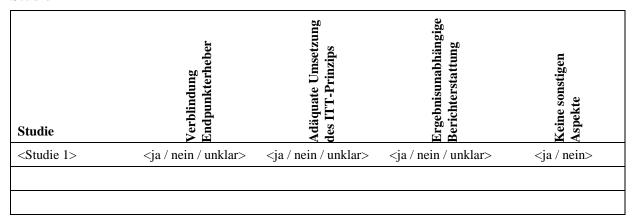
Tabelle 4-188: Operationalisierung von < Endpunkt xxx>

Studie	Operationalisierung
<studie 1=""></studie>	

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 4-189: Verzerrungsaspekte für < Endpunkt xxx> – nicht randomisierte vergleichende Studien



Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht relevant.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht relevant.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.32.

Nicht relevant.

#### 4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

#### 4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Es wurden keine weiteren Untersuchungen zum Nachweis des Zusatznutzens herangezogen.

# 4.3.2.3.2 Charakteristika der weiteren Untersuchungen

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

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

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

Nicht relevant.

# 4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

#### 4.3.2.3.3.1 < Endpunkt xxx> – weitere Untersuchungen

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

Tabelle 4-190: Operationalisierung von < Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<studie 1=""></studie>	

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

Nicht relevant.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht relevant.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

# 4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.32.

Nicht relevant.

#### 4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht relevant, da keine weiteren Unterlagen zur Bewertung des Zusatznutzens herangezogen werden.

#### 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.

### Studienqualität:

Für die Darstellung des Zusatznutzens wurden die Ergebnisse von Subpopulationen einer randomisierten kontrollierten Studie (RCT) von hoher Qualität herangezogen. Die Studien ermöglicht einen direkten Vergleich der zu bewertenden Arzneimittels Tofacitinib mit der zweckmäßigen Vergleichstherapie. Bei der eingeschlossenen Studie handelt es sich um eine doppelblinde Studie mit niedrigem Verzerrungspotenzial.

# Validität der herangezogenen Endpunkte:

Alle Endpunkte, die für die Darstellung der Wirksamkeit und Sicherheit im Rahmen der Nutzenbewertung von Tofacitinib herangezogen wurden, sind patientenrelevant. Bezüglich der Validität der einzelnen Endpunkte siehe Abschnitt 4.2.5.2 in diesem Modul 4A.

#### Evidenzstufe:

Der Nachweis für den Zusatznutzen beruht auf den Ergebnissen einer randomisierten kontrollierten Studie und entspricht somit dem Evidenzgrad Ib.

# 4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- erheblicher Zusatznutzen
- beträchtlicher Zusatznutzen
- geringer Zusatznutzen
- nicht quantifizierbarer Zusatznutzen
- kein Zusatznutzen belegbar
- der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Tabelle 4-191: Übersicht über Wahrscheinlichkeit und Ausmaß des Zusatznutzens aus RCT auf Endpunktebene – Population c)-Kombi

Stand: 21.04.2017

Studie: A3921064 (ORAL STANDARD)	RR [95% KI], p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
DAS28-4(ESR) < 2,6	0,8 [0,3-1,7], p=0,484	Kein Zusatznutzen belegt
$DAS28-4(ESR) \le 3,2$	0,8 [0,4-1,4], p=0,370	Kein Zusatznutzen belegt
DAS28-4(CRP) < 2,6	1,1 [0,7-1,8], p=0,622	Kein Zusatznutzen belegt
$DAS28-4(CRP) \le 3,2$	1,1 [0,8-1,6], p=0,512	Kein Zusatznutzen belegt
EULAR-Response (ESR): gute Verbesserung	0,8 [0,4-1,4], p=0,370	Kein Zusatznutzen belegt
EULAR-Response (ESR): gute oder moderate Verbesserung	1,1 [0,9-1,4], p=0,366	Kein Zusatznutzen belegt
EULAR-Response (CRP): gute Verbesserung	1,1 [0,8-1,6], p=0,506	Kein Zusatznutzen belegt
EULAR-Response (CRP): gute oder moderate Verbesserung	1,0 [0,8-1,2], p=0,765	Kein Zusatznutzen belegt
ACR70	1,7 [0,9-3,3], p=0,084	Kein Zusatznutzen belegt
ACR50	1,4 [0,9-2,0], p=0,130	Kein Zusatznutzen belegt
ACR/EULAR-Remission	1,7 [0,5-5,5], p=0,402	Kein Zusatznutzen belegt
$SDAI \leq 3,3$	1,9 [0,8-4,8], p=0,176	Kein Zusatznutzen belegt
SDAI ≤ 11	1,2 [0,8-1,6], p=0,407	Kein Zusatznutzen belegt
CDAI ≤ 2,8	1,6 [0,7-3,9], p=0,276	Kein Zusatznutzen belegt
CDAI ≤ 10	1,3 [0,9-1,8], p=0,190	Kein Zusatznutzen belegt
Schmerzbeurteilung durch den Patienten (Verbesserung ≥ 10)	0,9 [0,7-1,2], p=0,573	Kein Zusatznutzen belegt
Gesamtbeurteilung der Krankheitsaktivität durch den Patienten [PatGA] (Verbesserung ≥ 10)	0,9 [0,7-1,2], p=0,463	Kein Zusatznutzen belegt
Gesamtbeurteilung der Krankheitsaktivität durch den Arzt [PhysGA] (Verbesserung ≥ 10)	1,0 [0,8-1,2], p=0,667	Kein Zusatznutzen belegt
Anzahl empfindlicher/druckschmerzhafter Gelenke $\leq 1-28$ untersuchte Gelenke	1,0 [0,6-1,7], p=0,887	Kein Zusatznutzen belegt
Anzahl empfindlicher/druckschmerzhafter Gelenke ≤ 1 – 68 untersuchte Gelenke	0,9 [0,5-1,7], p=0,717	Kein Zusatznutzen belegt

Tofacitinib (XELJANZ®)

Seite 355 von 442

Anzahl geschwollener Gelenke ≤ 1 – 28 untersuchte Gelenke	1,0 [0,7-1,4], p=0,977	Kein Zusatznutzen belegt
Anzahl geschwollener Gelenke ≤ 1 – 66 untersuchte Gelenke	1,1 [0,8-1,7], p=0,574	Kein Zusatznutzen belegt
SF-36		
- Körperlicher Komponentenscore (Verbesserung ≥ 2,5)	1,0 [0,7-1,2], p=0,686	Kein Zusatznutzen belegt
- Psychischer Komponentenscore (Verbesserung ≥ 2,5)	0,8 [0,6-1,2], p=0,350	Kein Zusatznutzen belegt
- Vitalität (Verbesserung ≥ 5)	0,9 [0,7-1,3], p=0,527	Kein Zusatznutzen belegt
- Körperliche Funktion (Verbesserung ≥ 5)	1,0 [0,7-1,4], p=0,997	Kein Zusatznutzen belegt
- Körperliche Schmerzen (Verbesserung ≥ 5)	0,9 [0,7-1,2], p=0,522	Kein Zusatznutzen belegt
- Allgemeine Gesundheit (Verbesserung $\geq 5$ )	1,1 [0,7-1,7], p=0,670	Kein Zusatznutzen belegt
- Körperliche Rollenfunktion (Verbesserung ≥ 5)	1,1 [0,7-1,6], p=0,687	Kein Zusatznutzen belegt
- Emotionale Rollenfunktion (Verbesserung $\geq 5$ )	1,0 [0,7-1,6], p=0,920	Kein Zusatznutzen belegt
- Psychisches Wohlbefinden (Verbesserung≥5)	0,9 [0,6-1,4], p=0,675	Kein Zusatznutzen belegt
- Soziale Funktion (Verbesserung ≥ 5)	0,8 [0,6-1,1], p=0,253	Kein Zusatznutzen belegt
FACIT-F (Verbesserung ≥ 4)	0,9 [0,6-1,2], p=0,332	Kein Zusatznutzen belegt
HAQ-DI ≥ 0,22	1,0 [0,8-1,2], p=0,670	Kein Zusatznutzen belegt
HAQ-DI ≥ 0,3	0,8 [0,6-1,1], p=0,183	Kein Zusatznutzen belegt
<b>HAQ-DI</b> ≥ 0,5	0,8 [0,6-1,1], p=0,206	Kein Zusatznutzen belegt
WLQ		
- Indexscore (Verbesserung $\geq 5$ )	1,1 [0,5-2,7], p=0,798	Kein Zusatznutzen belegt
- Zeitmanagement (Verbesserung ≥ 13)	0,5 [0,2-0,9], p=0,030	Kein Zusatznutzen belegt
- Körperliche Anforderungen (Verbesserung ≥ 13)	0,7 [0,3-1,8], p=0,446	Kein Zusatznutzen belegt
- Mentale/Interpersonelle Anforderungen (Verbesserung ≥ 13)	1,2 [0,5-2,5], p=0,732	Kein Zusatznutzen belegt
- Ergebnisorientierte Anforderungen (Verbesserung ≥ 13)	0,7 [0,3-1,4], p=0,289	Kein Zusatznutzen belegt
MOS Schlafscore (Verbesserung ≥ 6)	0,8 [0,6-1,2], p=0,354	Kein Zusatznutzen belegt
Unerwünschte Ereignisse		

Tofacitinib (XELJANZ®)

männlich	0,5 [0,2-0,96], p=0,039	Hinweis auf geringen Zusatznutzen
weiblich	1,2 [0,99-1,5], p=0,057	Kein Zusatznutzen belegt
Schwerwiegende unerwünschte Ereignisse		Kein Zusatznutzen belegt
Alter ≤ 65 Jahre	6,2 [1,4-26,6], p=0,014	Hinweis auf geringeren Nutzen
Alter > 65 Jahre	0,3 [0,0-3,1], p=0,324	Kein Zusatznutzen belegt
Unerwünschte Ereignisse, die zum Therapieabbruch geführt hatten	1,5 [0,5-4,4], p=0,457	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen	4,7 [0,2-96,4], p=0,315	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung	1,3 [0,9-2,0], p=0,189	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster	6,6 [0,3-125,4], p=0,210	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs	0,3 [0,0-7,6], p=0,475	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs)	2,8 [0,1-68,2], p=0,523	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte	3,8 [0,4-32,9], p=0,232	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Neutropenie	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Anämie	2,8 [1,2-68,2], p=0,523	Kein Zusatznutzen belegt

Stand: 21.04.2017

Tofacitinib (XELJANZ®) Seite 357 von 442

Begründung für die Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens von Tofacitinib im Anwendungsgebiet A – erwachsene Patienten mit mittelschwerer bis schwerer RA, die bisher unzureichend nur auf MTX angesprochen haben und für die keine MTX-Unverträglichkeit vorliegt – Teilpopulation "c)-Kombi":

Insgesamt sind bei den dargestellten patientenrelevanten Endpunkten zur Morbidität von Tofacitinib nur sehr wenige statistisch signifikante Ergebnisse aufgetreten. Dies könnte angesichts einiger positiver Trends zugunsten von Tofacitinib (+ MTX) im Vergleich zur ZVT (Adalimumab + MTX) dadurch begründet sein, dass nur eine Subgruppe der Studie ORAL STANDARD für die Teilpopulation "c)-Kombi" zur Bewertung zur Verfügung steht. Hinsichtlich der Sicherheit/Verträglichkeit zeigen sich wenige positive Ergebnisse zugunsten von Tofacitinib (+ MTX) im Vergleich zur ZVT (Adalimumab + MTX), jedoch auch mehrere Trends zuungunsten von Tofacitinib. Auch hinsichtlich der Sicherheit/Verträglichkeit traten insgesamt nur sehr wenige statistisch signifikante Ergebnisse bei den dargestellten Endpunkten auf.

Da es sich bei der Evidenzbasis um eine Teilpopulation der doppelblinden Studie ORAL STANDARD handelt (39% der Patienten), die auf Studienebene und auf der Ebene der meisten dargestellten Endpunkte niedriges Verzerrungspotenzial besitzt, ist für alle nachfolgend diskutierten statistisch signifikanten Ergebnisse die Aussagesicherheit eines "Hinweises" gerechtfertigt.

Im Detail konnte kein Zusatznutzen zur Wirksamkeit von Tofacitinib belegt werden, da es zu keinem der dargestellten Wirksamkeits-Endpunkte statistisch signifikante Ergebnisse zu verzeichnen gab. Damit ist hinsichtlich der Wirksamkeit kein Zusatznutzen für Tofacitinib belegt. Ein geringerer Nutzen von Tofacitinib im Vergleich zur ZVT liegt ebenfalls nicht vor.

Hinsichtlich der Morbidität trat ein statistisch signifikanter Unterschied zuungunsten von Tofacitinib bezüglich der Responder-Rate der **Dimension "Zeitmanagement" des WLQ** (work limitations questionnaire) auf (relatives Risiko für Response (Verbesserung): 0,5, 95%-Konfidenzintervall [0,2-0,9], p=0,03). Dieser vom Patienten ausgefüllte Fragebogen (WLQ) konnte naturgemäß nur von den Studienpatienten ausgefüllt werden, die im Studienverlauf berufstätig waren. Dadurch ist die Zahl der vorhandenen Datensätze zum WLQ deutlich geringer, als es dem FAS (full analysis set) entsprochen hätte. Da jedoch die Randomisierung in der Studie nicht hinsichtlich des Faktors "Berufstätigkeit" stratifiziert worden war, ist für diesen Endpunkt insgesamt ein hohes Verzerrungspotenzial zu erwarten. Auch der hohe Anteil von fehlenden Werten (ca. 40-60% in den jeweiligen Studienarmen) führt dazu, dass eine sinnvolle Interpretation der Ergebnisse des WLQ nicht möglich ist. Deshalb ist in der Gesamteinschätzung dieser Endpunkt für die Ableitung eines Zusatznutzens nicht interpretierbar.

Zu keinem Endpunkt der gesundheitsbezogenen Lebensqualität lagen statistisch signifikante Ergebnisse vor.

Zur Domäne der Sicherheit/Verträglichkeit gab es hinsichtlich der **Häufigkeit** unerwünschter Ereignisse bei einem für die Teilpopulation "c)-Kombi" statistisch nicht signifikanten Ergebnis einen Beleg auf Effektmodifikation für den Faktor Geschlecht (Interaktions-p-Wert = 0,012). Nur für die männlichen Patienten zeigte sich ein statistisch signifikanter Effekt zugunsten von Tofacitinib in einer Reduktion der Häufigkeit unerwünschter Ereignisse versus Adalimumab (relatives Risiko 0,5 [0,2-0,96], p=0,039). Für die weiblichen Patienten ist ein Trend zuungunsten von Tofacitinib zu erkennen, der jedoch nicht statistisch signifikant ist. Insofern sind die Ergebnisse der beiden Subgruppen nicht als gleichgereichtet anzusehen, und daher wird für die **Subgruppe der männlichen Patienten** das Vorliegen eines **Hinweises auf einen geringen Zusatznutzen** von Tofacitinib gesehen.

Bei der **Häufigkeit schwerwiegender unerwünschter Ereignisse** zeigte sich ein statistisch signifikanter Effekt zuungunsten von Tofacitinib in der Subgruppe der jüngeren Patienten (≤ 65 Jahre). Da sich kein plausibles Muster bei den schwerwiegenden unerwünschten Ereignissen erkennen lässt, welches ein relevantes Schadenspotenzial von Tofacitinib im Vergleich zu bDMARD befürchten ließe und ein positives Nutzen-Risiko-Verhältnis durch die EMA bestätigt wurde, ist die Ableitung eines generell geringeren Nutzens nicht gerechtfertigt. **Ein Zusatznutzen ist somit nicht belegt.** 

Unter Tofacitinib traten in dieser Teilpopulation c) insgesamt 13 schwerwiegende unerwünschte Ereignisse auf. Diese verteilten sich auf insgesamt zehn unterschiedliche System Organ Classes (SOC) entsprechend der MedDRA Terminologie. Auf keine einzelne SOC, wie beispielsweise "Infektionen und parasitäre Erkrankungen" (*Infections and Infestations*), entfielen mehr als zwei schwerwiegende unerwünschte Ereignisse. Des Weiteren fällt bei genauerer Betrachtung der schwerwiegenden unerwünschten Ereignisse auf, dass bei nur drei schwerwiegenden unerwünschten Ereignissen der Studienarzt einen kausalen Zusammenhang zwischen der Studienmedikation und dem Ereignis für wahrscheinlich hielt. Somit lässt sich keine plausibles Muster bei den schwerwiegenden unerwünschten Ereignissen erkennen, welches ein relevantes Schadenspotenzial von Tofacitinib im Vergleich zu bDMARD befürchten ließe.

Zu keinem weiteren Endpunkt der Kategorie Sicherheit/Verträglichkeit lagen statistisch signifikante Ergebnisse vor.

In der Gesamtschau besteht die Aufgabe, einen positiven Effekt in einer Subgruppe (Männer) gegen einen negativen Effekt in einer anderen Subgruppe (jüngere Patienten) bewertend in Relation zu setzen. Da der singuläre Effekt zuungunsten von Tofacitinib nur bei berufstätigen Patienten auftrat und das Verzerrungspotenzial dieses Endpunktes als hoch eingestuft wurde, kann dieses Ergebnis ebenfalls nicht ohne weiteres auf die gesamte Teilpopulation "c)-Kombi" übertragen werden. In der Gesamtschau scheint deshalb das Fehlen eines Nachweises für einen Zusatznutzen bei gleichzeitig ebenfalls fehlendem Nachweis für einen generell geringeren Nutzen von Tofacitinib eine gerechtfertigte Bewertung der vorliegenden Ergebnisse für die Teilpopulation "c)-Kombi". Für die Teilpopulation "c)-Kombi" lässt sich somit kein Zusatznutzen ableiten.

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

Tabelle 4-192: Übersicht über Wahrscheinlichkeit und Ausmaß des Zusatznutzens aus RCT auf Endpunktebene – Population d)-Kombi

Stand: 21.04.2017

Studie: A3921064 (ORAL STANDARD)	RR [95% KI], p-Wert	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
DAS28-4(ESR) < 2,6		Kein Zusatznutzen belegt
$DAS28-4(ESR) \le 5,1$	1,5 [0,7-3,5], p=0,344	Kein Zusatznutzen belegt
DAS28-4(ESR) > 5,1	0,4 [0,2-0,9], p=0,035	Kein Zusatznutzen belegt
$DAS28-4(ESR) \le 3,2$	0,7 [0,4-1,3], p=0,250	Kein Zusatznutzen belegt
DAS28-4(CRP) < 2,6	1,0 [0,6-1,5], p=0,849	Kein Zusatznutzen belegt
$DAS28-4(CRP) \le 3,2$	1,0 [0,8-1,4], p=0,854	Kein Zusatznutzen belegt
EULAR-Response (ESR): gute Verbesserung	0,7 [0,4-1,3], p=0,250	Kein Zusatznutzen belegt
EULAR-Response (ESR): gute oder moderate Verbesserung	1,0 [0,8-1,2], p=0,828	Kein Zusatznutzen belegt
EULAR-Response (CRP): gute Verbesserung	1,1 [0,8-1,5], p=0,705	Kein Zusatznutzen belegt
EULAR-Response (CRP): gute oder moderate Verbesserung	0,9 [0,8-1,1], p=0,336	Kein Zusatznutzen belegt
ACR70	1,2 [0,7-2,1], p=0,430	Kein Zusatznutzen belegt
ACR50	1,0 [0,7-1,4], p=0,971	Kein Zusatznutzen belegt
ACR/EULAR-Remission	1,1 [0,5-2,6], p=0,763	Kein Zusatznutzen belegt
SDAI ≤ 3,3	0,9 [0,4-1,6], p=0,621	Kein Zusatznutzen belegt
SDAI ≤ 11	1,3 [0,9-1,8], p=0,190	Kein Zusatznutzen belegt
CDAI ≤ 2,8	1,0 [0,5-2,1], p=0,933	Kein Zusatznutzen belegt
CDAI ≤ 10	1,2 [0,8-1,7], p=0,376	Kein Zusatznutzen belegt
Schmerzbeurteilung durch den Patienten (Verbesserung $\geq 10$ )	0,1 [0,8-1,2], p=0,871	Kein Zusatznutzen belegt
Gesamtbeurteilung der Krankheitsaktivität durch den Patienten [PatGA] (Verbesserung ≥ 10)	1,1 [0,9-1,3], p=0,485	Kein Zusatznutzen belegt

Tofacitinib (XELJANZ®)

Seite 360 von 442

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

Gesamtbeurteilung der Krankheitsaktivität durch den Arzt [PhysGA] (Verbesserung ≥ 10)	1,0 [0,8-1,1], p=0,626	Kein Zusatznutzen belegt	
männlich	1,4 [0,9-2,0], p=0,102	Kein Zusatznutzen belegt	
weiblich	0,9 [0,7-1,1], p=0,180	Kein Zusatznutzen belegt	
Anzahl empfindlicher/druckschmerzhafter Gelenke $\leq 1-28$ untersuchte Gelenke	0,8 [0,6-1,3], p=0,433	Kein Zusatznutzen belegt	
Anzahl empfindlicher/druckschmerzhafter Gelenke $\leq 1-68$ untersuchte Gelenke	0,7 [0,4-1,2], p=0,215	Kein Zusatznutzen belegt	
Anzahl geschwollener Gelenke $\leq 1-28$ untersuchte Gelenke	1,2 [0,8-1,7], p=0,376	Kein Zusatznutzen belegt	
Anzahl geschwollener Gelenke $\leq 1-66$ untersuchte Gelenke	1,1 [0,7-1,6], p=0,655	Kein Zusatznutzen belegt	
SF-36			
- Körperlicher Komponentenscore (Verbesserung $\geq 2,5$ )	1,0 [0,8-1,2], p=0,862	Kein Zusatznutzen belegt	
- Psychischer Komponentenscore (Verbesserung $\geq 2,5$ )	1,1 [0,8-1,5], p=0,640	Kein Zusatznutzen belegt	
- Vitalität (Verbesserung $\geq 5$ )	1,1 [0,8-1,4], p=0,741	Kein Zusatznutzen belegt	
- Körperliche Funktion (Verbesserung $\geq 5$ )	0,9 [0,7-1,1], p=0,292	Kein Zusatznutzen belegt	
- Körperliche Schmerzen (Verbesserung $\geq 5$ )	0,9 [0,7-1,2], p=0,445	Kein Zusatznutzen belegt	
- Allgemeine Gesundheit (Verbesserung $\geq 5$ )		Kein Zusatznutzen belegt	
Europa	0,6 [0,4-0,95], p=0,029	Hinweis auf geringeren Nutzen	
USA/Kanada	6,1 [0,9-43,3], p=0,070	Kein Zusatznutzen belegt	
Lateinamerika	0,5 [0,2-1,8], p=0,298	Kein Zusatznutzen belegt	
Sonstige	3,5 [1,1-10,8], p=0,030	Hinweis auf geringen Zusatznutzen	
- Körperliche Rollenfunktion (Verbesserung $\geq 5$ )	1,0 [0,7-1,4], p=0,918	Kein Zusatznutzen belegt	
- Emotionale Rollenfunktion (Verbesserung $\geq 5$ )	1,1 [0,7-1,5], p=0,759	Kein Zusatznutzen belegt	
- Psychisches Wohlbefinden (Verbesserung $\geq 5$ )	(Verbesserung $\geq$ 5) 1,1 [0,8-1,5], p=0,753 Kein Zusatznutzen belegt		
- Soziale Funktion (Verbesserung ≥ 5)	Soziale Funktion (Verbesserung ≥ 5) 1,0 [0,7-1,3], p=0,826 Kein Zusatznutzen b		
FACIT-F (Verbesserung ≥ 4)	<b>T-F (Verbesserung ≥ 4)</b> 1,1 [0,8-1,4], p=0,620  Kein Zusatznutzen bele		
HAQ-DI ≥ 0,22	0,9 [0,7-1,1], p=0,304	Kein Zusatznutzen belegt	
HAQ-DI ≥ 0,3	0,9 [0,7-1,2], p=0,369	Kein Zusatznutzen belegt	
$HAQ-DI \ge 0.5$	0,9 [0,7-1,2], p=0,435	Kein Zusatznutzen belegt	

Stand: 21.04.2017

Tofacitinib (XELJANZ®) Seite 361 von 442

WLQ		
- Indexscore (Verbesserung ≥ 5)	2,9 [0,9-10,0], p=0,075	Kein Zusatznutzen belegt
- Zeitmanagement (Verbesserung ≥ 13)	1,7 [0,95-2,9], p=0,077	Kein Zusatznutzen belegt
- Körperliche Anforderungen (Verbesserung ≥ 13)	0,5 [0,2-1,3], p=0,151	Kein Zusatznutzen belegt
- Mentale/Interpersonelle Anforderungen (Verbesserung ≥ 13)	2,9 [1,2-7,1], p=0,016	Kein Zusatznutzen belegt
- Ergebnisorientierte Anforderungen (Verbesserung ≥ 13)	1,0 [0,6-1,8], p=0,931	Kein Zusatznutzen belegt
MOS Schlafscore (Verbesserung $\geq 6$ )	0,9 [0,6-1,2], p=0,417	Kein Zusatznutzen belegt
Unerwünschte Ereignisse	1,0 [0,8-1,1], p=0,721	Kein Zusatznutzen belegt
Schwerwiegende unerwünschte Ereignisse	1,3 [0,7-2,6], p=0,415	Kein Zusatznutzen belegt
Unerwünschte Ereignisse, die zum Therapieabbruch geführt hatten	0,8 [0,4-1,6], p=0,568	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Schwere Infektionen	4,0 [0,5-35,5], p=0,208	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Jedwede Infektion und parasitäre Erkrankung	0,9 [0,7-1,3], p=0,706	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Herpes Zoster	0,8 [0,2-3,3], p=0,711	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Tuberkulose	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Nicht-Melanom Hautkrebs	0,3 [0,0-8,2], p=0,503	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Gastrointestinale Perforation	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Malignität (ohne Nicht-Melanom Hautkrebs)	0,5 [0,1-5,5], p=0,574	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Leberwerte	0,5 [0,1-5,5], p=0,574	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Erhöhte Kreatininwerte im Blut	1,0 [0,1-15,9], p=0,995	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Neutropenie	3,0 [0,1-73,5], p=0,496	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Lymphopenie	nicht definiert	Kein Zusatznutzen belegt
Unerwünschte Ereignisse von besonderem Interesse: Anämie	3,0 [0,1-73,5], p=0,496	Kein Zusatznutzen belegt

Tofacitinib (XELJANZ®) Seite 362 von 442

Begründung für die Ableitung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens von Tofacitinib im Anwendungsgebiet A – erwachsene Patienten mit mittelschwerer bis schwerer RA, die bisher unzureichend auf MTX und mindestens ein weiteres csDMARD angesprochen haben und für die keine MTX-Unverträglichkeit vorliegt – Teilpopulation "d)-Kombi":

Insgesamt sind bei den dargestellten patientenrelevanten Endpunkten zur Wirksamkeit und gesundheitsbezogenen Lebensqualität von Tofacitinib nur sehr wenige statistisch signifikante Ergebnisse aufgetreten. Dies könnte dadurch begründet sein, dass nur eine Subgruppe der Studie ORAL STANDARD für die Teilpopulation "d)-Kombi" zur Bewertung zur Verfügung steht. Da deren separate Auswertung bei der Planung dieser RCT nicht absehbar war, besitzen diese Analysen keine ausreichende Power zum Nachweis patientenrelevanter Effekte. Hinsichtlich der Verträglichkeit zeigen sich sowohl einige positive Ergebnisse zugunsten von Tofacitinib (+ MTX) im Vergleich zur ZVT (Adalimumab + MTX), als auch einige Trends zuungunsten von Tofacitinib. Hinsichtlich der Verträglichkeit traten insgesamt keinerlei statistisch signifikante Ergebnisse bei den dargestellten Endpunkten auf.

Da es sich bei der Evidenzbasis um eine Teilpopulation der doppelblinden Studie ORAL STANDARD handelt (51% der Patienten), die auf Studienebene und auf der Ebene der meisten dargestellten Endpunkte niedriges Verzerrungspotenzial besitzt ist, für alle nachfolgend diskutierten statistisch signifikanten Ergebnisse die Aussagesicherheit eines "Hinweises" gerechtfertigt.

Hinsichtlich der Wirksamkeit konnte ein statistisch signifikanter Effekt hinsichtlich der **DAS28-Response** (DAS28-4(ESR) < 2,6) zuungunsten von Tofacitinib beobachtet werden. In den durchgeführten Sensitivitätsanalysen erweist sich die Aussagekraft dieses Ergebnisses jedoch nicht als robust, da in sechs der sieben Sensitivitätsanalysen mit unterschiedlichen Imputationsmethoden für fehlende Werte (siehe oben Tabelle 4-16) kein statistisch signifikanter Effekt auftritt. Lediglich bei einer reinen LOCF-Methode, die im konkreten Falle für diesen Endpunkt zu identischen Ergebnissen wie in der primären Analyse führt, ergibt sich zwangsläufig ebenfalls statistische Signifikanz. Die Schätzer für den Behandlungseffekt erweisen sich über alle Analysen hinweg jedoch als robust und weitgehend stabil.

Es wurde zudem ein Beleg auf Effektmodifikation bezüglich des Schweregrades der RA (DAS28-4(ESR)  $\leq 5,1$  vs. > 5,1) festgestellt (Interaktions-p-Wert = 0,029). Ein statistisch signifikanter Effekt zuungunsten von Tofacitinib wurde nur bei den Patienten festgestellt, die bei Baseline eine hohe Krankheitsaktivität aufwiesen (DAS28-4(ESR) > 5,1). In dieser Subgruppe betrug das relatives Risiko für Ansprechen (Response) = 0,4 [0,2-0,9], p=0,035. In der komplementären Subgruppe der Patienten ohne hohe Krankheitsaktivität gibt es einen Trend in die entgegengesetzte Richtung (zugunsten von Tofacitinib), der jedoch nicht statistisch signifikant ist. Dieser signifkante Effekt spiegelt sich jedoch nicht in anderen vergleichbaren Morbiditätsendpunkten, wie beispielsweise SDAI  $\leq 3,3$ , wider. Des Weiteren führt der hohe Anteil von fehlenden Werten (> 30%) dazu, dass eine sinnvolle Interpretation der Ergebnisse des DAS28-4(ESR) (siehe auch Abschnitt 4.3.1.3.1) generell nicht möglich ist.

Deshalb ist in der Gesamteinschätzung der Endpunkt DAS28-4(ESR) für die Ableitung eines Zusatznutzens nicht interpretierbar.

Hinsichtlich der Morbidität trat ein statistisch signifikanter Unterschied zugunsten von Tofacitinib bezüglich der Responder-Rate der **Dimension "Mentale/Interpersonelle Anforderungen" des WLQ** (work limitations questionnaire) auf (relatives Risiko für Response (Verbesserung): 2,9, 95%-Konfidenzintervall [1,2-7,1], p=0,016). Dieser vom Patienten ausgefüllte Fragebogen (WLQ) konnte naturgemäß nur von den Studienpatienten ausgefüllt werden, die im Studienverlauf berufstätig waren. Dadurch ist die Zahl der vorhandenen Datensätze zum WLQ deutlich geringer, als es dem FAS (full analysis set) entsprochen hätte. Da jedoch die Randomisierung in der Studie nicht hinsichtlich des Faktors "Berufstätigkeit" stratifiziert worden war, ist für diesen Endpunkt insgesamt ein hohes Verzerrungspotenzial zu erwarten. Auch der hohe Anteil von fehlenden Werten (ca. 30-35% in den jeweiligen Studienarmen) führt dazu, dass eine sinnvolle Interpretation der Ergebnisse des WLQ nicht möglich ist. Deshalb ist in der Gesamteinschätzung dieser Endpunkt für die Ableitung eines Zusatznutzens nicht interpretierbar.

Hinsichtlich der gesundheitsbezogenen Lebensqualität trat bei einem statistisch nicht signifikanten Unterschied bezüglich der Responder-Rate der Dimension "Allgemeine Gesundheit" des SF-36 ein Beleg auf Effektmodifikation hinsichtlich der Region der Studiendurchführung auf (Interaktions-p-Wert = 0,006). Die Effekte in jeder der vier Regionen (Europa, USA/Kanada, Lateinamerika und 'sonstige') sind dabei sehr heterogen und nur in den Regionen Europa und 'sonstige' statistisch signifikant. Bei der Region sonstige' handelt es sich um die Zusammenfassung von Patienten aus unterschiedlichen Regionen, sofern sie keiner der drei anderen Regionen zuzuordnen waren. Eine plausible Interpretation dieser heterogenen Effekte kann insgesamt nicht gegeben werden, es könnte sich auch um einen falsch positiven Effekt durch die hohe Zahl an Subgruppen-Effekttestungen handeln, da insgesamt 228 Interaktionstests ohne Kontrolle des Fehlers 1. Art für die Teilpopulation "d)-Kombi" vorgenommen wurden. Außerdem liegt in keiner der anderen sieben Einzeldimensionen und auch in beiden Komponentenscores des SF-36 kein weiterer Beleg auf Effektmodifikation bezüglich der Region vor, was ebenfalls die Plausibilität dieses Effektes fraglich erscheinen lässt. Für Europa ergibt sich jedoch ein statistisch signifikanter Effekt zuungunsten von Tofacitinib (relatives Risiko für Response (Verbesserung): 0,6, 95%-Konfidenzintervall [0,4-0,95], p=0,029). Für die zusammengefasste Region ,sonstige' ergibt sich ein statistisch signifikantes Ergebnis zugunsten von Tofacitinib (relatives Risiko für Response (Verbesserung): 3,5, 95%-Konfidenzintervall [1,1-10,8], p=0,030). Wegen der fehlenden Plausibilität werden daher der Hinweis auf einen geringeren Nutzen (Europa) und der Hinweis auf einen geringen Zusatznutzen ("sonstige") nicht interpretiert, sondern das Ergebnis für die gesamte Teilpopulation "d)-Kombi", welches nicht statistisch signifikant ist. Damit liegt für die Dimension "Allgemeine Gesundheit" des SF-36 weder ein belegter Zusatznutzen noch ein belegter geringerer Nutzen von Tofacitinib vor.

Zu keinem der Endpunkte der Kategorie Sicherheit/Verträglichkeit lagen statistisch signifikante Ergebnisse vor.

In der Gesamtschau ist ein positiver Effekt hinsichtlich der gesundheitsbezogenen Lebensqualität für alle Patienten in der Teilpopulation "d)-Kombi" einem negativen Effekt hinsichtlich der Wirksamkeit (DAS28-4(ESR)-Response) in der Subgruppe der bei Baseline schwerer erkrankten Patienten in Relation zu setzen. In der Gesamtschau scheint deshalb das Fehlen eines Nachweises für einen Zusatznutzen bei gleichzeitig ebenfalls fehlendem Nachweis für einen generell geringeren Nutzen von Tofacitinib eine gerechtfertigte Bewertung der vorliegenden Ergebnisse für die Teilpopulation "d)-Kombi". Für die Teilpopulation "d)-Kombi" lässt sich somit kein Zusatznutzen ableiten.

## 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 4-193: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Anwendungsgebiet A – Population b)-Mono	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population b)-Kombi	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population c)-Mono	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population c)-Kombi	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population d)-Mono	Kein Zusatznutzen belegt
Anwendungsgebiet A – Population d)-Kombi	Kein Zusatznutzen belegt

#### 4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

#### 4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Es wurde kein indirekter Vergleich für die Teilpopulationen c)-Kombi und d)-Kombi durchgeführt, da eine direkt vergleichende randomisierte Studie vorliegt.

Für die Teilpopulationen b)-Mono, b)-Kombi, c)-Mono und d)-Mono konnte trotz des Fehlens direkt vergleichender Evidenz aus RCT kein adjustierter indirekter Vergleich durchgeführt werden, da keine der in Frage kommenden Studien mit Tofacitinib Patienten ohne das Vorliegen ungünstiger Prognosefaktoren eingeschlossen hatte (für b)-Mono und b)-Kombi).

Des Weiteren liegen (für c)-Mono und d)-Mono) keine Daten mit ausreichender, als patientenrelevant eingestufter Therapiedauer von mindestens 6 Monaten unter einem möglichen "Brückenkomparator" (es kommt aufgrund der bekannten Tofacitinib-Studien nur Placebo dafür in Betracht) vor. Nicht-adjustierte indirekte Vergleiche für diese Teilpopulationen wurden wegen ihrer limitierten Aussagekraft als nicht zielführend für die Nutzenbewertung eingestuft und daher ebenfalls nicht 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.

Nicht relevant, da keine nicht randomisierten vergleichenden Studien oder weitere Untersuchungen vorgelegt wurden.

# 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.

Bei den in diesem Dossier dargestellten Ergebnissen der eingeschlossenen randomisierten klinischen Studie handelt es sich um valide Daten zu patientenrelevanten Endpunkten.

#### 4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005<sup>13</sup>, Molenberghs 2010<sup>14</sup>). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006<sup>15</sup>) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt ("individuelle Ebene") sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt ("Studienebene"). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006<sup>16</sup>) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle

Datenherkunft

verwendete Methodik

<sup>13</sup> Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.

\_

<sup>&</sup>lt;sup>14</sup> Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.

<sup>&</sup>lt;sup>15</sup> Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.

<sup>&</sup>lt;sup>16</sup> Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Nicht relevant, da keine Ergebnisse zu Surrogatendpunkten in diesem Dossier dargestellt werden.

#### 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).

Für die Nutzenbewertung wurden folgende Quellen der Studie ORAL STANDARD verwendet:

- Studienbericht (54)
- Zusatzanalysen für die Nutzenbewertung (38-41)
- Studienregistereinträge (42-46)
- Publikationen (55-57)

#### 4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

- 1. Gemeinsamer Bundesausschuss (G-BA). Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV Beratungsanforderung 2016-B-088 Wirkstoff: Tofacitinib. 2016.
- 2. EMA. EPAR: XELJANZ, Anhang I Zusammenfassung der Merkmale des Arzneimittels. 2017.
- 3. Wong, SS, Wilczynski, NL, Haynes, RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc. 2006;94(4):451-5.
- 4. Fransen, J, van Riel, PL. The Disease Activity Score and the EULAR response criteria. Clin Exp Rheumatol. 2005;23(5 Suppl 39):S93-9.
- 5. van Gestel, AM, Prevoo, ML, van 't Hof, MA, van Rijswijk, MH, van de Putte, LB, van Riel, PL. 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. 1996;39(1):34-40.
- 6. Aletaha, D, Landewe, R, Karonitsch, T, Bathon, J, Boers, M, Bombardier, C, et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Arthritis Rheum. 2008;59(10):1371-7.
- 7. Committee for Medicinal Products for Human Use. Guideline on clinical investigation of medicinal products 4 other than NSAIDs for treatment of rheumatoid arthritis Draft. 2015.
- 8. Singh, JA, Christensen, R, Wells, GA, Suarez-Almazor, ME, Buchbinder, R, Lopez-Olivo, MA, et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. Cochrane Database Syst Rev. 2009(4):CD007848.
- 9. Singh JA, Hossain A, Tanjong Ghogomu E, Kotb A, Christensen R, Mudano AS, et al. Biologics or tofacitinib for rheumatoid arthritis in incomplete responders tomethotrexate or other traditional diseasemodifying anti-rheumatic drugs: a systematic review and networkmeta-analysis (Review). Cochrane Database of Systematic Reviews. 2016(5).
- 10. Aletaha, D, Smolen, J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol. 2005;23(5 Suppl 39):S100-8.
- 11. Cuomo, G, Molinaro, G, La Montagna, G, Migliaresi, S, Valentini, G. [A comparison between the Simplified Disease Activity Index (SDAI) and the Disease Activity Score (DAS28) as measure of response to treatment in patients undergoing different therapeutic regimens]. Reumatismo. 2006;58(1):22-5.
- 12. Smolen, JS, Breedveld, FC, Schiff, MH, Kalden, JR, Emery, P, Eberl, G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology (Oxford). 2003;42(2):244-57.
- 13. Aletaha, D, Nell, VP, Stamm, T, Uffmann, M, Pflugbeil, S, Machold, K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther. 2005;7(4):R796-806.
- 14. Linde, L, SØRensen, JAN, ØStergaard, M, HØRslev-Petersen, KIM, Hetland, ML. Health-Related Quality of Life: Validity, Reliability, and Responsiveness of SF-36, EQ-15D,

- EQ-5D, RAQoL, and HAQ in Patients with Rheumatoid Arthritis. The Journal of Rheumatology. 2008;35(8):1528.
- 15. Egsmose, EL, Madsen, OR. Interplay between patient global assessment, pain, and fatigue and influence of other clinical disease activity measures in patients with active rheumatoid arthritis. Clinical Rheumatology. 2015;34(7):1187-94.
- 16. McHorney, CA, Ware, JE, Jr., Raczek, AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care. 1993;31(3):247-63.
- 17. Kosinski, M, Zhao, SZ, Dedhiya, S, Osterhaus, JT, Ware, JE. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis & Rheumatism. 2000;43(7):1478-87.
- 18. Talamo, J, Frater, A, Gallivan, S, Young, A. Use of the short form 36 (SF36) for health status measurement in rheumatoid arthritis. Rheumatology. 1997;36(4):463-9.
- 19. Lubeck, DP. Patient-reported outcomes and their role in the assessment of rheumatoid arthritis. Pharmacoeconomics. 2004;22(2 Suppl 1):27-38.
- 20. Strand, V, Burmester, GR, Ogale, S, Devenport, J, John, A, Emery, P. Improvements in health-related quality of life after treatment with tocilizumab in patients with rheumatoid arthritis refractory to tumour necrosis factor inhibitors: results from the 24-week randomized controlled RADIATE study. Rheumatology (Oxford). 2012;51(10):1860-9.
- 21. Cella, D, Yount, S, Sorensen, M, Chartash, E, Sengupta, N, Grober, J. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. J Rheumatol. 2005;32(5):811-9.
- 22. Bruce, B, Fries, JF. The Stanford Health Assessment Questionnaire: dimensions and practical applications. Health Qual Life Outcomes. 2003;1:20.
- 23. van Groen, MM, ten Klooster, PM, Taal, E, van de Laar, MA, Glas, CA. Application of the health assessment questionnaire disability index to various rheumatic diseases. Qual Life Res. 2010;19(9):1255-63.
- 24. Fries, JF, Spitz, P, Kraines, RG, Holman, HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23(2):137-45.
- 25. Fries, JF, Spitz, PW, Young, DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol. 1982;9(5):789-93.
- 26. Ramey, DR, Raynauld, JP, Fries, JF. The health assessment questionnaire 1992: status and review. Arthritis Care Res. 1992;5(3):119-29.
- 27. Wells, GA, Tugwell, P, Kraag, GR, Baker, PR, Groh, J, Redelmeier, DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. J Rheumatol. 1993;20(3):557-60.
- 28. Lerner, D, Amick, BC, 3rd, Rogers, WH, Malspeis, S, Bungay, K, Cynn, D. The Work Limitations Questionnaire. Med Care. 2001;39(1):72-85.
- 29. Walker, N, Michaud, K, Wolfe, F. Work limitations among working persons with rheumatoid arthritis: results, reliability, and validity of the work limitations questionnaire in 836 patients. J Rheumatol. 2005;32(6):1006-12.
- 30. Roy, JS, MacDermid, JC, Amick, BC, 3rd, Shannon, HS, McMurtry, R, Roth, JH, et al. Validity and responsiveness of presenteeism scales in chronic work-related upper-extremity disorders. Phys Ther. 2011;91(2):254-66.
- 31. Detert, J, Dziurla, R, Hoff, P, Gaber, T, Klaus, P, Bastian, H, et al. Effects of treatment with etanercept versus methotrexate on sleep quality, fatigue and selected immune parameters in patients with active rheumatoid arthritis. Clin Exp Rheumatol. 2016;34(5):848-56.

- 32. Cappuccio, FP, D'Elia, L, Strazzullo, P, Miller, MA. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep. 2010;33(5):585-92.
- 33. Stewart, AL, Hays, RD, Wells, KB, Rogers, WH, Spritzer, KL, Greenfield, S. Longterm functioning and well-being outcomes associated with physical activity and exercise in patients with chronic conditions in the Medical Outcomes Study. J Clin Epidemiol. 1994;47(7):719-30.
- 34. Wells, G, Li, T, Tugwell, P. Investigation into the impact of abatacept on sleep quality in patients with rheumatoid arthritis, and the validity of the MOS-Sleep questionnaire Sleep Disturbance Scale. Ann Rheum Dis. 2010;69(10):1768-73.
- 35. Pope, J, Bingham, CO, 3rd, Fleischmann, RM, Dougados, M, Massarotti, EM, Wollenhaupt, J, et al. Impact of certolizumab pegol on patient-reported outcomes in rheumatoid arthritis and correlation with clinical measures of disease activity. Arthritis Res Ther. 2015;17:343.
- 36. Wells, G, Li, T, Maxwell, L, MacLean, R, Tugwell, P. Determining the minimal clinically important differences in activity, fatigue, and sleep quality in patients with rheumatoid arthritis. J Rheumatol. 2007;34(2):280-9.
- 37. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden, Version 4.2 vom 22.04.2015 2015 [cited 2016 May 06]. Available from: https://www.iqwig.de/download/IQWiG\_Methoden\_Version\_4-2.pdf.
- 38. Pfizer. Zusatzanalysen der Studie A3921064 (ORAL STANDARD). Sensitivitätsanalysen. 2017.
- 39. Pfizer. Zusatzanalysen der Studie A3921064 (ORAL STANDARD). Demographie. 2017.
- 40. Pfizer. Zusatzanalysen der Studie A3921064 (ORAL STANDARD). Wirksamkeit. 2017.
- 41. Pfizer. Zusatzanalysen der Studie A3921064 (ORAL STANDARD). Sicherheit/Verträglichkeit. 2017.
- 42. ClinicalTrials.gov. 2013. A Phase 3 Study Comparing 2 Doses Of CP-690,550 And The Active Comparator, Humira (Adalimumab) Vs. Placebo For Treatment Of Rheumatoid Arthritis; Abrufbar unter: <a href="https://ClinicalTrials.gov/show/NCT00853385">https://ClinicalTrials.gov/show/NCT00853385</a> [Zugriff am: 03.02.2017].
- 43. ICTRP. 2015. Phase 3 Randomized, Double-Blind, Active Comparator, Placebo-Controlled Study Of The Efficacy And Safety Of 2 Doses Of CP690,550 In Patients With Active Rheumatoid Arthritis On Background Methotrexate; Abrufbar unter: <a href="http://apps.who.int/trialsearch/">http://apps.who.int/trialsearch/</a> [Zugriff am: 03.02.2017].
- 44. ICTRP. 2015. PHASE 3 RANDOMIZED, DOUBLE-BLIND, ACTIVE COMPARATOR, PLACEBO-CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP 690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND METHOTREXATE; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].
- 45. EU Clinical Trials Register. 2009. PHASE 3 RANDOMIZED, DOUBLE-BLIND, ACTIVE COMPARATOR, PLACEBO-CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP 690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND METHOTREXATE; Abrufbar unter: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract\_number:2008-008338-35">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract\_number:2008-008338-35</a> [Zugriff am: 06.02.2017].

- 46. PharmNet.Bund. 2008. PHASE 3 RANDOMIZED, DOUBLE-BLIND, ACTIVE COMPARATOR, PLACEBO-CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP 690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND METHOTREXATE; [Zugriff am: 06.02.2017].
- 47. ClinicalTrials.gov. 2016. Safety Study Of Tofacitinib Versus Tumor Necrosis Factor (TNF) Inhibitor In Subjects With Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02092467 [Zugriff am: 07.06.2016].
- 48. ICTRP. 2014. PHASE 3B/4 RANDOMIZED SAFETY ENDPOINT STUDY OF 2 DOSES OF TOFACITINIB IN COMPARISON TO A TUMOR NECROSIS FACTOR (TNF) INHIBITOR IN SUBJECTS WITH RHEUMATOID ARTHRITIS; Abrufbar unter: <a href="http://apps.who.int/trialsearch/">http://apps.who.int/trialsearch/</a> [Zugriff am: 07.06.2016].
- 49. EU Clinical Trials Register. 2014. PHASE 3B/4 RANDOMIZED SAFETY ENDPOINT STUDY OF 2 DOSES OF TOFACITINIB IN COMPARISON TO A TUMOR NECROSIS FACTOR (TNF) INHIBITOR IN SUBJECTS WITH RHEUMATOID ARTHRITIS; Abrufbar unter: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search/query=eudract\_number:2013-003177-99">https://www.clinicaltrialsregister.eu/ctr-search/search/query=eudract\_number:2013-003177-99</a> [Zugriff am: 07.06.2016].
- 50. ClinicalTrials.gov. 2016. An Efficacy And Safety Study Evaluating Tofacitinib With And Without Methotrexate Compared To Adalimumab With Methotrexate (ORAL STRATEGY); Abrufbar unter: <a href="https://ClinicalTrials.gov/show/NCT02187055">https://ClinicalTrials.gov/show/NCT02187055</a> [Zugriff am: 07.06.2016].
- 51. ICTRP. 2016. A Phase 3b/4 Randomized Double Blind Study Of 5 Mg Of Tofacitinib With And Without Methotrexate In Comparison To Adalimumab With Methotrexate In Subjects With Moderately To Severely Active Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 07.06.2016].
- 52. EU Clinical Trials Register. 2014. A PHASE 3b/4 RANDOMIZED DOUBLE BLIND STUDY OF 5 MG OF TOFACITINIB WITH AND WITHOUT METHOTREXATE IN COMPARISON TO ADALIMUMAB WITH METHOTREXATE IN SUBJECTS WITH MODERATELY TO SEVERELY ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract\_number:2014-000358-13">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract\_number:2014-000358-13</a> [Zugriff am: 07.06.2016].
- 53. PharmNet.Bund. 2014. A PHASE 3b/4 RANDOMIZED DOUBLE BLIND STUDY OF 5 MG OF TOFACITINIB WITH AND WITHOUT METHOTREXATE IN COMPARISON TO ADALIMUMAB WITH METHOTREXATE IN SUBJECTS WITH MODERATELY TO SEVERELY ACTIVE RHEUMATOID ARTHRITIS; [Zugriff am: 07.06.2016].
- 54. Pfizer. Complete Study Report of Phase 3 Randomized, Double-Blind, Active Comparator, Placebo-Controlled Study of the Efficacy and Safety of 2 Doses of CP-690,550 in Patients with Active Rheumatoid Arthritis on Background Methotrexate. Study A3921064 (ORAL STANDARD). Version: 2014 October 07. 2014.
- 55. Strand, V, Van Vollenhoven, RF, Lee, EB, Fleischmann, R, Zwillich, SH, Gruben, D, et al. Tofacitinib or adalimumab versus placebo: patient-reported outcomes from a phase 3 study of active rheumatoid arthritis. Rheumatology (Oxford). 2016;55(6):1031-41.
- 56. Van Vollenhoven, R, Cohen, S, Mendelsohn, A, Bananis, E, Fan, H, Takiya, L, et al. Efficacy of adalimumab and tofacitinib in rheumatoid arthritis: Post-HOC analyses from a phase 3 study. Annals of the rheumatic diseases. 2016;75:1042.
- 57. van Vollenhoven, RF, Fleischmann, R, Cohen, S, Lee, EB, Garcia Meijide, JA, Wagner, S, et al. Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. NEJM. 2012;367(6):508-19.

#### 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):

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

Daten	enbankname EMBASE		
Sucho	noberfläche Ovid		
Datun	n der Suche	08.12.2010	
Zeitse	egment	1980 to 2010 week 50	
Suchf	ilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Que Strategy minimizing difference between sensitivity and specificity	lle <sup>17</sup> ] –
#	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		32
6	(105816-04-4 or 135062-02-1).rn. 2854		2854
7	or/1-6 3467		3467
8	Diabetes mellitus/ 224164		224164
9	Non Insulin depe	ndent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti. 379777		379777
11	or/8-10 454517		454517
12	(random* or double-blind*).tw. 650136		650136
13	placebo*.mp. 243550		243550
14	or/12-13 773621		773621
15	and/7,11,14 719		719

<sup>&</sup>lt;sup>17</sup> Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. 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

Daten	nbankname EMBASE Alert EMBASE		
Sucho	berfläche	DIMDI	
Datun	n der Suche	07.06.2016	
Zeitse	gment	keine zeitliche Einschränkung: ab 2008 (EMBASE Alert) bzw. 1974 (EMBASE)	
Suchfi	Filter RCT nach: Wong SS et al. Comparison of top-performing search strategic for detecting clinically sound treatment studies and systematic reviews MEDLINE and EMBASE. J Med Libr Assoc. 2006; 94: 451-5. Table 1: "Strate minimizing difference between sensitivity and specificity". Angepasst and Kommandosprache von DIMDI.		tic reviews in ole 1: "Strategy
#	Suchbegriffe		Ergebnis
1	EA08; EM74		27910389
2	RHEUMATOID,	? ARTHRITIS.	176480
3	CT D RHEUMATOID ARTHRITIS 16182		161828
4	2 OR 3		178447
5	CT=TOFACITINIB 1593		1593
6	TOFACITINIB OR TASOCITINIB 1460		1460
7		CP690550 OR CP 690550 OR CP-690,550 OR CP690,550 OR CP 690 550 OR CP-690 550 OR CP690 550	527
8	5 OR 6 OR 7 1707		1707
9	RANDOM?/TI OR RANDOM?/AB 1085924		1085924
10	PLACEBO? 372707		372707
11	DOUBLE-BLIND?/TI OR DOUBLE-BLIND?/AB 169084		169084
12	9 OR 10 OR 11 131347		1313478
13	4 AND 8 AND 1	2	399
14	check duplicates:	unique in s=13; state=update	299

Datenbankname	EMBASE Alert EMBASE	
Suchoberfläche	ProQuest Dialog	
Datum der Suche	02.02.2017 (Update-Recherche)	
Zeitsegment	04.06.2016 bis 02.02.2017	
Suchfilter  Filter RCT nach: Wong SS et al. Comparison of top-performing search strat for detecting clinically sound treatment studies and systematic review MEDLINE and EMBASE. J Med Libr Assoc. 2006; 94: 451-5. Table 1: "Straminimizing difference between sensitivity and specificity".		c reviews in
# Suchbegriffe		Ergebnis

1	(rheumatoid NEAR/5 arthritis)	201064
2	(EMB.EXACT.EXPLODE("rheumatoid arthritis"))	190200
3	S1 OR S2	203319
4	EMB.EXACT.EXPLODE("tofacitinib")	1707
5	(tofacitinib OR tasocitinib)	1804
6	("CP-690550" OR CP690550 OR "CP 690550" OR "CP-690,550" OR "CP690,550" OR "CP 690 550" OR "CP-690 550" OR "CP690 550")	537
7	S4 OR S4 OR S6	2030
8	pub(random*) OR ab(random*) OR ti(random*)	1160039
9	pub(double NEAR/1 blind*) OR ab(double NEAR/1 blind*) OR ti(double NEAR/1 blind*)	178380
10	placebo*	392710
11	EMB.EXACT.EXPLODE("placebo")	328702
12	S8 OR S9 OR S10 OR S11	1401090
13	(S3 AND S7 AND S12) and (pd(20160604-20170202))	54

Date	enbankname	MEDLINE	
Such	oberfläche DIMDI		
Datu	ım der Suche	07.06.2016	
Zeits	segment	keine zeitliche Einschränkung: ab 1960	
Such	nfilter	Filter RCT nach: Wong SS et al. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc. 2006; 94: 451-5. Table 1: "Strategy minimizing difference between sensitivity and specificity". Angepasst an die Kommandosprache von DIMDI.	
#	Suchbegriffe		Ergebnis
1	ME60		26111876
2	RHEUMATOID	, ? ARTHRITIS.	115322
3	CT D ARTHRIT	TIS, RHEUMATOID	98850
4	2 OR 3 126430		126430
5	CT=TOFACITINIB (		0
6	TOFACITINIB OR TASOCITINIB 432		432
7	CP-690550 OR CP690550 OR CP 690550 OR CP-690,550 OR CP690,550 OR CP 690,550 OR CP 690,550 OR CP690 550 OR CP690 550		100
8	5 OR 6 OR 7 453		453
9	DT=RANDOMI	ZED CONTROLLED TRIAL	412461
10	RANDOMIZED OR RANDOMISED 672930		672930
11	PLACEBO 176591		176591
12	9 OR 10 OR 11 723852		723852
13	CT D ANIMALS NOT CT D HUMANS 4218003		4218003
14	12 NOT 13		690639
15	4 AND 8 AND 1	4	59
16	check duplicates: unique in s=15; state=update 37		37

Dater	nbankname	MEDLINE	
	hoberfläche ProQuest Dialog		
	m der Suche	02.02.2017 (Update-Recherche)	
	egment	04.06.2016 bis 02.02.2017	
	Filter RCT nach: Wong SS et al. Comparison of top-performing search strategi for detecting clinically sound treatment studies and systematic reviews MEDLINE and EMBASE. J Med Libr Assoc. 2006; 94: 451-5. Table 1: "Strates minimizing difference between sensitivity and specificity".		tic reviews in
#	Suchbegriffe		Ergebnis
1	(rheumatoid NEA	AR/5 arthritis)	118483
2	(MESH.EXACT.	EXPLODE("Arthritis, Rheumatoid"))	100806
3	S1 OR S2 129839		129839
4	tofacitinib OR tasocitinib 545		545
5	"CP-690550" OR CP690550 OR "CP 690550" OR "CP-690,550" OR "CP690,550" OR "CP690,550" OR "CP690 550" OR "CP690 550"		106
6	S4 OR S5 568		568
7	DTYPE("random	nized controlled trial")	427660
8	randomized OR randomised 706062		706062
9	placebo 196847		196847
10	S7 OR S8 OR S9 768999		768999
11	MESH.EXACT.EXPLODE("ANIMALS") NOT MESH.EXACT.EXPLODE("HUMANS")  4293502		4293502
12	S10 NOT S11 733988		733988
13	(S3 AND S6 AND S12) and (pd(20160604-20170202)) 24		

Datenb	abankname Cochrane Central Register of Controlled Trials		
Suchob	perfläche	Cochrane Library	
Datum	um der Suche 02.02.2017		
Zeitseg	keine zeitliche Einschränkung: ab 1993		
Suchfil	lter	kein Suchfilter verwendet	
#	Suchbegriffe		Ergebnis
1	TOFACITINIB or TASOCITINIB:ti,ab,kw in Trials (word variations have been searched)  174		174
2	"CP-690550" or CP690550 or "CP 690550" or "CP-690,550" or "CP690,550" or "CP 690,550" or "CP 690,550" or "CP-690 550" or "CP690 550": ti,ab,kw in Trials (word variations have been searched)		50
3	1 OR 2		181
4	rheumatoid near arthritis:ti,ab,kw in Trials (word variations have been searched)		7778
5	MeSH descriptor: [Arthritis, Rheumatoid] explode all trees		4606
6	4 OR 5 in Trials		7937
7	3 AND 6 in Trial	s	91

#### Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht relevant, da keine indirekten Vergleiche durchgeführt wurden.

#### Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant, da keine nicht randomisierten vergleichenden Studien in Ergebnisdarstellung dieses Dossiers einbezogen wurden.

#### Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht relevant, da keine weiteren Untersuchungen dargestellt wurden.

#### Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. clinicaltrials.gov), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. http://www.clinicaltrials.gov), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase II" OR "Phase IV") [PHASE]
Treffer	23

### Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	02.02.2017
Suchstrategie	Advanced Search rheumatoid arthritis [condition]   TOFACITINIB OR TASOCITINIB OR CP- 690,550 OR CP-690550 OR CP690550 [intervention]
Treffer	43

Studienregister	EU Clinical Trials Register
Internetadresse	www.clinicaltrialsregister.eu
Datum der Suche	02.02.2017
Suchstrategie	TOFACITINIB OR TASOCITINIB OR CP-690,550 OR CP-690550 OR CP690550
Treffer	40

Studienregister	WHO – ICTRP-Portal
Internetadresse	http://apps.who.int/trialsearch/Default.aspx
Datum der Suche	02.02.2017
Suchstrategie	Basic Search: TOFACITINIB OR TASOCITINIB OR CP-690,550 OR CP-690550 OR CP690550
Treffer	125

Studienregister	PharmnetBund
Internetadresse	http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm
Datum der Suche	02.02.2017
Suchstrategie	Tofacitinib [Active substance] OR Tasocitinib [Active substance] OR CP-690,550 [Active substance] OR CP-690550 [Active substance] OR CP-690550 [Active substance] OR CP-690,550 [Textfelder] OR CP-690550 [Textfelder] OR CP690550 [Textfelder]
Treffer	29

#### Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht relevant, da keine indirekten Vergleiche durchgeführt wurden.

#### Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant, da keine nicht randomisierten vergleichenden Studien in die Ergebnisdarstellung dieses Dossiers einbezogen wurden.

#### Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht relevant, da keine weiteren Untersuchungen dargestellt wurden.

# Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der / den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nummer	Publikation	Ausschluss- kriterium
1	Corrections: Tofacitinib or adalimumab versus placebo in rheumatoid arthritis NEJM. 2013;369(3):293.	A7
2	Ahadieh, S, Checchio, T, Tensfeldt, T, French, J, Geier, J, Riese, R, et al. Meta- analysis of malignancies, serious infections, and serious adverse events with tofacitinib or biologic treatment in rheumatoid arthritis clinical trials. Journal of Pharmacokinetics and Pharmacodynamics. 2013;40(1):93-4.	A7
3	Al-Maini, M, Li, ZG, Liu, Y, Al Emadi, S, Zwillich, SH, Bradley, JD, et al. Tofacitinib, an oral Janus kinase inhibitor, in the treatment of rheumatoid arthritis: Summary of five Phase 3 studies. International journal of rheumatic diseases. 2012;15:67.	A7
4	Alten, R, Strand, V, Fleischmann, R, Koncz, T, Zwillich, S, Bradley, J, et al. Effects of tofacitinib monotherapy versus methotrexate on patient-reported outcomes in the 2-year phase 3 oral start trial in methotrexate-naive patients with rheumatoid arthritis. Ann Rheum Dis. 2014;73.	A1
5	An, Y, Li, Z, Wu, Q, Kwok, K, Wang, L. Efficacy and safety of tofacitinib in chinese patients with active rheumatoid arthritis: Subgroup analysis from a phase 3 study of tofacitinib in combination with nonbiologic disease-modifying antirheumatic drugs. Ann Rheum Dis. 2015;74:1071.	A7
6	Berhan, A. Efficacy, safety and tolerability of tofacitinib in patients with an inadequate response to disease modifying anti-rheumatic drugs: a meta-analysis of randomized double-blind controlled studies. BMC musculoskeletal disorders. 2013;14:332.	A7
7	Bird, P, Bensen, W, El-Zorkany, B, Kaine, J, Manapat-Reyes, BH, Pascual-Ramos, V, et al. Comprehensive summary of the efficacy and safety of tofacitinib 5mg twice daily in patients with rheumatoid arthritis and an inadequate response to disease-modifying antirheumatic drugs. Arthritis Rheumatol. 2014;66:197-8.	A7
8	Boyce, EG, Vyas, D, Rogan, EL, Valle-Oseguera, CS, O'Dell, KM. Impact of tofacitinib on patient outcomes in rheumatoid arthritis - review of clinical studies. Patient related outcome measures. 2016;7:1-12.	A7
9	Buckley, F, Best, JH, Dejonckheere, F, Finckh, A, Huizinga, T, Jansen, JP, et al. Comparative efficacy of novel DMARDs as monotherapy and in combination with methotrexate in rheumatoid arthritis patients with an inadequate response to traditional DMARDs-a network meta-analysis. Value in Health. 2014;17(3):41-2.	A7

10	Buckley, F, Finckh, A, Huizinga, TW, Dejonckheere, F, Jansen, JP. Comparative Efficacy of Novel DMARDs as Monotherapy and in Combination with Methotrexate in Rheumatoid Arthritis Patients with Inadequate Response to Conventional DMARDs: A Network Meta-Analysis. Journal of managed care & specialty pharmacy. 2015;21(5):409-23.	A7
11	Buckley, F, Finckh, A, Huizinga, TWJ, Dejonckheere, F, Jansen, JP. Comparative efficacy of novel disease-modifying antirheumatic drugs as monotherapy and in combination with methotrexate in rheumatoid arthritis patients with an inadequate response to traditional disease-modifying antirheumatic drugs: A network meta-analysis. Arthritis Rheum. 2013;65:201-2.	A7
12	Buckley, F, Finckh, A, Huizinga, TWJ, Dejonckheere, F, Jansen, JP. Comparative efficacy of novel diseasemodifying antirheumatic drugs as monotherapy and in combination with methotrexate in rheumatoid arthritis patients with an inadequate response to traditional dmards: A network metaanalysis. Rheumatology (United Kingdom). 2014;53:88-9.	A7
13	Burmester, G, Blanco, R, Rubbert-Roth, A, Hendrikx, T, Kwok, K, Bradley, J, et al. Short-and long-term efficacy of tofacitinib, an oral janus kinase inhibitor, in the treatment of patients with rheumatoid arthritis and an inadequate response to TNF inhibitors: Analyses of pooled phase 2, phase 3, and long-term extension studies. Arthritis Rheum. 2012;64(12):4172.	A7
14	Burmester, G, Charles-Schoemann, C, Isaacs, J, Hendrikx, T, Kwok, K, Zwillich, SH, et al. Tofacitinib, an oral janus kinase inhibitor: Safety comparison in patients with rheumatoid arthritis and an inadequate response to nonbiologic or biologic disease-modifying anti-rheumatic drugs. Ann Rheum Dis. 2013;72.	A7
15	Burmester, GR, Blanco, R, Charles-Schoeman, C, Wollenhaupt, J, Zerbini, C, Benda, B, et al. Tofacitinib (CP-690,550) in combination with methotrexate in patients with active rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitors: a randomised phase 3 trial. Lancet (London, England). 2013;381(9865):451-60.	A3
16	Burmester, GR, Blanco, R, Charles-Schoeman, C, Wollenhaupt, J, Zerbini, CAF, Benda, B, et al. Tofacitinib (CP-690,550), an oral janus kinase inhibitor, in combination with methotrexate, in patients with active rheumatoid arthritis with an inadequate response to tumor necrosis factor-inhibitors: A 6-month phase 3 study. Arthritis Rheum. 2011;63(10 SUPPL. 1).	A7
17	Burmester, GR, Van Der Heijde, D, Strand, V, Zerbini, CAF, Connell, CA, Mebus, CA, et al. Effects of tofacitinib on patient-reported outcomes in patients with active rheumatoid arthritis receiving stable-dose methotrexate: Results of two phase 3 studies. Arthritis Rheum. 2012;64:549.	A7
18	Chang, S, Sawyer, L, Dejonckheere, F. Indirect comparison of tocilizumab and tofacitinib in patients with rheumatoid arthritis. Arthritis Rheumatol. 2014;66:1093.	A7
19	Charles-Schoeman, C, Burmester, G, Nash, P, Zerbini, CA, Soma, K, Kwok, K, et al. Efficacy and safety of tofacitinib following inadequate response to conventional synthetic or biological disease-modifying antirheumatic drugs. Ann Rheum Dis. 2015.	A7
20	Charles-Schoeman, C, Burmester, G, Nash, P, Zerbini, CAF, Anway, S, Kwok, K, et al. Efficacy and safety of tofacitinib following inadequate response to nonbiologic DMARD or biologic DMARD. Arthritis Rheumatol. 2014;66:212-3.	A7
21	Charles-Schoeman, C, Burmester, GR, Koenig, A, Maurey, J, Naik, S, Bourret, J, et al. Efficacy and safety of tofacitinib in patients with rheumatoid arthritis that had inadequate response to non-biologic DMARDs or one or more biologic DMARDs including tumor necrosis inhibitors: Pooled analysis from Phase 2 and 3 studies. Pharmacotherapy. 2014;34(10):225.	A7

22	Charles-Schoeman, C, Wicker, P, Gonzalez-Gay, MA, Wood, SP, Boy, MG, Geier, J, et al. Cardiovascular safety findings in rheumatoid arthritis patients treated with tofacitinib, a novel, oral janus kinase inhibitor. Arthritis Rheum. 2013;65:191.	A7
23	Charles-Schoeman, C, Wicker, P, Sechtem, U, Gonzalez-Gay, MA, Wood, S, Boy, M, et al. Cardiovascular safety findings in rheumatoid arthritis patients treated with tofacitinib (CP-690,550), a novel, oral jak inhibitor. Ann Rheum Dis. 2013;71.	A7
24	Cohen, S, Krishnaswami, S, Benda, B, Riese, R, Boy, MG, Gruben, D, et al. Tofacitinib, an oral janus kinase inhibitor: Analyses of efficacy and safety of 10 versus 5mg twice daily in a pooled phase 3 and long-term extension rheumatoid arthritis population. Arthritis Rheum. 2012;64:1048-9.	A7
25	Cohen, S, Radominski, SC, Asavatanabodee, P, Wood, SP, Soma, K, Nduaka, CI, et al. Tofacitinib, an oral Janus kinase inhibitor: Analysis of infections and all-cause mortality across Phase 3 and long-term extension studies in patients with rheumatoid arthritis. International journal of rheumatic diseases. 2012;15:68.	A7
26	Cohen, S, Tanaka, Y, Mariette, X, Curtis, J, Kwok, K, Lee, E, et al. Integrated safety analysis of tofacitinib in RA clinical trials with a cumulative exposure of 12,664 patient-years. Ann Rheum Dis. 2014;73.	A7
27	Cohen, S, Tanaka, Y, Marriette, X, Curtis, JR, Kwok, K, Lee, EB, et al. Integrated safety analysis of tofacitinib in rheumatoid arthritis clinical trials with a cumulative exposure of 12,664 patient-years. Pharmacotherapy. 2014;34(10):226.	A7
28	Cohen, SB, Fleischmann, RM, Kremer, JM, Koenig, A, Kwok, K, Wang, L, et al. Efficacy and safety analyses of tofacitinib from pooled phase 2, phase 3 and long-term extension rheumatoid arthritis studies: Us compared with non-us populations. Arthritis Rheum. 2013;65:192-3.	A7
29	Cohen, SB, Koenig, A, Wang, L, Kwok, K, Mebus, CA, Riese, R, et al. Efficacy and safety of tofacitinib in US and non-US rheumatoid arthritis patients: pooled analyses of phase II and III. Clinical and experimental rheumatology. 2016;34(1):32-6.	A7
30	Cohen, SS, Radominski, SC, Asavatanabodee, P, Wood, SP, Soma, K, Nduaka, CI, et al. Tofacitinib (CP-690,550), an oral janus kinase inhibitor: Analysis of infections and all-cause mortality across phase 3 and long-term extension studies in patients with rheumatoid arthritis. Arthritis Rheum. 2011;63(10).	A7
31	Coombs, JH, Bloom, BJ, Breedveld, FC, Fletcher, MP, Gruben, D, Kremer, JM, et al. Improved pain, physical functioning and health status in patients with rheumatoid arthritis treated with CP-690,550, an orally active Janus kinase (JAK) inhibitor: results from a randomised, double-blind, placebo-controlled trial. Ann Rheum Dis. 2010;69(2):413-6.	A7
32	Curtis, JR, Lee, EB, Kaplan, IV, Kwok, K, Geier, J, Benda, B, et al. Tofacitinib, an oral Janus kinase inhibitor: analysis of malignancies across the rheumatoid arthritis clinical development programme. Ann Rheum Dis. 2016;75(5):831-41.	A7
33	Curtis, JR, Schulze-Koops, H, Takiya, L, Mebus, C, Terri, K, Biswas, P, et al. Efficacy and safety of tofacitinib in older and younger patients with rheumatoid arthritis. Pharmacotherapy. 2014;34(10):225.	A7
34	Curtis, JR, Schulze-Koops, H, Takiya, L, Mebus, CA, Terry, K, Chew, R, et al. Efficacy and safety of tofacitinib in older and younger patients with rheumatoid arthritis. Arthritis Rheum. 2013;65:994-5.	A7
35	Cutolo, M. The kinase inhibitor tofacitinib in patients with rheumatoid arthritis: latest findings and clinical potential. Therapeutic advances in musculoskeletal disease. 2013;5(1):3-11.	A7
36	Dikranian, A, Soma, K, Riese, R, Gruben, D, Jones, TV. Tolerability and non-serious adverse events in rheumatoid arthritis patients treated with tofacitinib as monotherapy or in combination therapy. Arthritis Rheum. 2013;65:192.	A7

37	Fleischmann, R, Cutolo, M, Genovese, MC, Lee, EB, Kanik, KS, Sadis, S, et al. Phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) or adalimumab monotherapy versus placebo in patients with active rheumatoid arthritis with an inadequate response to disease-modifying antirheumatic drugs. Arthritis Rheum. 2012;64(3):617-29.	A5
38	Fleischmann, R, Kremer, J, Cush, J, Schulze-Koops, H, Connell, CA, Bradley, JD, et al. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. NEJM. 2012;367(6):495-507.	A3
39	Fleischmann, R, Mease, PJ, Schwartzman, S, Hwang, LJ, Patel, A, Soma, K, et al. Effect of methotrexate dose on the efficacy of tofacitinib: Treatment outcomes from a phase 3 clinical trial of patients with rheumatoid arthritis. Arthritis Rheumatol. 2015;67.	A7
40	Fleischmann, R, Strand, V, Wilkinson, B, Kwok, K, Bananis, E. Relationship between different clinical measurements and patient-reported outcomes: Results from a phase 3 study of tofacitinib or methotrexate in methotrexate-naïve patients with rheumatoid arthritis. Ann Rheum Dis. 2015;74:259-60.	A1
41	Fleischmann, R, Strand, V, Wilkinson, B, Kwok, K, Bananis, E. Relationship between clinical and patient-reported outcomes in a phase 3 trial of tofacitinib or MTX in MTX-naïve patients with rheumatoid arthritis. RMD open. 2016;2(1):000232.	A1
42	Fleischmann, R, Strand, V, Wilkinson, B, Kwok, K, Bananis, E, Fleischmann, R, et al. Relationship between different clinical measurements and patient-reported outcomes. Arthritis Rheumatol. 2014;66:1086.	A1
43	Fleischmann, RM, Genovese, MC, Gruben, D, Kanik, KS, Wallenstein, GV, Wilkinson, B, et al. Safety and efficacy after 24 week (Wk) dosing of the oral JAK Inhibitor CP-690,550 (CP) as monotherapy in patients (pts) with active rheumatoid arthritis (RA). Arthritis Rheum. 2009;60:1924.	A7
44	Fleischmann, RM, Krishnaswami, S, Gruben, D, Zwillich, SH, Mebus, C, Bradley, JD. Improvements in disease activity score by baseline status: Pooled analysis of five phase 3 studies with tofacitinib (CP-690,550) in patients with active rheumatoid arthritis. Ann Rheum Dis. 2013;71.	A7
45	Fleischmann, RM, Lee, EB, Hall, S, Wilkinson, BE, Bradley, JD, Gruben, D, et al. Tofacitinib monotherapy versus methotrexate in methotrexate-naive patients with rheumatoid arthritis: Radiographic, clinical and functional comparison. Arthritis Rheum. 2013;65(12):3323.	A1
46	Gaujoux-Viala, C, Nam, J, Ramiro, S, Landewé, R, Buch, MH, Smolen, JS, et al. Efficacy of conventional synthetic disease-modifying antirheumatic drugs, glucocorticoids and tofacitinib: a systematic literature review informing the 2013 update of the EULAR recommendations for management of rheumatoid arthritis. Ann Rheum Dis. 2014;73(3):510-5.	A7
47	Haraoui, B, Fallon, L, Riese, R, Kwok, K, Valdez, H. Tofacitinib, an oral Janus Kinase inhibitor: Effects of baseline glucocorticoid use on selected infections in pooled phase 3 and longterm extension rheumatoid arthritis study populations. Journal of Rheumatology. 2013;40(6):961.	A7
48	Hazlewood, GS, Barnabe, C, Tomlinson, G, Marshall, D, Devoe, D, Bombardier, C. Methotrexate monotherapy and methotrexate combination therapy with traditional and biologic disease modifying antirheumatic drugs for rheumatoid arthritis: abridged Cochrane systematic review and network meta-analysis. BMJ (Clinical research ed). 2016;353:1777.	A7

49	Hazlewood, GS, Barnabe, C, Tomlinson, GA, Marshall, D, Devoe, D, Bombardier, C, et al. Methotrexate monotherapy and methotrexate combination therapy with traditional and biologic DMARDs for rheumatoid arthritis: A cochrane systematic review and network meta-analysis. Arthritis Rheumatol. 2015;67.	A7
50	He, Y, Wong, A, Chan, E, Lau, W, Man, K, Chui, C, et al. Safety of tofacitinib in the treatment of rheumatoid arthritis: A systematic review and meta-analysis. Drug Safety. 2013;36(9):852-3.	A7
51	He, Y, Wong, AY, Chan, EW, Lau, WC, Man, KK, Chui, CS, et al. Efficacy and safety of tofacitinib in the treatment of rheumatoid arthritis: a systematic review and meta-analysis. BMC musculoskeletal disorders. 2013;14:298.	A7
52	Ingham, MP, Song, P, Cartier, S, Lawson-Remer, K, Murray, E. Assessing janus kinase inhibitor's place in therapy in established rheumatoid arthritis patients-from a simplified indirect comparison versus tumor necrosis factor inhibitors to a bayesian probability of response-the value of transparency. Arthritis Rheum. 2013;65:607-8.	A7
53	Kaur, K, Kalra, S, Kaushal, S. Systematic review of tofacitinib: a new drug for the management of rheumatoid arthritis. Clinical therapeutics. 2014;36(7):1074-86.	A7
54	Kawalec, P, Mikrut, A, Wisniewska, N, Pilc, A. The effectiveness of tofacitinib, a novel Janus kinase inhibitor, in the treatment of rheumatoid arthritis: a systematic review and meta-analysis. Clinical rheumatology. 2013;32(10):1415-24.	A7
55	Keystone, E, Fleischmann, R, Van Vollenhoven, R, Kremer, J, Gruben, D, Bradley, J, et al. Tofacitinib, an oral janus kinase inhibitor: Post-hoc analyses of efficacy and safety of monotherapy versus combination therapy in a phase 3 rheumatoid arthritis population. Ann Rheum Dis. 2013;72.	A7
56	Kivitz, AJ, Haraoui, B, Kaine, J, Castellano, V, Bananis, E, Connell, CA, et al. A safety analysis of Tofacitinib 5mg twice daily administered as monotherapy or in combination with background conventional synthetic DMARDs in a phase 3 rheumatoid arthritis population. Arthritis Rheumatol. 2015;67.	A7
57	Kremer, J, Cohen, S, Wilkinson, B, Gruben, D, Wallenstein, GV, Kanik, KS, et al. Safety and efficacy after 24 week (Wk) dosing of the oral JAK inhibitor CP-690,550 (CP) in combination with methotrexate (MTX) in patients (PTS) with active rheumatoid arthritis (RA). Arthritis Rheum. 2009;60:1925.	A7
58	Kremer, J, Li, ZG, Hall, S, Fleischmann, R, Genovese, M, Martin-Mola, E, et al. Tofacitinib in combination with nonbiologic disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis: a randomized trial. Ann Intern Med. 2013;159(4):253-61.	A3
59	Kremer, J, Zerbini, C, Lee, EB, Gruben, D, Krishnaswami, S, Zwillich, SH, et al. Tofacitinib, an oral Janus kinase inhibitor: Analyses of efficacy endpoints by subgroups in a pooled Phase 2 and 3 rheumatoid arthritis study population. International journal of rheumatic diseases. 2012;15:67-8.	A7
60	Kremer, J, Zerbini, C, Lee, EB, Gruben, D, Krishnaswami, S, Zwillich, SH, et al. Tofacitinib (CP-690,550), an oral janus kinase inhibitor: Analyses of efficacy endpoints by subgroups in a pooled phase 2 and 3 rheumatoid arthritis study population. Ann Rheum Dis. 2013;71.	A7
61	Kremer, JM, Bloom, BJ, Breedveld, FC, Coombs, JH, Fletcher, MP, Gruben, D, et al. The safety and efficacy of a JAK inhibitor in patients with active rheumatoid arthritis: Results of a double-blind, placebo-controlled phase IIa trial of three dosage levels of CP-690,550 versus placebo. Arthritis Rheum. 2009;60(7):1895-905.	A5
62	Kremer, JM, Cohen, S, Wilkinson, BE, Connell, CA, French, JL, Gomez-Reino, J, et al. A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. Arthritis Rheum. 2012;64(4):970-81.	A3

63	Lee, YH, Bae, SC. Comparative efficacy and safety of tocilizumab, rituximab, abatacept and tofacitinib in patients with active rheumatoid arthritis that inadequately responds to tumor necrosis factor inhibitors: a Bayesian network meta-analysis of randomized controlled trials. International journal of rheumatic diseases. 2015.	A7
64	Lee, YH, Bae, SC, Song, GG. Comparative efficacy and safety of tofacitinib, with or without methotrexate, in patients with active rheumatoid arthritis: a Bayesian network meta-analysis of randomized controlled trials. Rheumatology international. 2015;35(12):1965-74.	A7
65	Lopez-Olivo, MA, Suarez-Almazor, ME, Bavineni, M. Tofacitinib for rheumatoid arthritis: A systematic review and meta-analysis. Arthritis Rheum. 2013;65:625.	A7
66	Mariette, X, Curtis, JR, Lee, EB, Benda, B, Kaplan, I, Soma, K, et al. Tofacitinib, an oral janus kinase inhibitor: Analysis of malignancies across the rheumatoid arthritis clinical program. Arthritis Rheum. 2013;65:340.	A7
67	Mariette, X, Curtis, JR, Lee, EB, Riese, R, Kaplan, I, Chew, R, et al. Tofacitinib (CP-690,550), an oral janus kinase inhibitor: Analysis of malignancies across the rheumatoid arthritis clinical programme. Ann Rheum Dis. 2013;71.	A7
68	Romero, ACG, Donath, E, Osorno, H, Kumar, S. An indirect comparisons analysis between biologic disease modifiers in the treatment of rheumatoid arthritis to evaluate for efficacy and safety. Arthritis Rheum. 2013;65:203.	A7
69	Salgado, E, Gomez-Reino, JJ. The JAK inhibitor tofacitinib for active rheumatoid arthritis: Results from Phase III trials. International Journal of Clinical Rheumatology. 2013;8(3):315-26.	A7
70	Salgado, E, Maneiro, JR, Carmona, L, Gomez-Reino, JJ. Safety profile of protein kinase inhibitors in rheumatoid arthritis: Systematic review and meta-analysis. Ann Rheum Dis. 2014;73(5):871-82.	A7
71	Schulze-Koops, H, Rubbert-Roth, A, Behmer, O, Kerkmann, U, Kwok, K, Burkhardt, H. Tofacitinib (CP-690,550), an oral janus kinase inhibitor: Pooled phase 3 analysis in a German rheumatoid arthritis study population. Zeitschrift fur Rheumatologie. 2012;71:90.	A7
72	Singh, JA, Hossain, A, Tanjong Ghogomu, E, Kotb, A, Christensen, R, Mudano, AS, et al. Biologics or tofacitinib for rheumatoid arthritis in incomplete responders to methotrexate or other traditional disease-modifying anti-rheumatic drugs: A systematic review and network meta-analysis. Cochrane Database of Systematic Reviews. 2016;2016(5).	A7
73	Song, GG, Bae, SC, Lee, YH. Efficacy and safety of tofacitinib for active rheumatoid arthritis with an inadequate response to methotrexate or disease-modifying antirheumatic drugs: a meta-analysis of randomized controlled trials. The Korean journal of internal medicine. 2014;29(5):656-63.	A7
74	Strand, V, Ahadieh, S, French, J, Geier, J, Krishnaswami, S, Menon, S, et al. Meta- analysis of serious infections with tofacitinib and biological treatment in rheumatoid arthritis clinical trials. Pharmacotherapy. 2014;34(10):224.	A7
75	Strand, V, Ahadieh, S, French, J, Geier, J, Krishnaswami, S, Menon, S, et al. Meta- analysis of serious infections with tofacitinib and biological treatment in rheumatoid arthritis clinical trials. Arthritis Rheumatol. 2014;66:195-6.	A7
76	Strand, V, Ahadieh, S, French, J, Geier, J, Krishnaswami, S, Menon, S, et al. Systematic review and meta-analysis of serious infections with tofacitinib and biologic disease-modifying antirheumatic drug treatment in rheumatoid arthritis clinical trials. Arthritis research & therapy. 2015;17(1):362.	A7
77	Strand, V, Alten, RE, Nduaka, CI, Riese, R, Gruben, D, Zwillich, SH, et al. Improvements in physical function correlate with improvements in health related quality of life: Reported outcomes in rheumatoid arthritis patients treated with tofacitinib: Results from 3 randomized phase 3 trials. Arthritis Rheum. 2013;65:993-	A7

	4.	
78	Strand, V, Burmester, GR, Zerbini, CA, Mebus, CA, Zwillich, SH, Gruben, D, et al. Tofacitinib with methotrexate in third-line treatment of patients with active rheumatoid arthritis: patient-reported outcomes from a phase III trial. Arthritis care & research. 2015;67(4):475-83.	A7
79	Strand, V, Dikranian, A, Beal, J, Kwok, K, Krishnaswami, S, Wood, S, et al. Analysis of early neutropenia, clinical response, and serious infection events in patients receiving tofacitinib for rheumatoid arthritis. Arthritis Rheumatol. 2014;66:1086-7.	A7
80	Strand, V, Fleischmann, R, Alten, RE, Koncz, T, Zwillich, SH, Bradley, JD, et al. Oral start: Effects of the oral JAK inhibitor tofacitinib monotherapy versus methotrexate on patient-reported outcomes in the phase 3 oral start trial of active rheumatoid arthritis. Ann Rheum Dis. 2013;72.	A1
81	Strand, V, Kremer, J, Wallenstein, G, Kanik, KS, Connell, C, Gruben, D, et al. Effects of tofacitinib monotherapy on patient-reported outcomes in a randomized phase 3 study of patients with active rheumatoid arthritis and inadequate responses to DMARDs. Arthritis research & therapy. 2015;17:307.	A7
82	Strand, V, Kremer, JM, Li, ZG, Hall, S, Fleischmann, RM, Genovese, MC, et al. Tofacitinib (CP-690,550) in combination with traditional disease-modifying anti-rheumatic drugs: Phase 3 study patient-reported outcomes in patients with active rheumatoid arthritis and an inadequate response to disease-modifying anti-rheumatic drugs. Arthritis Rheum. 2011;63(10 SUPPL. 1).	A7
83	Strand, V, Van Der Heijde, D, Landewe, R, Lee, EB, Wilkinson, B, Zwillich, SH, et al. Remission at 3 or 6 months and radiographic non-progression at 12 months in methotrexate-naive rheumatoid arthritis patients treated with tofacitinib or methotrexate: A post-HOC analysis of the oral start trial. Arthritis Rheum. 2013;65:842-3.	A1
84	Strand, V, Van Der Heijde, D, Zerbini, CAF, Connell, CA, Gruben, D, Riese, R, et al. Oral scan: Effects of the oral JAK inhibitor tofacitinib in combination with methotrexate on patient reported outcomes in a 24-month phase 3 trial of active rheumatoid arthritis. Arthritis Rheum. 2013;65:996-7.	A7
85	Strand, V, van Vollenhoven, RF, Lee, EB, Fleischmann, R, Zwillich, SH, Gruben, D, et al. Tofacitinib or adalimumab versus placebo: patient-reported outcomes from a phase 3 study of active rheumatoid arthritis. Rheumatology (Oxford, England). 2016;55(6):1031-41.	A7
86	Takeuchi, T, Tanaka, Y, Yamanaka, H, Suzuki, M, Nakamura, H, Yazawa, KK, et al. Tofacitinib (CP-690,550), an oral janus kinase inhibitor, as monotherapy in patients with active rheumatoid arthritis: A comparison between japanese and global populations over 12 weeks of dosing in phase 2b studies. Arthritis Rheum. 2011;63(10 SUPPL. 1).	A7
87	Tanaka, Y, Suzuki, M, Nakamura, H, Toyoizumi, S, Zwillich, SH. Phase II study of tofacitinib (CP-690,550) combined with methotrexate in patients with rheumatoid arthritis and an inadequate response to methotrexate. Arthritis care & research. 2011;63(8):1150-8.	A3
88	Tanaka, Y, Takeuchi, T, Yamanaka, H, Hirose, T, Toyoizumi, S, Zwillich, S, et al. Tofacitinib, an oral janus kinase inhibitor: Analysis of malignancies in japanese patients across the rheumatoid arthritis clinical programme. Ann Rheum Dis. 2014;73.	A7
89	Tanaka, Y, Takeuchi, T, Yamanaka, H, Nakamura, H, Toyoizumi, S, Zwillich, S. Efficacy and Safety of Tofacitinib as Monotherapy in Japanese Patients With Active Rheumatoid Arthritis: A 12-Week, Randomized, Phase 2 Study. Modern Rheumatology / the Japan Rheumatism Association. 2014:1-25.	A7

90	Tanaka, Y, Takeuchi, T, Yamanaka, H, Nakamura, H, Toyoizumi, S, Zwillich, S. Efficacy and safety of tofacitinib as monotherapy in japanese patients with active rheumatoid arthritis: A 12-week, randomized, phase 2 study. Modern Rheumatology. 2015;25(4):514-21.	A3
91	Tanaka, Y, Takeuchi, T, Yamanaka, H, Sugiyama, N, Yoshinaga, T, Togo, K, et al. Malignancy data in tofacitinib-treated Japanese patients with rheumatoid arthritis. Arthritis Rheumatol. 2015;67.	A7
92	Tanaka, Y, Takeuchi, T, Yamanaka, H, Suzuki, M, Nakamura, H, Toyoizumi, S, et al. Tofacitinib (CP-690,550), an oral janus kinase inhibitor, as monotherapy in japanese patients with active rheumatoid arthritis: A 12-week phase 2b study. Arthritis Rheum. 2011;63(10 SUPPL. 1).	A7
93	Tarp, S, Furst, DE, Luta, G, Boers, M, Tarp, U, Asmussen, KH, et al. Risk of serious adverse effects associated with different biological and targeted synthetic disease-modifying anti-rheumatic drugs in patients with rheumatoid arthritis: A systematic review and meta-analysis of randomised trials. Ann Rheum Dis. 2015;74:176-7.	A7
94	van der Heijde, D, Tanaka, Y, Fleischmann, R, Keystone, E, Kremer, J, Zerbini, C, et al. Tofacitinib (CP-690,550) in patients with rheumatoid arthritis receiving methotrexate: twelve-month data from a twenty-four-month phase III randomized radiographic study. Arthritis Rheum. 2013;65(3):559-70.	A3
95	Van Der Heijde, D, Tanaka, Y, Fleischmann, R, Keystone, EC, Kremer, JM, Zerbini, CAF, et al. Tofacitinib (CP-690,550), an oral janus kinase inhibitor, in combination with methotrexate reduced the progression of structural damage in patients with rheumatoid arthritis: A 24-month phase 3 study. Arthritis Rheum. 2011;63(10 SUPPL. 1).	A7
96	Van Der Heijde, D, Tanaka, Y, Fleischmann, RM, Keystone, E, Kremer, JM, Zerbini, C, et al. Tofacitinib, an oral janus kinase inhibitor, in combination with methotrexate reduced the progression of structural damage in patients with rheumatoid arthritis: Year 2 efficacy and safety results from a 24-month phase 3 study. Arthritis Rheum. 2012;64:546-7.	A7
97	Van Vollenhoven, RF, Fleischmann, RM, Cohen, SB, Lee, EB, Meijide, G, Wagner, S, et al. Tofacitinib (CP-690,550), an oral janus kinase inhibitor, or adalimumab versus placebo in patients with rheumatoid arthritis on background methotrexate: A phase 3 study. Arthritis Rheum. 2011;63(10 SUPPL. 1).	A7
98	Van Vollenhoven, RF, Krishnaswami, S, Benda, B, Gruben, D, Wilkinson, B, Mebus, CA, et al. Tofacitinib and adalimumab achieve similar rates of low disease activity in rheumatoid arthritis-lack of improvement in disease activity score by 3 months predicts low likelihood of low disease activity at 1 year. Arthritis Rheum. 2012;64:556.	A7
99	Van Vollenhoven, RF, Wallenstein, G, Lee, EB, Fleischmann, R, Zwillich, SH, Gruben, D, et al. Effects of tofacitinib (CP-690,550), an oral janus kinase inhibitor, or adalimumab on patient reported outcomes in a phase 3 study of active rheumatoid arthritis. Ann Rheum Dis. 2013;71.	A7
100	Vieira, MC, Wallenstein, G, Bradley, J, Gruben, D, Koncz, T, Zwillich, SH, et al. Tofacitinib versus biologic treatments in patients with active rheumatoid arthritis who have had an inadequate response to tumour necrosis factor inhibitors-a network meta-analysis. Ann Rheum Dis. 2013;72.	A7
101	Vieira, MC, Wallenstein, GV, Bradley, JD, Gruben, D, Koncz, T, Zwillich, SH, et al. Tofacitinib versus biologic treatments with and without methotrexate in patients with active rheumatoid arthritis who have had an inadequate response to traditional disease modifying anti-rheumatic drugs-a network meta-analysis. Ann Rheum Dis. 2013;71.	A7

102	Wallenstein, GV, Cohen, S, Cutolo, M, Gruben, D, Kanik, KS, Kremer, JM, et al. Effects of the oral JAK inhibitor CP-690,550 (CP) On pain, physical functioning (PF), fatigue, and health-related quality of life (HRQoL) in patients (pts) with active rheumatoid arthritis (RA). Arthritis Rheum. 2009;60:418.	A7
103	Wallenstein, GV, Kanik, KS, Wilkinson, B, Cohen, S, Cutolo, M, Fleishmann, R, et al. Effects of the oral Janus kinase inhibitor tofacitinib on patient-reported outcomes in patients with active rheumatoid arthritis: results of two Phase 2 randomised controlled trials. Clinical and experimental rheumatology. 2016.	A7
104	Yamaoka, K, Kubo, S, Sonomoto, K, Hirata, S, Cavet, G, Bolce, R, et al. Correlation of a multi-biomarker disease activity (vectra DA) score with clinical disease activity and its components with radiographic progression in rheumatoid arthritis patients treated with tofacitinib. Arthritis Rheum. 2012;64:914-5.	A7
105	Zhang, X, Liang, F, Yin, X, Xiao, X, Shi, P, Wei, D, et al. Tofacitinib for acute rheumatoid arthritis patients who have had an inadequate response to disease-modifying antirheumatic drug (DMARD): a systematic review and meta-analysis. Clinical rheumatology. 2014;33(2):165-73.	A7
106	Bergman, MJ, Yazici, Y, Dikranian, A, Bourret, J, Zang, C, Mojcik, CF, et al. Evaluation of disease activity in patients with rheumatoid arthritis treated with tofacitinib by RAPID3: An analysis of data from 6 phase 3 studies. Arthritis and Rheumatology. 2016;68:1985-6.	A7
107	Bergrath, E, Wallenstein, G, Gerbert, R, Gruben, D, Makin, C. Tofacitinib versus biologic treatments in moderate to severe rheumatoid arthritis patients who have had an inadequate response to nonbiologic DMARDS: Systematic literature review and network meta-analysis. Annals of the rheumatic diseases. 2016;75:726.	A4
108	Charles-Schoeman, C, Wicker, P, Gonzalez-Gay, MA, Boy, M, Zuckerman, A, Soma, K, et al. Cardiovascular safety findings in patients with rheumatoid arthritis treated with tofacitinib, an oral Janus kinase inhibitor. Seminars in arthritis and rheumatism. 2016;46(3):261-71.	A7
109	Chatzidionysiou, K, Emamikia, S, Nam, JL, Ramiro, S, Smolen, J, Van Der Heijde, D, et al. Efficacy and safety of conventional and targeted synthetic disease-modifying antirheumatic drugs as well as glucocorticoids: A systematic literature review informing the 2016 update of the eular recommendations for the management of rheumatoid arthritis. Arthritis and Rheumatology. 2016;68:3434-7.	A7
110	Cohen, SB, Koenig, A, Wang, L, Kwok, K, Mebus, CA, Riese, RJ, et al. Efficacy and safety of tofacitinib in US and non-US rheumatoid arthritis patients: Pooled analyses of phase II and III. Clinical and experimental rheumatology. 2016;34(1):32-6.	A7
111	Curtis, JR, Schulze-Koops, H, Takiya, L, Mebus, CA, Terry, KK, Biswas, P, et al. Efficacy and safety of tofacitinib in older and younger patients with rheumatoid arthritis. Clinical and experimental rheumatology. 2017. 0392-856X 0392-856X (ISSNLinking)	A7
112	Desai, RJ, Thaler, KJ, Mahlknecht, P, Gartlehner, G, McDonagh, MS, Mesgarpour, B, et al. Comparative Risk of Harm Associated With the Use of Targeted Immunomodulators: A Systematic Review. Arthritis Care and Research. 2016;68(8):1078-88.	A7
113	Fleischmann, R, Landew, R, J.S, S. Review of head-to-head study designs in rheumatoid arthritis. Seminars in arthritis and rheumatism. 2016;46(3):279-85.	A7
114	Hall, S, Nash, P, Rischmueller, M, Wiid, Z, Witcombe, D, Gruben, D, et al. Efficacy of tofacitinib in patients who are inadequate responders to disease-modifying antirheumatic drugs according to early versus late duration of rheumatoid arthritis: Post-hoc analysis of data from phase 3 trials. Arthritis and Rheumatology. 2016;68:1999-2001.	A3

115	Kivitz, A, Haraoui, B, Kaine, J, Castellano, V, Bananis, E, Connell, CA, et al. A safety analysis of tofacitinib 5 mg twice daily administered as monotherapy or in combination with background conventional synthetic DMARDS in a phase 3 rheumatoid arthritis population. Annals of the rheumatic diseases. 2016;75:247-8.	A7
116	Lee, YH, Bae, SC. Comparative efficacy and safety of tocilizumab, rituximab, abatacept and tofacitinib in patients with active rheumatoid arthritis that inadequately responds to tumor necrosis factor inhibitors: a Bayesian network meta-analysis of randomized controlled trials. International journal of rheumatic diseases. 2016;19(11):1103-11.	A3
117	Lopez-Olivo, MA, Matusevich, A, Suarez-Almazor, M. Therapy in patients with rheumatoid arthritis (RA) with inadequate response to tumor necrosis factor alpha inhibitors (TNFi): A systematic review and meta-analysis of randomized controlled trials (RCTs). Arthritis and Rheumatology. 2016;68:809-11.	A7
118	Park, SK, Lee, MY, Jang, EJ, Kim, HL, Ha, DM, Lee, EK. A comparison of discontinuation rates of tofacitinib and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis: a systematic review and Bayesian network meta-analysis. Clinical and experimental rheumatology. 2017. 0392-856X 0392-856X (ISSNLinking)	A7
119	Radominski, SoC, Cardiel, MH, Citera, G, Goecke, A, Jaller, JJ, Lomonte, ABV, et al. Tofacitinib, an oral Janus kinase inhibitor, for the treatment of Latin American patients with rheumatoid arthritis: Pooled efficacy and safety analyses of Phase 3 and long-term extension studies. Reumatologia clinica. 2016. 1699-258X (ISSNLinking)	A7
120	Schulze-Koops, H, Strand, V, Nduaka, C, DeMasi, R, Wallenstein, G, Kwok, K, et al. Analysis of haematological changes in tofacitinib-treated patients with rheumatoid arthritis across phase 3 and long-term extension studies. Rheumatology (Oxford, England). 2017;56(1):46-57.	A7
121	Singh, JA, Hossain, A, Tanjong Ghogomu, E, Mudano, AS, Tugwell, P, Wells, GA. Biologic or tofacitinib monotherapy for rheumatoid arthritis in people with traditional disease-modifying anti-rheumatic drug (DMARD) failure: a Cochrane Systematic Review and network meta-analysis (NMA). The Cochrane database of systematic reviews. 2016;11:CD012437.	A3
122	Tarp, S, Eric Furst, D, Boers, M, Luta, G, Bliddal, H, Tarp, U, et al. Risk of serious adverse effects of biological and targeted drugs in patients with rheumatoid arthritis: a systematic review meta-analysis. Rheumatology (Oxford, England). 2016. 1462-0324 (ISSNLinking)	A7
123	Vieira, MC, Zwillich, SH, Jansen, JP, Smiechowski, B, Spurden, D, Wallenstein, GV. Tofacitinib Versus Biologic Treatments in Patients With Active Rheumatoid Arthritis Who Have Had an Inadequate Response to Tumor Necrosis Factor Inhibitors: Results From a Network Meta-analysis. Clinical therapeutics. 2016;38(12):2628-41.	A3
124	Wallenstein, GV, Kanik, KS, Wilkinson, B, Cohen, S, Cutolo, M, Fleishmann, R, et al. Effects of the oral Janus kinase inhibitor tofacitinib on patient-reported outcomes in patients with active rheumatoid arthritis: results of two Phase 2 randomised controlled trials. Clinical and experimental rheumatology. 2016;34(3):430-42.	A3
125	Fleischmann, R, Kremer, J, Tanaka, Y, Gruben, D, Kanik, K, Koncz, T, et al. Efficacy and safety of tofacitinib in patients with active rheumatoid arthritis: review of key Phase 2 studies. International journal of rheumatic diseases. 2017;(no pagination).	A5
126	Kivitz, AJ, Haraoui, B, Kaine, J, Castellano, V, Bananis, E, Connell, CA, et al. A safety analysis of Tofacitinib 5mg twice daily administered as monotherapy or in combination with background conventional synthetic DMARDs in a phase 3 rheumatoid arthritis population. Arthritis and Rheumatology. 2015;67.	A7

127	Vieira, MC, Wallenstein, G, Bradley, J, Gruben, D, Koncz, T, Zwillich, SH, et al. Tofacitinib versus biologic treatments in patients with active rheumatoid arthritis who have had an inadequate response to tumour necrosis factor inhibitors-a network meta-analysis. Annals of the rheumatic diseases. 2013;72.	A5
128	Vieira, MC, Wallenstein, GV, Bradley, JD, Gruben, D, Koncz, T, Zwillich, SH, et al. Tofacitinib versus biologic treatments with and without methotrexate in patients with active rheumatoid arthritis who have had an inadequate response to traditional disease modifying anti-rheumatic drugs-a network meta-analysis. Annals of the Rheumatic Disease. 2013;71.	A5
129	Vollenhoven, RF, Tanaka, Y, Lamba, M, Collinge, M, Hendrikx, T, Hirose, T, et al. Relationship between NK cell count and important safety events in rheumatoid arthritis patients treated with tofacitinib. Annals of the rheumatic diseases. 2015;74:258-9.	A3
130	Zerbini, C, Radominski, S, Cardiel, MH, Castaneda, O, Romero, F, Citera, G, et al. Efficacy and safety of tofacitinib monotherapy versus combination therapy in a latin american subpopulation of patients with rheumatoid arthritis: A pooled phase 3 analysis. Arthritis and Rheumatology. 2015;67.	A7

#### Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht relevant, da keine indirekten Vergleiche durchgeführt wurden.

#### Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant, da keine nicht randomisierten vergleichenden Studien in die Ergebnisdarstellung dieses Dossiers einbezogen wurden.

#### Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht relevant, da keine weiteren Untersuchungen dargestellt wurden.

# Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nummer	Registereintrag	Ausschluss- kriterium
1	ClinicalTrials.gov. 2010. A Study To Compare The Amount Of Tasocitinib (CP-690,550) That Is Absorbed Into The Blood Of Healthy Subjects Following Oral Administration Of Three Different Tablet Versions Of Tasocitinib (CP-690,550); Abrufbar unter: https://ClinicalTrials.gov/show/NCT01184092 [Zugriff am: 03.02.2017].	A1
2	ClinicalTrials.gov. 2010. A Study To Evaluate The Effect Of Food On How Tasocitinib (CP-690,550) Is Absorbed And Moves Through The Body Following Oral Administration Of Tasocitinib (CP-690,550) Tablets To Healthy Subjects; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01184001 [Zugriff am: 03.02.2017].	A1
3	ClinicalTrials.gov. 2010. A Study To Estimate The Amount Of CP-690,550 (Study Drug) That Is Absorbed Into The Blood Of Healthy Subjects Following Oral Administration Of CP-690,550 In Tablet Form; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01143805 [Zugriff am: 03.02.2017].	A1
4	ClinicalTrials.gov. 2010. A Study Of The Taste Of Different Oral Liquid Preparations Of CP-690,550 In Trained Adult Product Testers; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01023100 [Zugriff am: 03.02.2017].	A1
5	ClinicalTrials.gov. 2011. Effect Of CP-690,550 On The Pharmacokinetics Of Metformin In Healthy Volunteers; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01405118 [Zugriff am: 03.02.2017].	A1
6	ClinicalTrials.gov. 2011. A Study To Compare The Amount Of CP-690,550 That Is Absorbed Into The Blood Of Healthy Subjects Following Oral Administration Of Two Different Strength Tablets Of CP-690,550; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01277991 [Zugriff am: 03.02.2017].	A1
7	ClinicalTrials.gov. 2011. CP-690,550 Pharmacokinetics In Healthy Chinese Subjects; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01101919 [Zugriff am: 03.02.2017].	A1
8	ClinicalTrials.gov. 2012. Study Of The Effects Of Atorvastatin On Cholesterol Levels In Rheumatoid Arthritis Patients Taking CP-690,550; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01059864 [Zugriff am: 03.02.2017].	A6
9	ClinicalTrials.gov. 2012. Effects Of CP-690,550 (Tasocitinib) On Cholesterol Metabolism In Patients With Active Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01262118 [Zugriff am: 03.02.2017].	A6
10	ClinicalTrials.gov. 2012. A Study To Evaluate The Mechanism Of Action Of CP-690,550 In Patients With Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00976599 [Zugriff am: 03.02.2017].	A6

11	ClinicalTrials.gov. 2012. Three Dose Levels of CP-690,550 Monotherapy Versus Placebo, Administered Orally Twice Daily (BID) for 6 Weeks; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00147498 [Zugriff am: 03.02.2017].	A3
12	ClinicalTrials.gov. 2012. Long-Term Safety Follow-Up Of Subjects Previously Enrolled In Rheumatoid Arthritis Studies Of CP-690,550; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00414661 [Zugriff am: 03.02.2017].	A6
13	ClinicalTrials.gov. 2013. A Study To Assess the Immune Response Following Administration Of Influenza and Pneumococcal Vaccines To Subjects With Rheumatoid Arthritis Receiving CP-690,550 Or Placebo; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01359150 [Zugriff am: 03.02.2017].	A1
14	ClinicalTrials.gov. 2013. Co-Administration Of Methotrexate And CP-690,550; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01745055 [Zugriff am: 03.02.2017].	A6
15	ClinicalTrials.gov. 2013. Comparing The Effectiveness And Safety Of 2 Doses Of An Experimental Drug, CP-690,550, To Methotrexate (MTX) In Patients With Rheumatoid Arthritis Who Have Not Previously Received MTX; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01039688 [Zugriff am: 03.02.2017].	A1
16	ClinicalTrials.gov. 2014. A Study in Healthy Adults and Adult Subjects With Rheumatoid Arthritis to Evaluate the Safety, Tolerability and Pharmacokinetics After Multiple Doses of ABT-494; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01741493 [Zugriff am: 03.02.2017].	A1
17	ClinicalTrials.gov. 2014. A Study To Evaluate The Effect Of CP-690,550 On Measures Of Kidney Function In Patients With Active Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01484561 [Zugriff am: 03.02.2017].	A1
18	ClinicalTrials.gov. 2015. Effects of Tofacitinib (CP-690,550) on Magnetic Resonance Imaging (MRI)- Assessed Joint Structure In Early Rheumatoid Arthritis (RA); Abrufbar unter: https://ClinicalTrials.gov/show/NCT01164579 [Zugriff am: 03.02.2017].	A1
19	ClinicalTrials.gov. 2015. Long-Term, Open-Label Study Of CP-690,550 For Treatment Of Rheumatoid Arthritis In Japan; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00661661 [Zugriff am: 03.02.2017].	A6
20	ClinicalTrials.gov. 2015. A Study to Assess Immune Response Following Zoster Vaccination to Subjects With Rheumatoid Arthritis Receiving Tofacitinib or Placebo With Background Methotrexate; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02147587 [Zugriff am: 03.02.2017].	A3
21	ClinicalTrials.gov. 2015. An Evaluation of the Optimal Dose of Tofacitinib Needed to Achieve Low Disease Activity (LDA) or Clinical Remission in Patients With Active Rheumatoid Arthritis (RA) as Measured From a Clinical and Structural Perspective; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02566967 [Zugriff am: 03.02.2017].	A6
22	ClinicalTrials.gov. 2016. Musculoskeletal Ultrasound Assessment of Therapeutic Response of Tofacitinib in Rheumatoid Arthritis Patients; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02321930 [Zugriff am: 03.02.2017].	A6
23	ClinicalTrials.gov. 2016. A Study To Evaluate The Safety And Efficacy Of Tofacitinib Modified Release Tablets Compared To Tofacitinib Immediate Release Tablets In Adult Patients With Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02281552 [Zugriff am: 03.02.2017].	A1
24	ClinicalTrials.gov. 2016. Evaluation of the Condition of Rheumatoid Arthritis After Treatment; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02157012 [Zugriff am: 03.02.2017].	A6

25	ClinicalTrials.gov. 2016. Tofacitinib (Xeljanz) Special Investigation for Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT01932372 [Zugriff am: 03.02.2017].	A6
26	ClinicalTrials.gov. 2016. Long-Term Effectiveness And Safety Of CP-690,550 For The Treatment Of Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00413699 [Zugriff am: 03.02.2017].	A6
27	ClinicalTrials.gov. 2016. Safety Study Of Tofacitinib Versus Tumor Necrosis Factor (TNF) Inhibitor In Subjects With Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02092467 [Zugriff am: 03.02.2017].	A8
28	ClinicalTrials.gov. 2016. An Efficacy And Safety Study Evaluating Tofacitinib With And Without Methotrexate Compared To Adalimumab With Methotrexate (ORAL STRATEGY); Abrufbar unter: https://ClinicalTrials.gov/show/NCT02187055 [Zugriff am: 03.02.2017].	A8
29	ClinicalTrials.gov. 2012. A Phase 2 Study For Patients With A Physician's Diagnosis Of Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00550446 [Zugriff am: 02.02.2017].	A5
30	ClinicalTrials.gov. 2012. A Study Comparing 2 Doses Of CP-690,550 Vs. Placebo For The Treatment Of Rheumatoid Arthritis In Patients On Other Background Arthritis Medications; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00856544 [Zugriff am: 02.02.2017].	A3
31	ClinicalTrials.gov. 2012. Study of CP-690,550 Versus Placebo In Rheumatoid Arthritis Patients On Background Methotrexate With Inadequate Response To Tumor Necrosis Factor (TNF) Inhibitors; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00960440 [Zugriff am: 02.02.2017].	A3
32	ClinicalTrials.gov. 2012. Comparison Of 4 CP-690,550 Doses Vs. Placebo, Each Combined With Methotrexate, For The Treatment Of Rheumatoid Arthritis in Japan; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00603512 [Zugriff am: 03.02.2017].	A3
33	ClinicalTrials.gov. 2013. A Phase 3 Study Comparing 2 Doses of CP-690,550 vs. Placebo for Treatment of Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00814307 [Zugriff am: 03.02.2017].	A3
34	ClinicalTrials.gov. 2013. A 2-Year Phase 3 Study Of CP-690,550 In Patients With Active Rheumatoid Arthritis On Background Methotrexate; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00847613 [Zugriff am: 03.02.2017].	A3
35	ClinicalTrials.gov. 2013. Comparison Of 5 CP-690,550 Doses Vs. Placebo, For The Treatment Of Rheumatoid Arthritis In Japan; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00687193 [Zugriff am: 03.02.2017].	A3
36	ClinicalTrials.gov. 2013. Comparison Of 6 CP-690,550 Doses Vs.Placebo, Each Combined With Methotrexate, For The Treatment Of Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT00413660 [Zugriff am: 03.02.2017].	A3
37	Tocilizumab to Other Biologics in Patients With Rheumatoid Arthritis (RA); Abrufbar unter: https://ClinicalTrials.gov/show/NCT02797769 [Zugriff am: 03.02.2017].	A6
38	ClinicalTrials.gov. 2017. Effectiveness and Safety of Tofacitinib in Korean Patients With Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT03011281 [Zugriff am: 03.02.2017].	A8
39	ClinicalTrials.gov. 2017. Methotrexate Withdrawal Study of Tofacitinib Modified Release Formulation in Subjects With Rheumatoid Arthritis; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02831855 [Zugriff am: 03.02.2017].	A2

40	ClinicalTrials.gov. 2017. Safety and Efficacy of Pf-06650833 In Subjects With Rheumatoid Arthritis, With An Inadequate Response To Methotrexate; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02996500 [Zugriff am: 03.02.2017].	A2
41	ClinicalTrials.gov. 2017. Korean Post-marketing Surveillance for Xeljanz; Abrufbar unter: https://ClinicalTrials.gov/show/NCT02984020 [Zugriff am: 03.02.2017].	A6
42	ClinicalTrials.gov. 2017. Safety and Tolerability of Herpes Zoster Vaccine Rheumatologic Patients; Abrufbar unter: https://ClinicalTrials.gov/show/NCT03016884 [Zugriff am: 03.02.2017].	A8
43	EU Clinical Trials Register. 2005. A PHASE 2A, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO COMPARE 3 DOSE LEVELS OF CP 690,550 VERSUS PLACEBO, ADMINISTERED ORALLY TWICE DAILY (BID) FOR 6 WEEKS, IN THE TREATMENT OF THE SIGNS AND SYMPTOMS OF SUBJECTS WITH ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2004-002846-36 [Zugriff am: 06.02.2017].	A3
44		A6
45	ACTIVE COMPARATOR-CONTROLLED TRIAL TO EVALUATE THE SAFETY AND EFFICACY OF COADMINISTRATION OF CP-690,550 AND MYCOPHENOLATE MOFETIL / MYCOPHENOLATE SODIUM IN DE NOVO KIDNEY ALLOGRAFT RECIPIENTS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2007-001984-31 [Zugriff am: 06.02.2017].	A1
46	EU Clinical Trials Register. 2007. ESTUDIO OBSERVACIONAL PROSPECTIVO PARA EVALUAR LA SEGURIDAD A LARGO PLAZO Y EL ESTADO FUNCIONAL DE SUJETOS CON ARTRITIS REUMATOIDE INCLUIDOS PREVIAMENTE EN ESTUDIOS DE CP-690,550 [A PROSPECTIVE OBSERVATIONAL STUDY TO EVALUATE LONG TERM SAFETY AND FUNCTIONAL STATUS OF SUBJECTS WITH RHEUMATOID ARTHRITIS PREVIOUSLY ENROLLED IN STUDIES OF CP-690,550]; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2006-006373-25 [Zugriff am: 06.02.2017].	A6
47	EU Clinical Trials Register. 2008. A PHASE 2, MULTICENTER, OPEN-LABEL, ACTIVE COMPARATOR CONTROLLED, EXTENSION TRIAL TO EVALUATE THE LONG-TERM SAFETY AND EFFICACY OF CP-690,550 IN RENAL ALLOGRAFT RECIPIENTS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-002345-23 [Zugriff am: 06.02.2017].	A1
48	EU Clinical Trials Register. 2008. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTER STUDY TO INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CROHN'S DISEASE; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-003571-45 [Zugriff am: 06.02.2017].	A1
49	EU Clinical Trials Register. 2009. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTER STUDY TO INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-004564-40 [Zugriff am: 06.02.2017].	A1

1 50	ELI Clinical Trials Decistor, 2010, A DUASE 2, MULTI SITE, ODEN I ADEL	I A 1
50	EU Clinical Trials Register. 2010. A PHASE 3, MULTI-SITE, OPEN-LABEL STUDY OF THE LONG TERM SAFETY AND TOLERABILITY OF 2 ORAL	A1
	DOSES OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE	
	CHRONIC PLAQUE PSORIASIS; Abrufbar unter:	
	https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-	
	020002-15 [Zugriff am: 06.02.2017].	
51	EU Clinical Trials Register. 2010. A PHASE 3, MULTI SITE, RANDOMIZED,	A1
	DOUBLE BLIND, PLACEBO CONTROLLED, PARALLEL GROUP STUDY OF	
	THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 AND 1	
	SUBCUTANEOUS DOSE OF ETANERCEPT IN SUBJECTS WITH MODERATE	
	TO SEVERE CHRONIC PLAQUE PSORIASIS; Abrufbar unter:	
	https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-	
	020004-30 [Zugriff am: 06.02.2017].	
52	,	A1
	STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP-690,550	
	COMPARED TO METHOTREXATE IN METHOTREXATE-NAØVE PATIENTS	
	WITH RHEUMATOID ARTHRITIS; Abrufbar unter:	
	https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-	
	016987-34 [Zugriff am: 06.02.2017].	1
53	EU Clinical Trials Register. 2011. A PHASE 3, MULTI-SITE, RANDOMIZED,	A1
	MIXED-BLIND, PARALLEL-GROUP TREATMENT WITHDRAWAL AND RE-	
	TREATMENT STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES	
	OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC	
	PLAQUE PSORIASIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-	
	search/search?query=eudract_number:2010-020005-32 [Zugriff am: 06.02.2017].	
54	EU Clinical Trials Register. 2011. A PHASE 3, MULTI-SITE, RANDOMIZED,	A1
	DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF	
	THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 IN	
	SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE; Abrufbar	
	unter: https://www.clinicaltrialsregister.eu/ctr-	
	search/search?query=eudract_number:2010-020003-73 [Zugriff am: 06.02.2017].	
55	EU Clinical Trials Register. 2011. A PHASE 3, MULTI-SITE, RANDOMIZED,	A1
	DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF	
	THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 IN	
	SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE; Abrufbar	
	unter: https://www.clinicaltrialsregister.eu/ctr-	
	search/search?query=eudract_number:2010-019988-10 [Zugriff am: 06.02.2017].	
56	EU Clinical Trials Register. 2011. AN EXPLORATORY PHASE 2,	A1
	RANDOMIZED, DOUBLE-BLIND, MULTICENTER STUDY TO ASSESS THE	1
	EFFECTS OF TOFACITINIB (CP-690,550) ON MAGNETIC RESONANCE	
	IMAGING ENDPOINTS, IN METHOTREXATE NAÏVE SUBJECTS WITH	
	EARLY ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter:	
	https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-	1
	020890-18 [Zugriff am: 06.02.2017].	ļ
57		A1
	FOLLOW-UP CLINICAL DATA FROM KIDNEY TRANSPLANT RECIPIENTS	1
	WHO RECEIVED TOFACITINIB (CP-690,550) IN COMPLETED PHASE 2	1
	STUDIES; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-	
	search/search?query=eudract_number:2011-000903-41 [Zugriff am: 06.02.2017].	
58		A1
	690,550 AS MAINTENANCE THERAPY IN PATIENTS WITH CROHN'S	1
	DISEASE; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-	1
	search/search?query=eudract_number:2011-003622-27 [Zugriff am: 06.02.2017].	

9 EU Clinical Trials Register. 2012. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTRE STUDY TO INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 FOR	A1
INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 FOR	
·	
INDUCTION THED ADVIN CHILIECTS WITH MODED ATE TO SEVEDE	
INDUCTION THERAPY IN SUBJECTS WITH MODERATE TO SEVERE	
CROHN'S DISEASE; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-	
<u> </u>	
	A1
	A1
PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTRE STUDY TO	
INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 FOR	
MAINTENANCE THERAPY IN SUBJECTS WITH MODERATE TO SEVERE	
CROHN'S DISEASE; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-	
	A1
	111
	A1
E ,	Ai
	A 1
	A1
	A1
search/search?query=eudract_number:2011-004915-22 [Zugriff am: 06.02.2017].	
6 EU Clinical Trials Register. 2012. AN OPEN-LABEL MULTIPLE DOSE STUDY	A1
TO EVALUATE THE PHARMACOKINETICS, SAFETY AND TOLERABILITY	
OF CP-690,550 IN PEDIATRIC PATIENTS FROM 2 TO LESS THAN 18 YEARS	
OF AGE WITH JUVENILE IDIOPATHIC ARTHRITIS (JIA); Abrufbar unter:	
	A1
	A1
	AI
DOSES OF TOFACITINIB (CP-690,550) OR ADALIMUMAB IN SUBJECTS	I
WITH ACTIVE PSORIATIC ARTHRITIS. Paedi; Abrufbar unter:	
5	search/search?query=eudract_number:2011-001733-16 [Zugriff am: 06.02.2017].  60 EU Clinical Trials Register. 2012. A MULTICENTRE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF ORAL CP-690,550 AS A MAINTENANCE THERAPY IN SUBJECTS WITH ULCERATIVE COLITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-004580-79 [Zugriff am: 06.02.2017].  61 EU Clinical Trials Register. 2012. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTRE STUDY TO INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 FOR MAINTENANCE THERAPY IN SUBJECTS WITH MODERATE TO SEVERE CROHN'S DISEASE; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-001754-28 [Zugriff am: 06.02.2017].  62 EU Clinical Trials Register. 2012. A MULTI-CENTRE, OPEN-LABEL STUDY OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-004581-14 [Zugriff am: 03.02.2017].  63 EU Clinical Trials Register. 2012. A MULTICENTRE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF ORAL CP-690,550 AS AN INDUCTION THERAPY IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-004579-35 [Zugriff am: 06.02.2017].  64 EU Clinical Trials Register. 2012. A MULTICENTRE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF ORAL CP-690,550 AS AN INDUCTION THERAPY IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-004578-27 [Zugriff am: 06.02.2017].  64 EU Clinical Trials Register. 2012. A MULTICENTRE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF ORAL CP-690,550 AS AN INDUCTION THERAPY IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; Abrufbar unter: https://www.clinicaltrialsre

	EU Clinical Trials Register. 2013. A PHASE 2B, MULTI-SITE, RANDOMIZED, DOUBLE-BLIND, VEHICLE-CONTROLLED, PARALLEL-GROUP STUDY OF THE EFFICACY, SAFETY, LOCAL TOLERABILITY AND PHARMACOKINETICS OF 2 DOSE STRENGTHS AND 2 REGIMENS OF TOFACITINIB OINTMENT IN SUBJECTS WITH CHRONIC PLAQUE PSORIASIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-005645-20 [Zugriff am: 06.02.2017]. EU Clinical Trials Register. 2013. A PHASE 2, RANDOMIZED, DOUBLE-BLIND,	A1 A1
	PLACEBO-CONTROLLED, DOSE-RANGING STUDY OF THE EFFICACY AND SAFETY OF TOFACITINIB IN SUBJECTS WITH ACTIVE ANKYLOSING SPONDYLITIS (AS); Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search/query=eudract_number:2011-005689-39 [Zugriff am: 06.02.2017].	
	EU Clinical Trials Register. 2014. A LONG-TERM, OPEN-LABEL EXTENSION STUDY OF TOFACITINIB (CP-690,550) FOR THE TREATMENT OF PSORIATIC ARTHRITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctrsearch/search/query=eudract_number:2011-002169-39 [Zugriff am: 06.02.2017].	A1
72	EU Clinical Trials Register. 2014. PHASE 3B/4 RANDOMIZED SAFETY ENDPOINT STUDY OF 2 DOSES OF TOFACITINIB IN COMPARISON TO A TUMOR NECROSIS FACTOR (TNF) INHIBITOR IN SUBJECTS WITH RHEUMATOID ARTHRITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-003177-99 [Zugriff am: 06.02.2017].	A8
73	EU Clinical Trials Register. 2014. A PHASE 3b/4 RANDOMIZED DOUBLE BLIND STUDY OF 5 MG OF TOFACITINIB WITH AND WITHOUT METHOTREXATE IN COMPARISON TO ADALIMUMAB WITH METHOTREXATE IN SUBJECTS WITH MODERATELY TO SEVERELY ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-000358-13 [Zugriff am: 06.02.2017].	A8
74	EU Clinical Trials Register. 2014. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PHASE 2 STUDY TO ASSESS THE IMMUNE RESPONSE FOLLOWING ADMINISTRATION OF ZOSTER VACCINE TO SUBJECTS WITH RHEUMATOID ARTHRITIS RECEIVING TOFACITINIB (CP-690,550) OR PLACEBO WITH BACKGROUND METHOTREXATE TREATMENT; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-000706-34 [Zugriff am: 06.02.2017].	A3
75	EU Clinical Trials Register. 2016. Effects of JAK inhibition on rheumatoid arthritis-related comorbidities: link between bone and vascular effects (investigator-initiated grant proposal); Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search/query=eudract_number:2015-002523-26 [Zugriff am: 06.02.2017].	A6
76	EU Clinical Trials Register. 2007. A PHASE 2B, RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED, MULTICENTER STUDY TO COMPARE 6 DOSE REGIMENS OF CP 690,550 VS. PLACEBO, EACH COMBINED WITH METHOTREXATE, ADMINISTERED FOR 6 MONTHS IN THE TREATMENT OF SUBJECTS WITH ACTIVE RHEUMATOID ARTHRITIS WHO HAVE HAD AN INADEQUATE RESPONSE TO METHOTREXATE ALONE.; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2006-005036-24 [Zugriff am: 06.02.2017].	A3
77	EU Clinical Trials Register. 2008. A PHASE 2B, RANDOMIZED, DOUBLE-BLIND, PLACEBO CONTROLLED ACTIVE COMPARATOR, MULTICENTER STUDY TO COMPARE 5 DOSE REGIMENS OF CP-690,550 AND ADALIMUMAB VERSUS PLACEBO, ADMINISTERED FOR 6 MONTHS IN THE TREATMENT OF SUBJECTS WITH ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2007-002066-35 [Zugriff am: 06.02.2017].	A5

78	EU Clinical Trials Register. 2009. PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY OF THE SAFETY AND EFFICACY OF 2 DOSES OF CP-690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND METHOTREXATE WITH INADEQUATE RESPONSE TO TNF INHIBITORS; Abrufbar unter:	A3
79	https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-014296-40 [Zugriff am: 06.02.2017].  EU Clinical Trials Register. 2009. PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY OF THE SAFETY AND EFFICACY OF 2	
	DOSES OF CP-690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND DMARDS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-008337-11 [Zugriff am: 06.02.2017].	A3
80	EU Clinical Trials Register. 2009. PHASE 3 RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP 690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND METHOTREXATE; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-007023-26 [Zugriff am: 06.02.2017].	А3
81	EU Clinical Trials Register. 2009. PHASE 3, RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP 690,550 MONOTHERAPY IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-007788-17 [Zugriff am: 06.02.2017].	А3
82	ICTRP. 2012. A PHASE 2A, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO COMPARE 3 DOSE LEVELS OF CP 690,550 VERSUS PLACEBO, ADMINISTERED ORALLY TWICE DAILY (BID) FOR 6 WEEKS, IN THE TREATMENT OF THE SIGNS AND SYMPTOMS OF SUBJECTS WITH ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
83	ICTRP. 2012. ESTUDIO OBSERVACIONAL PROSPECTIVO PARA EVALUAR LA SEGURIDAD A LARGO PLAZO Y EL ESTADO FUNCIONAL DE SUJETOS CON ARTRITIS REUMATOIDE INCLUIDOS PREVIAMENTE EN ESTUDIOS DE CP-690,550 A PROSPECTIVE OBSERVATIONAL STUDY TO EVALUATE LONG TERM SAFETY AND FUNCTIONAL STATUS OF SUBJECTS WITH RHEUMATOID ARTHRITIS PREVIOUSLY ENROLLED IN STUDIES OF CP-690,550; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6
84	ICTRP. 2012. A PHASE 2 RANDOMIZED, MULTICENTER, ACTIVE COMPARATOR-CONTROLLED TRIAL TO EVALUATE THE SAFETY AND EFFICACY OF COADMINISTRATION OF CP-690,550 AND MYCOPHENOLATE MOFETIL / MYCOPHENOLATE SODIUM IN DE NOVO KIDNEY ALLOGRAFT RECIPIENTS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
85	ICTRP. 2012. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTER STUDY TO INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CROHN'S DISEASE; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
86	ICTRP. 2012. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTER STUDY TO INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1

	ICTRP. 2012. AN OBSERVATIONAL STUDY TO COLLECT FOLLOW-UP CLINICAL DATA FROM KIDNEY TRANSPLANT RECIPIENTS WHO RECEIVED TOFACITINIB (CP-690,550) IN COMPLETED PHASE 2 STUDIES; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
88	ICTRP. 2013. PHASE 3 RANDOMIZED, DOUBLE-BLIND STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP-690,550 COMPARED TO METHOTREXATE IN METHOTREXATE-NAIVE PATIENTS WITH RHEUMATOID ARTHRITIS; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
89	ICTRP. 2013. A PHASE 3, MULTI-SITE, RANDOMIZED, MIXED-BLIND, PARALLEL-GROUP TREATMENT WITHDRAWAL AND RE-TREATMENT STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
90	ICTRP. 2014. A PHASE 3, MULTI SITE, RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED, PARALLEL GROUP STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 AND 1 SUBCUTANEOUS DOSE OF ETANERCEPT IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
91	DOSES OF TOFACITINIB IN COMPARISON TO A TUMOR NECROSIS FACTOR (TNF) INHIBITOR IN SUBJECTS WITH RHEUMATOID ARTHRITIS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A8
92	ICTRP. 2015. A PHASE 3, MULTI-SITE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
93	ICTRP. 2015. A PHASE 3, MULTI-SITE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
94	ICTRP. 2015. A 6-Month, Phase 2, Multicenter, Randomized, Open-Label, Comparative Study Of 2 Dose Levels Of CP-690,550 Administered Concomitantly With IL-2 Receptor Antagonist Induction Therapy, Mycophenolate Mofetil And Corticosteroids Versus A Tacrolimus-Based Immunosuppressive Regimen For The Prevention Of Allograft Rejection In De Novo Renal Allograft Recipients; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
95	ICTRP. 2015. A Phase 2a, Randomized, Double-Blind, Placebo-Controlled, Multicenter Study To Compare 3 Dose Levels Of CP-690,550 Versus Placebo, Administered Orally Twice Daily (BID) For 6 Weeks, In The Treatment Of The Signs And Symptoms Of Subjects With Active Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
96	ICTRP. 2015. A Multicenter, Phase 2, Open-Label, Controlled, Extension Study For Stage 1 Subjects Of Study A3921009 To Evaluate The Long-Term Safety And Efficacy Of CP-690,550 Versus Tacrolimus, When Co-Administered With Mycophenolate Mofetil In Renal Allograft Recipients; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
97	ICTRP. 2015. A Prospective Observational Study To Evaluate Long-Term Safety And Functional Status Of Subjects With Rheumatoid Arthritis Previously Enrolled In Studies Of CP-690,550; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6

98	ICTRP. 2015. A Phase 2 Randomized, Multicenter, Active Comparator-Controlled Trial to Evaluate the Safety and Efficacy of Co-administration of CP-690,550 and Mycophenolate Mofetil / Mycophenolate Sodium in De Novo Renal Allograft Recipients; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
99	ICTRP. 2015. A Pilot Randomized, Multicenter, Open-Label, Active Comparator-Controlled Trial Of A CP-690,550-Based Regimen In De Novo Kidney Allograft Recipients; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
100	ICTRP. 2015. A Randomized, Double-Blind, Placebo-Controlled, Parallel Group, Multi-Center Study To Investigate The Safety And Efficacy Of CP-690,550 In Subjects With Moderate To Severe Crohn's Disease; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
101	ICTRP. 2015. A Phase 2, Multicenter, Open-label, Active Comparator-controlled, Extension Trial To Evaluate The Long-term Safety And Efficacy Of Cp-690,550 In Renal Allograft Recipients; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
102	ICTRP. 2015. A Long-Term, Open-Label Study Of CP-690,550 To Confirm The Safety Following Long Term Administration Of CP-690,550 In The Treatment Of Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6
103	ICTRP. 2015. A Phase 2B, Multicenter, Randomized, Double-Blind, Parallel-Group, Placebo-Controlled Trial Evaluating The Efficacy And Safety Of Dose Regimens With Oral CP-690,550 In The Treatment Of Subjects With Moderate To Severe Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
104	ICTRP. 2015. Phase 2A Randomized, Double-Blind, Vehicle-Controlled, Intra-Individual Comparison Trial Assessing Safety, Toleration, Pharmacokinetics And Pilot Efficacy Of 4 Weeks Treatment With Topical CP-690,550 In Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
105	ICTRP. 2015. A Phase I/II Prospective, Randomized, Double Masked, Vehicle And Comparator Controlled, Dose Ranging Study Of CP-690,550 In Subjects With Dry Eye Disease; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
106	ICTRP. 2015. A Phase I, Open Label, Fixed-Sequence Study to Estimate the Effect of Tacrolimus and Cyclosporine on the Pharmacokinetics of CP-690,550 in Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
107	ICTRP. 2015. A Phase 1, Randomized, 2-Way Crossover, Multiple Dose, Open Label Study Of The Effect Of CP-690,550 On Midazolam Pharmacokinetics In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
108	ICTRP. 2015. A Phase 1, Non-Randomized, Open-Label, Single-Dose Study to Evaluate the Pharmacokinetics, Safety and Tolerability of CP-690,550 in Subjects With Hepatic Impairment and Normal Hepatic Function; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
109	ICTRP. 2015. An Exploratory Phase 2a, Randomized, Double-Blind, Placebo-Controlled, Multicenter Study To Assess The Pharmacodynamics Of CP-690,550, Administered Orally Twice Daily (BID) For 4 Weeks, In Subjects With Active Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A5
110	ICTRP. 2015. A Study In Trained Taste Panel Healthy Adult Volunteers To Investigate The Palatability Of Select Formulations Of CP-690,550 Oral Solution; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
111	ICTRP. 2015. Phase 3 Randomized, Double-Blind Study Of The Efficacy And Safety Of 2 Doses Of CP-690,550 Compared To Methotrexate In Methotrexate-Naive Patients With Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1

112	ICTRP. 2015. Phase 2 Study Of The Effects Of Open-Label CP-690,550 And Double-Blind Atorvastatin On Lipids In Patients With Active Rheumatoid Arthritis;	A1
113	Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].  ICTRP. 2015. An Open Label, Single And Multiple Dose Study To Investigate The Pharmacokinetics, Safety And Tolerability Of CP-690,550 In Healthy Chinese Subjects; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
114		A1
115	ICTRP. 2015. A Phase 1, Randomized, Open Label, 2-Way Crossover Study To Assess The Effect Of Multiple Dose CP-690,550 On The Pharmacokinetics Of Single Dose Oral Contraceptive Steroids In Healthy Female Subjects; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
116	ICTRP. 2015. A Phase 1, Open Label, Single Dose, Randomized, Cross Over Study To Estimate The Absolute Oral Bioavailability Of CP-690,550 In Healthy Subjects; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
117		A1
118	ICTRP. 2015. A Phase 1, Randomized, 2-Period, 2-Sequence, Open Label, Single Dose, Cross-Over Study To Evaluate The Effect Of Food On Pharmacokinetics Of Tasocitinib (CP-690,550) Tablets In Healthy Subjects; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
119	ICTRP. 2015. Phase 1, Open-Label, Randomized, Single-Dose, 3-Treatment, 3-Period, Cross-Over, Bioequivalence Study Comparing Phase 2b, Phase 3 And Commercial Image Tablet Formulations Of Tasocitinib (CP-690,550) Under Fasted Conditions; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
120	ICTRP. 2015. A Phase 1, Randomized, 3-Period, Open Label, Single Dose, Cross Over Study To Evaluate The Pharmacokinetics And Safety Of Two Controlled Release Formulations Of CP-690,550; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
121	ICTRP. 2015. A Phase 3, Multi-Site, Randomized, Mixed-Blind, Parallel-Group Treatment Withdrawal And Re-Treatment Study Of The Efficacy And Safety Of 2 Oral Doses Of CP-690,550 In Subjects With Moderate To Severe Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
122	ICTRP. 2015. A Phase 1, Open Label, Fixed-Sequence Study To Estimate The Effect Of Ketoconazole On The Pharmacokinetics Of Tasocitinib (CP-690,550) In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
123	ICTRP. 2015. A Phase 1, Open Label, Fixed-Sequence Study To Estimate The Effect Of Repeat-Dose Rifampin On The Pharmacokinetics Of CP-690,550 In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
124	ICTRP. 2015. A Phase II, Randomized, Double Masked, Parallel Group, Vehicle Controlled, Multiple-Dose Study Of Tasocitinib (CP-690,550) In Subjects With Dry Eye Disease; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
125	ICTRP. 2015. A Phase 3, Multi Site, Randomized, Double Blind, Placebo Controlled, Parallel Group Study Of The Efficacy And Safety Of 2 Oral Doses Of CP- 690,550 And 1 Subcutaneous Dose Of Etanercept In Subjects With Moderate To Severe Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	Al

126	ICTRP. 2015. A Phase 2a, Multi Site, Randomized, Double Blind, Placebo Controlled, Parallel Group Study Of The Pilot Efficacy, Safety, And Pharmacokinetics Of 2 Ointment Formulations Of CP-690,550 In Subjects With Mild To Moderate Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
127	ICTRP. 2015. An Exploratory Phase 1, Fixed Sequence, Open-Label Study To Assess The Effects Of CP-690,550 On The Kinetics Of Cholesterol Flux Through The High Density Lipoprotein/Reverse Cholesterol Transport Pathway In Patients With Active Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
128	ICTRP. 2015. Phase 3 Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study Of The Efficacy And Safety Of 2 Oral Doses Of CP-690,550 In Subjects With Moderate To Severe Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
129	ICTRP. 2015. Phase 1, Open-label, Randomized, Single-dose, 2-treatment, 2-period, Crossover, Bioequivalence Study Comparing Phase 2b 1 Mg And 5 Mg Tablet Formulations Of Cp-690,550 Under Fasted Conditions; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
130	ICTRP. 2015. A Phase 3, Multi-Site, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study Of The Efficacy And Safety Of 2 Oral Doses Of CP-690,550 In Subjects With Moderate To Severe Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
131	ICTRP. 2015. A Randomized, Double Blind, Placebo Controlled Phase 2 Study To Assess The Immune Response Following Administration Of Influenza And Pneumococcal Vaccines To Subjects With Rheumatoid Arthritis Receiving Cp-690,550 Or Placebo Cp-690,550 With And Without Background Methotrexate; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
132	ICTRP. 2015. An Observational Study To Collect Follow-up Clinical Data From Kidney Transplant Recipients Who Received Tofacitinib (CP-690,550) In Completed Phase 2 Studies; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
133	ICTRP. 2015. A Randomized, Double-blind, Placebo-controlled, Parallel Group, Multi-centre Study To Investigate The Safety And Efficacy Of Cp-690,550 For Induction Therapy In Subjects With Moderate To Severe Crohn's Disease; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
134	ICTRP. 2015. A Randomized, Double-blind, Placebo-controlled, Parallel Group, Multi-centre Study To Investigate The Safety And Efficacy Of Cp-690,550 For Maintenance Therapy In Subjects With Moderate To Severe Crohn's Disease; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
135	ICTRP. 2015. A Phase 1, Open Labeled, Fixed Sequence Study To Estimate The Effect Of CP-690,550 On The Pharmacokinetics Of Metformin In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
136	ICTRP. 2015. A Multicentre, Randomized, Double-blind, Placebo-controlled, Parallel-group Study Of Oral Cp-690,550 As An Induction Therapy In Subjects With Moderate To Severe Ulcerative Colitis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
137	ICTRP. 2015. A Multicentre, Randomized, Double-blind, Placebo-controlled, Parallel-group Study Of Oral Cp-690,550 As An Induction Therapy In Subjects With Moderate To Severe Ulcerative Colitis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
138	ICTRP. 2015. A Phase 1, Randomized, Placebo-Controlled, Two-Period, Fixed Sequence Study To Evaluate The Effect Of CP-690,550 On Measured Glomerular Filtration Rate In Patients With Active Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A5

139	ICTRP. 2015. A Phase 1, Randomized, Open Label, Partial Crossover Study To Evaluate The Pharmacokinetics (PK) And Safety Of Three Modified Release (MR) And One Immediate Release (IR) Formulations Of Tofacitinib (CP-690,550) In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
140	ICTRP. 2015. A Phase 3, Multi Site, Randomized, Double Blind Study Of The Longterm Safety, Tolerability And Efficacy Of 2 Oral Doses Of Cp 690,550 In Subjects With Moderate To Severe Plaque Psoriasis And/or Psoriatic Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
141	ICTRP. 2015. Phase 1, Open-Label, Randomized, Single-Dose, Cross-Over, Bioequivalence Study Comparing Tablets (5 And 10 Mg) And Capsules (5 And 10 Mg) Of Tofacitinib Under Fasted Conditions In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
142	ICTRP. 2015. Phase 1, Open-Label Study Of The Pharmacokinetics, Non-Renal Clearance And Dialyzability Of CP-690,550 In Subjects With End-Stage Renal Disease Undergoing Hemodialysis; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
143	ICTRP. 2015. Phase 1, Placebo-controlled, Randomized, Sequential, Parallel-group, Dose Escalation Study to Evaluate 28-day Multiple Dose Safety, Tolerability, Pharmacokinetics and Pharmacodynamics of CP-690,550 in Stable Renal Allograft Recipients; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
144	ICTRP. 2015. An Exploratory Phase 2a, Randomized, Double-blind, Placebocontrolled, Multicenter Study To Assess Mechanism Of Action (Moa) Of Cp-690,550 In The Skin When Administered Orally At 10 Mg Twice Daily (Bid) For 12 Weeks In Subjects With Moderate To Severe Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
145	ICTRP. 2015. A Phase 1, Randomized, Open-Label, 2-Way Crossover Study To Evaluate The Pharmacokinetics (PK), Safety, And Bioavailability Of Tofacitinib Following Single Oral Dose Of MR 11 mg Compared To MR 22 mg In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
146	ICTRP. 2015. Phase 1, Investigator-Blind, Subject-Blind, Sponsor-Open, Placebo-Controlled, Two-Week, Multiple Dose Escalation Study In Medically Stable Subjects With Psoriasis To Evaluate The Safety, Tolerability, Pharmacokinetics And Pharmacodynamics Of CP-690,550; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
147	ICTRP. 2015. Phase 1, Open-Label Study To Evaluate Single Dose Pharmacokinetics, Safety And Tolerability Of CP-690,550 In Patients With Impaired Renal Function; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
	ICTRP. 2015. A Study in Healthy Adult Volunteers and Adult Subjects With Rheumatoid Arthritis to Evaluate the Safety, Tolerability, and Pharmacokinetics After Multiple Dosing of ABT-494; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
	ICTRP. 2015. A Phase 1, Randomized, Placebo- And Positive-Controlled Crossover Study To Determine The Effect Of Single-Dose CP-690,550 On QTc Interval In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
150	ICTRP. 2015. A Phase 1, Open Label Study Of The Pharmacokinetics Of Multiple Doses Of Oral CP-690,550 And Single Doses Of Oral Methotrexate In Rheumatoid Arthritis Subjects; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6

151	ICTRP. 2015. A Phase 2, Randomized, Double-blind, Placebo-controlled, Doseranging Study Of The Efficacy And Safety Of Tofacitinib In Subjects With Active Ankylosing Spondylitis (as); Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
152		A1
153	ICTRP. 2015. A Phase 2b, Multi-site, Randomized, Double-blind, Vehicle-controlled, Parallel-group Study Of The Efficacy, Safety, Local Tolerability And Pharmacokinetics Of 2 Dose Strengths And 2 Regimens Of Tofacitinib Ointment In Subjects With Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
154	ICTRP. 2015. A Phase 2a, Multi-site, Randomized, Double-blind, Vehicle-controlled, Parallel-group Study Of The Pilot Efficacy, Safety, Tolerability, And Pharmacokinetics Of 2% Tofacitinib Ointment In Subjects With Mild To Moderate Atopic Dermatitis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
155		A1
156	ICTRP. 2015. A Randomized, Double-blind, Placebo-controlled, Phase 2 Study To Assess The Immune Response Following Administration Of Zoster Vaccine To Subjects With Rheumatoid Arthritis Receiving Tofacitinib (Cp-690,550) Or Placebo With Background Methotrexate Treatment; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
157	ICTRP. 2015. A Phase 1, Single- Center, Randomized, Double-blind, Vehicle And Active Comparator-controlled Trial To Evaluate The Antipsoriatic Activity And Safety Of A Topically Applied Pf-06263276 Formulation In A Psoriasis Plaque Test; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
158	ICTRP. 2015. Tofacitnib for the Treatment of Alopecia Areata and Variants; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
159	ICTRP. 2015. An Open-Label Pilot Study To Evaluate The Efficacy Of Tofacitinib In Moderate To Severe Alopecia Areata, Totalis And Universalis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
160	ICTRP. 2015. A Phase 1, Randomized, Open Label, Single Dose, 2 Period Crossover Study To Evaluate The Effect Of Food On The Pharmacokinetics Of Tofacitinib Modified Release (mr) 22 Mg Tablets In Healthy Volunteers; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
161	ICTRP. 2015. An Evaluation of the Optimal Dose of Tofacitinib Needed to Achieve LDA or Clinical Remission in Patients With Active Rheumatoid Arthritis (RA) as Measured From a Clinical and Structural Perspective When Treating to Target; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6
162	ICTRP. 2016. Effects of JAK inhibition on rheumatoid arthritis-related comorbidities: link between bone and vascular effects (investigator-initiated grant proposal); Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6
163	ICTRP. 2016. Clinical study on drug holiday with regulated rheumatoid arthritis activity after treatment with tofacitinib (Xanadu study); Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A2

164	ICTRP. 2016. In methotrexate resistant rheumatoid arthritis, switching methotrexate to tofacitinib versus adding tofacitinib to methotrexate, open labeled, multicenter randomized non-inferiority study; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A8
165	ICTRP. 2016. A clinical study to evaluate the correlation between the efficacy and the immunological markers in patients with rheumatoid arthritis treated with Tofacitinib; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6
166	ICTRP. 2016. A Long-term, Open-label Follow-up Study Of Tofacitinib (Cp-690,550) For Treatment Of Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6
167	ICTRP. 2016. A Phase 3, Multi-Site, Open-Label Study Of The Long Term Safety And Tolerability Of 2 Oral Doses Of CP-690,550 In Subjects With Moderate To Severe Chronic Plaque Psoriasis; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A1
168	ICTRP. 2016. A Multicentre, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study Of Oral CP-690,550 As A Maintenance Therapy In Subjects With Ulcerative Colitis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
169	ICTRP. 2016. A Open-label Extension Study Of CP-690,550 As Maintenance Therapy In Patients With Crohn's Disease; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
170	ICTRP. 2016. A Multi-Center, Open-Label Study Of CP-690,550 In Subjects With Moderate To Severe Ulcerative Colitis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
171	ICTRP. 2016. A Long-Term, Open-Label Follow-Up Study Of Tofacitinib For Treatment Of Juvenile Idiopathic Arthritis (JIA); Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
172	ICTRP. 2016. An Open-label Multiple Dose Study To Evaluate The Pharmacokinetics, Safety And Tolerability Of Cp-690,550 In Pediatric Patients From 2 To Less Than 18 Years Of Age With Juvenile Idiopathic Arthritis (Jia); Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
173	ICTRP. 2016. A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study Of The Efficacy And Safety Of 2 Doses Of Tofacitinib (CP-690,550) Or Adalimumab In Subjects With Active Psoriatic Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
174	ICTRP. 2016. A Phase 3, Randomized, Double-blind, Placebo-controlled Study Of The Efficacy And Safety Of 2 Doses Of Tofacitinib (Cp-690,550) In Subjects With Active Psoriatic Arthritis And An Inadequate Response To At Least One Tnf Inhibitor; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
175	ICTRP. 2016. Tofacitinib Special Investigation Of Xeljanz Tablets 5mg. (Regulatory Post Marketing Commitment Plan); Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A6
176	ICTRP. 2016. A Long Term, Open-label Extension Study Of Tofacitinib (Cp-690,550) For The Treatment Of Psoriatic Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
177	ICTRP. 2016. A Phase 3b/4 Randomized Double Blind Study Of 5 Mg Of Tofacitinib With And Without Methotrexate In Comparison To Adalimumab With Methotrexate In Subjects With Moderately To Severely Active Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A8

178	ICTRP. 2016. A Multicenter, Randomized, Double-blind, Parallel-group, Phase 3 Study To Demonstrate Non-inferiority For The Efficacy Of A Once Daily Dose Of Tofacitinib Modified Release Tablet To A Twice Daily Dose Of The Immediate Release Tablet In Adult Patients With Rheumatoid Arthritis On Background Methotrexate; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
179	ICTRP. 2016. Tofacitinib for the Treatment of Alopecia Areata and Its Variants; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
180	ICTRP. 2016. Musculoskeletal Ultrasound Assessment of Therapeutic Response of Tofacitinib in Rheumatoid Arthritis Patients; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A6
181	ICTRP. 2016. Safety of Tofacitinib, an Oral Janus Kinase Inhibitor, in Systemic Lupus Erythematosus; a Phase Ib Clinical Trial and Associated Mechanistic Studies; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
182	ICTRP. 2016. Efficacy, Safety And Tolerability Of Tofacitinib For Treatment Of Polyarticular Course Juvenile Idiopathic Arthritis (Jia) In Children And Adolescent Subjects; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
183	ICTRP. 2012. PHASE 3 RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP 690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND METHOTREXATE; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
184	ICTRP. 2012. PHASE 3, RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP 690,550 MONOTHERAPY IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
185	ICTRP. 2012. PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY OF THE SAFETY AND EFFICACY OF 2 DOSES OF CP-690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND METHOTREXATE WITH INADEQUATE RESPONSE TO TNF INHIBITORS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	А3
186	ICTRP. 2012. A PHASE 2B, RANDOMIZED, DOUBLE-BLIND, PLACEBO CONTROLLED ACTIVE COMPARATOR, MULTICENTER STUDY TO COMPARE 5 DOSE REGIMENS OF CP-690,550 AND ADALIMUMAB VERSUS PLACEBO, ADMINISTERED FOR 6 MONTHS IN THE TREATMENT OF SUBJECTS WITH ACTIVE RHEUMATOID ARTHRITIS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A5
187	ICTRP. 2012. A PHASE 2B, RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED, MULTICENTER STUDY TO COMPARE 6 DOSE REGIMENS OF CP 690,550 VS. PLACEBO, EACH COMBINED WITH METHOTREXATE, ADMINISTERED FOR 6 MONTHS IN THE TREATMENT OF SUBJECTS WITH ACTIVE RHEUMATOID ARTHRITIS WHO HAVE HAD AN INADEQUATE RESPONSE TO METHOTREXATE ALONE; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	А3
188	ICTRP. 2012. PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY OF THE SAFETY AND EFFICACY OF 2 DOSES OF CP-690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND DMARDS; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
189	ICTRP. 2015. A Phase 2B, Randomized, Double Blind, Placebo-Controlled, Multicenter Study To Compare 6 Dose Regimens Of CP-690,550 Vs. Placebo, Each Combined With Methotrexate, Administered For 6 Months In The Treatment Of Subjects With Active Rheumatoid Arthritis Who Have Had An Inadequate Response To Methotrexate Alone; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	А3

190	ICTRP. 2015. A Phase 2b, Randomized, Double Blind, Placebo Controlled Active Comparator, Multicenter Study To Compare 5 Dose Regimens Of CP- 690,550 And Adalimumab Versus Placebo, Administered For 6 Months In The Treatment Of Subjects With Active Rheumatoid Arthritis; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A5
191	ICTRP. 2015. A Phase 2, Randomized, Double-blind, Placebo-controlled, Multicenter Study to Confirm Dose Responsiveness Following 12 Weeks of the Administration of CP-690,550 (4 Doses) or Placebo in Subjects With Active Rheumatoid Arthritis Inadequately Controlled With Methotrexate Alone; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
192	ICTRP. 2015. A Phase 2, Randomized, Double-Blind, Placebo-Controlled, Multicenter Study To Confirm Dose Responsiveness Following 12 Weeks Of The Administration Of CP-690,550 (5 Doses) Or Placebo In Subjects With Active Rheumatoid Arthritis Inadequately Responding To At Least 1 DMARD; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
193		A3
194	ICTRP. 2015. Phase 3 Randomized, Double Blind, Placebo Controlled Study Of The Efficacy And Safety Of 2 Doses Of CP-690,550 In Patients With Active Rheumatoid Arthritis On Background Methotrexate; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
195	ICTRP. 2015. Phase 3, Randomized, Double Blind, Placebo Controlled Study Of The Safety And Efficacy Of 2 Doses Of CP 690,550 In Patients With Active Rheumatoid Arthritis On Background DMARDS; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A3
196	ICTRP. 2015. Phase 3, Randomized, Double-Blind, Placebo-Controlled Study Of The Safety And Efficacy Of 2 Doses Of CP-690,550 In Patients With Active Rheumatoid Arthritis On Background Methotrexate With Inadequate Response To TNF Inhibitors; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A3
197	ICTRP. 2016. efficacy of aterosclerosis of tofacitinib in rheumatoid arthritis Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
198	ICTRP. 2017. Topical Tofacitinib for the Treatment of Alopecia Areata and Its Variants; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
199	ICTRP. 2017. A Phase 3b/4 Randomized Double Blind Placebo Controlled Study Of Methotrexate (Mtx) Withdrawal In Subjects With Rheumatoid Arthritis (ra) Treated With Tofacitinib 11mg Modified Release (mr) Formulation. A; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A2
200	ICTRP. 2017. A 12 Week Study To Assess The Efficacy And Safety Profile Of Pf-06650833 In Subjects With Active Rheumatoid Arthritis With An Inadequate Response To Methotrexate. Irak, B; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A2
201	Efficacy And Safety Of CP-690,550 In Children From 2 To Less Than 18 Years Old With Systemic Juvenile Idiopathic Arthritis (sJIA) With Active Systemic Features. A; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
202	ICTRP. 2017. Study of Tofacitinib in Refractory Dermatomyositis (STIR): Proof of Concept, Open-Label Study of 10 Patients. Irb; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A1
203	ICTRP. 2017. Effectiveness and Safety of Tofacitinib in Korean Patients With Rheumatoid Arthritis. Huhrd SPE; Abrufbar unter: http://apps.who.int/trialsearch/[Zugriff am: 03.02.2017].	A8

204	ICTRP. 2017. Safety and Tolerability of Herpes Zoster Vaccine in Patients With Rheumatoid Arthritis Immunized Prior to Biologics and Tofacitinib Therapy Initiation. C. EM; Abrufbar unter: http://apps.who.int/trialsearch/ [Zugriff am: 03.02.2017].	A8
205		A6
206	PharmNet.Bund. 2006. A LONG-TERM, OPEN-LABEL FOLLOW-UP STUDY OF TOFACITINIB (CP-690,550) FOR TREATMENT OF RHEUMATOID ARTHRITIS; [Zugriff am: 06.02.2017].	A6
207	PharmNet.Bund. 2007. A PHASE 2 RANDOMIZED, MULTICENTER, ACTIVE COMPARATOR-CONTROLLED TRIAL TO EVALUATE THE SAFETY AND EFFICACY OF COADMINISTRATION OF CP-690,550 AND MYCOPHENOLATE MOFETIL / MYCOPHENOLATE SODIUM IN DE NOVO KIDNEY ALLOGRAFT RECIPIENTS; [Zugriff am: 06.02.2017].	A1
208	PharmNet.Bund. 2008. A PHASE 2, MULTICENTER, OPEN-LABEL, ACTIVE COMPARATOR CONTROLLED, EXTENSION TRIAL TO EVALUATE THE LONG-TERM SAFETY AND EFFICACY OF CP-690,550 IN RENAL ALLOGRAFT RECIPIENTS; [Zugriff am: 06.02.2017].	A1
209	PharmNet.Bund. 2009. PHASE 3 RANDOMIZED, DOUBLE-BLIND STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP-690,550 COMPARED TO METHOTREXATE IN METHOTREXATE-NAÏVE PATIENTS WITH RHEUMATOID ARTHRITIS; [Zugriff am: 06.02.2017].	A1
210	PharmNet.Bund. 2010. A PHASE 3, MULTI-SITE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS; [Zugriff am: 06.02.2017].	A1
211	PharmNet.Bund. 2010. A PHASE 3, MULTI-SITE, OPEN-LABEL STUDY OF THE LONG TERM SAFETY AND TOLERABILITY OF 2 ORAL DOSES OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS; [Zugriff am: 06.02.2017].	A1
212	PharmNet.Bund. 2010. A PHASE 3, MULTI-SITE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS; [Zugriff am: 06.02.2017].	A1
213	PharmNet.Bund. 2010. A PHASE 3, MULTI SITE, RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED, PARALLEL GROUP STUDY OF THE EFFICACY AND SAFETY OF 2 ORAL DOSES OF CP-690,550 AND 1 SUBCUTANEOUS DOSE OF ETANERCEPT IN SUBJECTS WITH MODERATE TO SEVERE CHRONIC PLAQUE PSORIASIS; [Zugriff am: 06.02.2017].	A1
214	PharmNet.Bund. 2011. A LONG-TERM, OPEN-LABEL EXTENSION STUDY OF TOFACITINIB(CP-690,550) FOR THE TREATMENT OF PSORIATIC ARTHRITIS; [Zugriff am: 06.02.2017].	A1
215	PharmNet.Bund. 2011. A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDYOF THE EFFICACY AND SAFETY OF 2 DOSES OF TOFACITINIB (CP-690,550)OR ADALIMUMAB IN SUBJECTS WITH ACTIVE PSORIATIC ARTHRITIS.Paediatric investigation plan numbers - (P/144/2010),(P/162/2011) and(P/0064/2012); [Zugriff am: 06.02.2017].	A1
216	PharmNet.Bund. 2011. A LONG-TERM, OPEN-LABEL FOLLOW-UP STUDY OF CP-690,550 FOR TREATMENT OF JUVENILE IDIOPATHIC ARTHRITIS (JIA); [Zugriff am: 06.02.2017].	A1

217	PharmNet.Bund. 2011. AN OPEN-LABEL MULTIPLE DOSE STUDY TO EVALUATE THE PHARMACOKINETICS, SAFETY AND TOLERABILITY OF CP-690,550 IN PEDIATRIC PATIENTS FROM 2 TO LESS THAN 18 YEARS OF AGE WITH JUVENILE IDIOPATHIC ARTHRITIS (JIA); [Zugriff am: 06.02.2017].	A1
218	PharmNet.Bund. 2011. A MULTICENTRE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF ORAL CP-690,550 AS AN INDUCTION THERAPY IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; [Zugriff am: 06.02.2017].	A1
219	PharmNet.Bund. 2011. A OPEN-LABEL EXTENSION STUDY OF CP-690,550 AS MAINTENANCE THERAPY IN PATIENTS WITH CROHN'S DISEASE; [Zugriff am: 06.02.2017].	A1
220	PharmNet.Bund. 2011. A MULTI-CENTER, OPEN-LABEL STUDY OF CP-690,550 IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; [Zugriff am: 06.02.2017].	A1
221	PharmNet.Bund. 2011. A MULTICENTRE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF ORAL CP-690,550 AS A MAINTENANCE THERAPY IN SUBJECTS WITH ULCERATIVE COLITIS; [Zugriff am: 06.02.2017].	A1
222	PharmNet.Bund. 2011. A MULTICENTRE, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF ORAL CP-690,550 AS AN INDUCTION THERAPY IN SUBJECTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS; [Zugriff am: 06.02.2017].	A1
223	PharmNet.Bund. 2011. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLELGROUP, MULTI-CENTRE STUDY TO INVESTIGATE THE SAFETY ANDEFFICACY OF CP-690,550 FOR INDUCTION THERAPY IN SUBJECTS WITH MODERATE TO SEVERE CROHN'S DISEASE; [Zugriff am: 06.02.2017].	A1
224	PharmNet.Bund. 2011. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTRE STUDY TO INVESTIGATE THE SAFETY AND EFFICACY OF CP-690,550 FOR MAINTENANCE THERAPY IN SUBJECTS WITH MODERATE TO SEVERE CROHN'S DISEASE; [Zugriff am: 06.02.2017].	A1
225	PharmNet.Bund. 2011. AN OBSERVATIONAL STUDY TO COLLECT FOLLOW-UP CLINICAL DATAFROM KIDNEY TRANSPLANT RECIPIENTS WHO RECEIVED TOFACITINIB(CP-690,550) IN COMPLETED PHASE 2 STUDIES; [Zugriff am: 06.02.2017].	A1
226	PharmNet.Bund. 2012. A PHASE 2B, MULTI-SITE, RANDOMIZED, DOUBLE-BLIND, VEHICLE-CONTROLLED, PARALLEL-GROUP STUDY OF THE EFFICACY, SAFETY, LOCAL TOLERABILITY AND PHARMACOKINETICS OF 2 DOSE STRENGTHS AND 2 REGIMENS OF TOFACITINIB OINTMENT IN SUBJECTS WITH CHRONIC PLAQUE PSORIASIS; [Zugriff am: 06.02.2017].	A1
227	PharmNet.Bund. 2013. A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF TOFACITINIB (CP-690,550) IN SUBJECTS WITH ACTIVE PSORIATIC ARTHRITIS AND AN INADEQUATE RESPONSE TO AT LEAST ONE TNF INHIBITOR.Paediatric investigation plan numbers - (P/144/2010),(P/162/2011) and (P/0064/2012); [Zugriff am: 06.02.2017].	A1
228	PharmNet.Bund. 2014. A PHASE 3b/4 RANDOMIZED DOUBLE BLIND STUDY OF 5 MG OF TOFACITINIB WITH AND WITHOUT METHOTREXATE IN COMPARISON TO ADALIMUMAB WITH METHOTREXATE IN SUBJECTS WITH MODERATELY TO SEVERELY ACTIVE RHEUMATOID ARTHRITIS; [Zugriff am: 06.02.2017].	A8

229	PharmNet.Bund. 2009. PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY OF THE SAFETY AND EFFICACY OF 2 DOSES OF CP-690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND METHOTREXATE WITH INADEQUATE RESPONSE TO TNF INHIBITORS; [Zugriff am: 06.02.2017].	A3
230	PharmNet.Bund. 2008. PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY OF THE SAFETY AND EFFICACY OF 2 DOSES OF CP-690,550 IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS ON BACKGROUND DMARDS; [Zugriff am: 06.02.2017].	A3
231	PharmNet.Bund. 2008. PHASE 3, RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED STUDY OF THE EFFICACY AND SAFETY OF 2 DOSES OF CP 690,550 MONOTHERAPY IN PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS; [Zugriff am: 06.02.2017].	A3
232	PharmNet.Bund. 2007. A PHASE 2B, RANDOMIZED, DOUBLE-BLIND, PLACEBO CONTROLLED ACTIVE COMPARATOR, MULTICENTER STUDY TO COMPARE 5 DOSE REGIMENS OF CP-690,550 AND ADALIMUMAB VERSUS PLACEBO, ADMINISTERED FOR 6 MONTHS IN THE TREATMENT OF SUBJECTS WITH ACTIVE RHEUMATOID ARTHRITIS; [Zugriff am: 06.02.2017].	A5

# Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht relevant, da keine indirekten Vergleiche durchgeführt wurden.

# Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht relevant, da keine nicht randomisierten vergleichenden Studien in die Ergebnisdarstellung dieses Dossiers einbezogen wurden.

# Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht relevant, da keine weiteren Untersuchungen dargestellt wurden.

## Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-194 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 4-194 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-194 (Anhang): Studiendesign und -methodik für Studie A3921064 (ORAL STANDARD)

Item <sup>a</sup>	Charakteristikum	Studieninformation	
Studie	Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	primäre Ziele:	
		Vergleich der Wirksamkeit von beiden CP-690,550-Dosierungen (5 mg und 10 mg BID) gegenüber Placebo für die Behandlung von Krankheitszeichen und Symptomen von RA bei Patienten mit aktiver RA basierend auf einem stabilen Hintergrund von MTX; gemessen anhand von Ansprechraten des ACR 20 nach 6 Monaten.	
		Vergleich der körperlichen Funktion von Patienten nach Verabreichung von 5 mg bzw. 10 mg CP-690,550 gegenüber Placebo bei Patienten mit aktiver RA basierend auf einem stabilen Hintergrund von MTX; gemessen anhand des HAQ-DI zu Monat 3 im Vergleich zu Studienbeginn.	
		<ul> <li>Vergleich der Raten von Patienten mit aktiver RA basierend auf einem stabilen Hintergrund von MTX, die DAS28-4 ESR&lt; 2,6 zu Monat 6 erreichen nach Verabreichung von 5 mg bzw. 10 mg CP-690,550 gegenüber Placebo.</li> </ul>	
		Untersuchen der Sicherheit und Verträglichkeit über 12 Monate von 5 mg bzw. 10 mg CP-690,550 gegenüber Placebo bei Patienten mit aktiver RA basierend auf einem stabilen Hintergrund von MTX.	
		sekundäre Ziele:	
		Vergleich der Wirksamkeit von beiden CP-690,550-Dosierungen (5 mg und 10 mg BID) plus MTX gegenüber Placebo plus MTX für die Behandlung von Krankheitszeichen und Symptomen von RA zu allen anderen Studienzeitpunkten; gemessen anhand von ACR 20, ACR 50, ACR 70 und DAS28.	
		Vergleich der Wirksamkeit von Adalimumab gegenüber Placebo für die Behandlung von Krankheitszeichen und Symptomen von RA auf einem stabilen Hintergrund von MTX; gemessen anhand der Ansprechraten von ACR 20, ACR 50, ACR 70 und DAS28.  March 1918	
		Vergleich der Dauer des Ansprechens	

Itema	Charakteristikum	Studieninformation
		ACR 20, ACR 50, ACR 70 und DAS28.
		Vergleich der Inzidenz der DAS28- Remission und des Zustands niedriger Krankheitsaktivität zu jedem Studienzeitpunkt.
		Vergleich der Effekte aller in der Studie erhobenen <i>Health Outcomes</i> Parameter zu jedem Studienzeitpunkt gegenüber den jeweiligen Werten zu Studienbeginn.
		Vergleich der Wirksamkeit von Adalimumab gegenüber beiden CP- 690,550-Dosierungen (5 mg und 10 mg BID) zu allen anderen Studienzeitpunkten für die Behandlung von Krankheitszeichen und Symptomen von RA auf einem stabilen Hintergrund von MTX; gemessen anhand von ACR 20, ACR 50, ACR 70 und DAS28.
Metho	den	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte Phase III Studie; doppelblind; placebokontrolliert; Laufzeit: 1 Jahr; Zuteilungsverhältnis: 4:4:1:1:4 (Gruppe 1:Gruppe 2:Gruppe 3:Gruppe 4:
		Gruppe 5)
		Gruppenzuteilung:
		Gruppe 1: CP-690,550 5 mg BID Gruppe 2: CP-690,550 10 mg BID
		Gruppe 3: Placebo -> CP-690,550 5 mg BID (Wechsel nach spätestens 6 Monaten)
		Gruppe 4: Placebo -> CP-690,550 10 mg BID (Wechsel nach spätestens 6 Monaten)
		Gruppe 5: Adalimumab 40 mg jede 2. Woche
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Protokoll Amendment 1 (15. April 2010): Hinzufügen des DAS28(ESR) < 2,6 nach 6 Monaten als primären Endpunkt für den Vergleich beider CP-690,550 Dosierungen mit Placebo.
		<b>Protokoll Amendment 2</b> (30. November 2010):
		Überwiegend textliche Änderungen.
		<b>Protokoll Amendment 3</b> (7. Februar 2011):
		Länderspezifisches Amendment für Korea
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien:
		Vorliegen aktiver RA:
		1. Vorhandensein mind. 4 der 7 ACR

Itema	Charakteristikum	Studieninformation
		Kriterien für die Diagnose von RA
		2. Vorliegen von RA sowohl zu Screening als auch zu Baseline, wobei die beiden folgenden Kriterien erfüllt sein müssen:
		a) ≥ 6 empfindliche/druckschmerzhafte Gelenke (von 68 untersuchten Gelenken) und
		b) ≥ 6 geschwollene Gelenke (von 66 untersuchten Gelenken)
		3. Vorliegen eines der folgenden Kriterien:
		a) ESR (Westergren Methode) > 28 mm/h oder
		b) CRP > 7 mg/L via Zentrallabor
		4. Vorliegen der Klasse I, II oder III der ACR Kriterien von 1991
		Laufende Behandlung mit MTX:
		1. Orale oder parenterale MTX-Behandlung während mind. 4 Monate vor ersten Studienmedikationsgabe. In mind. 6 Wochen vor der ersten Studienmedikationsgabe sollte eine stabile Dosis von 7,5 – 25 mg wöchentlich gegeben sein. Dosierungen < 15 mg waren lediglich in Ausnahmefällen erlaubt (z.B. MTX-Intoleranz, Toxizitäten). Höhere MTX-Dosierungen waren nicht erlaubt.
		2. Gabe von Folsäure < 5 mg wöchentlich während mind. 4 Wochen vor der ersten Studienmedikationsgabe.
		3. Inadäquates Ansprechen auf MTX; offensichtliches Vorhandensein von Krankheitsaktivität.
		weitere Einschlusskriterien:
		Vorliegen einer unterzeichneten und datierten Einverständniserklärung des Patienten bzw. seines gesetzlichen Vertreters.
		2. Alter ≥ 18 Jahre.
		3. Absetzen aller nicht erlaubten Begleitmedikationen vor der ersten Studienmedikationsgabe.
		4. Negativer Schwangerschaftstest für Frauen im gebärfähigen Alter.
		5. Verwendung empfängnisverhütender Mittel für Frauen im gebärfähigen Alter bzw. für deren Partner.
		6. Kein Vorliegen einer Tuberkulose- Infektion.
		Ausschlusskriterien:
		Schwangere oder stillende Mütter.

Itema	Charakteristikum	Studieninformation
		2. Blutdyskrasien während der 3 Monate vor der ersten Studienmedikationsgabe:
		a) Hämoglobin < 9 g/dL oder Hämatokrit < 30%
		b) Anzahl weißer Blutkörperchen < 3,0 × 10 <sup>9</sup> /L.
		c) absolute Neutrophilen-Anzahl < 1,2 × 10 <sup>9</sup> /L.
		d) Thrombozytenzahl $< 100 \times 10^9/L$ .
		3. Glomuläre Filtrationsrate < 40 mL/min (nach Cockcroft-Gault).
		4. AST oder ALT > 1,5 über Normalbereich oder sonstige relevante Laborabweichungen, die gegen die Studienteilnahme sprechen.
		5. Aktuell vorliegende oder vorangegangene unkontrollierte renale, hepatische, hämatologische, gastrointestinale, metabolische, endokrine, pulmonale, kardiale oder neurologische Erkrankung.
		6. Rheumatische Autoimmunerkrankung (außer Sjogren's Syndrom) in der Vorgeschichte.
		7. Entzündete Gelenkprothese
		8. Lymphoproliferative Erkrankung (z.B. Epstein Barr Virus), Lymphom, Leukämie in der Vorgeschichte oder andere Anzeichen aktueller lymphatischer Erkrankung.
		9. Herpes Zoster (wiederholt oder disseminiert) oder Herpes Simplex in der Vorgeschichte.
		10. Infektion mit nachfolgender Hospitalisierung, parenteraler antimikrobieller Therapie oder von anderweitiger klinischer Relevanz in den 6 Monaten vor der ersten Studienmedikationsgabe.
		11. Infektion mit nötiger antimikrobieller Therapie innerhalb der 2 Wochen vor der ersten Studienmedikationsgabe.
		12. Versagen auf eine TNFi Therapie bzgl. Wirksamkeit bzw. Vorliegen eines TNFi spezifischen unerwünschten Ereignisses.
		13. Patienten, die Adalimumab in der Vorgeschichte verabreicht bekommen haben.
		14. Patienten, für die Adalimumab kontraindiziert ist gemäß Zulassung; Patienten, für die die Kriterien eines kongestiven Herzversagens der Klasse III oder IV (gemäß New York Heart Association) erfüllt sind.
		15. Therapie mit nicht-Betazellen spezifischer lymphatisch ablagernder Substanzen (z.B. Alemtuzumab) oder alkylierender Substanzen

Itema	Charakteristikum	Studieninformation
		(z.B. Cyclophosphamid) in der Vorgeschichte.
		16. Patienten, denen ein Lebend- oder abgeschwächter Impfstoff während der 6 Wochen vor der ersten Studienmedikationsgabe verabreicht wurde; oder Patienten, bei denen eine solche Impfung während der Studienlaufzeit oder 6 Wochen
		danach geplant ist.  17. Patienten, bei denen die orale Absorbierung der Medikation beeinflusst ist (z.B. nach Gastrektomie, klinisch signifikante diabetische Gastroenteropathie, Magenbypass).
		18. Alkohol- oder Drogenmissbrauch mit weniger als 6 Monate Abstinenz vor Verabreichen der ersten Studienmedikation.
		19. 12-Punkte EKG zu Screening, das klinisch relevante Abweichungen zeigt, die die Sicherheit des Patienten gefährden oder die Interpretation der Studienergebnisse beeinflussen.
		20. Vererblicher Immundefekt bei einem Verwandten ersten Grades.
		21. Aktuelle oder vorangegangene maligne Erkrankung; Ausnahmen: adäquat behandeltes oder entferntes nichtmetastasierendes Basalzell- oder Plattenepithelkarzinom der Haut oder Zervixkarzinom.
		22. Bedeutsame Verletzung oder Operation in dem Monat vor der ersten Studienmedikationsgabe.
		23. Patienten mit nicht erlaubter Begleitmedikation
		24. Patienten mit bekannter HIV-Infektion, Hepatitis B oder Hepatitis C Infektion.
		25. Patienten, die bereits an einer klinischen Studie für CP-690,550 teilgenommen haben.
		26. Patienten mit Teilnahme an klinischen Studien anderer Substanzen während 4 Wochen oder 5-mal der Halbwertszeit (je
		nachdem, was länger dauert) vor Studienbeginn. Patienten durften während der Studienteilnahme an keiner weiteren klinischen Studie teilnehmen.
		27. Patienten mit Allergie oder Hypersensitivität gegenüber MTX oder sonstiger früherer Unverträglichkeiten gegenüber MTX.
		28. Sonstiger bedeutender schwerer oder chronischer medizinischer oder psychiatrischer Zustand oder

Itema	Charakteristikum	Studieninformation
		Laborabweichungen, die gegen die Studienteilnahme sprechen.
4b	Studienorganisation und Ort der Studiendurchführung	weltweite Studiendurchführung an 115 Zentren
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Gruppe 1: CP-690,550 5 mg BID Gruppe 2: CP-690,550 10 mg BID Gruppe 3: Placebo -> CP-690,550 5 mg BID (Wechsel nach spätestens 6 Monaten) Gruppe 4: Placebo -> CP-690,550 10 mg BID (Wechsel nach spätestens 6 Monaten) Gruppe 5: Adalimumab 40 mg jede 2. Woche  Patienten in Gruppe 3 und 4 wechselten bei Nichtansprechen nach 3 Monaten auf die jeweilige aktive Substanz; spätestens nach 6 Monaten wurden alle Patienten auf die aktive Substanz umgestellt.  CP-690,550 (5 und 10 mg): Tabletten, die vom Patienten mit oder ohne Nahrungsaufnahme eingenommen werden sollten; 2 Tabletten täglich, eine morgens und eine abends (im Abstand von ca. 12 Stunden) Adalimumab (40 mg): Spritzen, die sich der Patient alle zwei Wochen subkutan verabreichen sollte.  Placebo: Äußerliche identische (zu den aktiven Substanzen) Tabletten und Spritzen wurden verabreicht. Jeder Patient bekam somit die gleiche Menge an Tabletten und Spritzen, um
		die doppelblinde Studiendurchführung zu gewährleisten. MTX als Begleitmedikation wurde (oral oder parenteral) gemäß den Anforderungen der Einschlusskriterien zusätzlich verabreicht.
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	<ul> <li>primäre Endpunkte:</li> <li>ACR20 nach 6 Monaten.</li> <li>HAQ-DI Änderung gegenüber Baseline zu Monat 3.</li> <li>Inzidenz von DAS28-4(ESR) &lt; 2,6 nach 6 Monaten.</li> </ul>
		<ul> <li>sekundäre Endpunkte:</li> <li>Ansprechraten von ACR20, ACR50 und ACR70 zu allen anderen Studienzeitpunkten.</li> <li>Änderungen der Einzelkomponenten des</li> </ul>

Itema	Charakteristikum	Studieninformation
		ACR-Kriteriums gegenüber Baseline zu allen Studienzeitpunkten.
		<ul> <li>Änderungen des DAS28-3(CRP), DAS28-4(CRP), DAS28-3(ESR) und DAS28-4(ESR) gegenüber Baseline zu allen Studienzeitpunkten.</li> </ul>
		<ul> <li>Inzidenzen von DAS28-3(CRP) ≤ 3,2,</li> <li>DAS28-4(CRP) ≤ 3,2, DAS28-3(ESR) ≤ 3,2 und DAS28-4(ESR) ≤ 3,2 gegenüber</li> <li>Baseline zu allen Studienzeitpunkten.</li> </ul>
		• Inzidenz von DAS28-4(ESR) zu allen anderen Studienzeitpunkten als Monat 6.
		• Inzidenzen von DAS28-3(CRP) < 2,6, DAS28-4(CRP) < 2,6 und DAS28- 3(ESR) < 2,6 gegenüber Baseline zu allen Studienzeitpunkten.
		DAS28 Ansprechraten (keine Verbesserung, Verbesserung (moderate oder gute)).
		ACR70 Ansprechraten für mindestens 6 Monate.
		<ul> <li>Dauerhaftigkeit des Ansprechens bzgl. ACR20, ACR50, ACR70, DAS28.</li> </ul>
		<ul> <li>Anteil der Patienten mit Ansprechen zu jedem Studienzeitpunkt, davon jeweils die Raten, die zum darauf folgenden Studienzeitpunkt weiterhin ansprachen.</li> </ul>
		<ul> <li>Änderungen des SF-36 gegenüber Baseline (8 Domänen und 2 Komponentenscores).</li> </ul>
		<ul> <li>Änderungen des WLQ gegenüber Baseline (4 Domänen und Arbeitsverlustscore).</li> </ul>
		<ul> <li>Änderungen des EQ-5D gegenüber Baseline.</li> </ul>
		<ul> <li>Änderungen des MOS Schlafscore gegenüber Baseline.</li> </ul>
		<ul> <li>Änderungen des FACIT gegenüber Baseline.</li> </ul>
		• Anteil der Patienten mit Verbesserung im HAQ-DI (Schwellenwerte: Abnahme um mind. 0,22, 0,3, 0,5 bzw. 0,8).
		Rate der Patienten, die in den Placebo- Armen zu Monat 3 auf die aktive Substanz wechseln (Nicht-Ansprecher).
		<ul> <li>Rate der Patienten, die in den Placebo- Armen zu Monat 3 fälschlicherweise auf die aktive Substanz wechseln.</li> </ul>
		Sicherheit/Verträglichkeit
		Pharmakokinetik

Item <sup>a</sup>	Charakteristikum	Studieninformation
		ACR:
		Berechnet sich aus den folgenden Einzelkomponenten:
		Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 68 Gelenken)
		2. Anzahl geschwollener Gelenke (von insgesamt 66 Gelenken)
		3. Schmerzbeurteilung durch den Patienten (visuelle Analogskala)
		4. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA)
		5. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PhysGA)
		6. CRP-Wert
		7. HAQ-DI
		Ein Ansprechen bzgl. des ACR50 liegt vor, wenn der Patient in den ersten der beiden Einzelkomponenten eine mind. 50%-ige Verbesserung erfährt sowie jeweils mind. eine 50%-ige Verbesserung in mind. 3 der 5 weiteren Einzelkomponenten. Die Definitionen des ACR20 und des ACR70 gelten analog.
		Die Erhebung aller Einzelkomponenten und des ACR erfolgte zu jedem Studienzeitpunkt.
		DAS28: Der DAS28 ist ein gewichteter Score, der in 4 Varianten berechnet wird: DAS28-3(CRP), DAS28-4(CRP), DAS28-3(ESR) und DAS28- 4(ESR). Berechnet werden die jeweiligen Scores aus den folgenden Einzelkomponenten:
		1. Anzahl empfindlicher/druckschmerzhafter Gelenke (von insgesamt 28 Gelenken)
		2. Anzahl geschwollener Gelenke (von insgesamt 28 Gelenken)
		3. CRP- bzw. ESR-Wert
		4. Gesamtbeurteilung der Krankheitsaktivität durch den Patienten (PatGA) (nur DAS28-4)
		Die Erhebung aller Einzelkomponenten und des DAS28 erfolgte zu jedem Studienzeitpunkt.
		SF-36 wurde zu jedem Studienzeitpunkt erhoben. MOS Schlafscore, FACIT und EQ- 5D zu jedem Studienzeitpunkt außer nach 9 Monaten. WLQ zu jedem Studienzeitpunkt außer nach 1 Monat und nach 9 Monaten.

<b>Item</b> <sup>a</sup>	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Hinzufügen des DAS28(ESR) < 2,6 nach 6 Monaten als primären Endpunkt, welcher bislang als sekundärer Endpunkt geplant war (Amendment 1 des SAP vom 6. April 2010).
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Es war geplant 700 Patienten in die Studie einzuschließen.  Die Fallzahl wurde anhand einer Step-Down Prozedur bestimmt, die die drei primären Endpunkte und die beiden Dosierungen in hierarchischer Reihenfolge berücksichtigt. Die zu testende Reihenfolge (für die jeweiligen Tests gegenüber Placebo) ist in folgender Grafik dargestellt:
		10 mg ACR20  10 mg ACR20  p <= 0.05  p <= 0.05  DAS < 2.6  Start  10 mg ACR20  5 mg ACR20  5 mg HAQ-DI  Both p <= 0.05  5 mg DAS < 2.6
		<ul> <li>Folgende Annahmen wurden pro Endpunkt für die geplante Fallzahl getroffen:</li> <li>ACR20: Differenz der Ansprechrate von mind. 20%; Rate unter Placebo: 30%; Power &gt; 90%.</li> <li>HAQ-DI: Differenz mind. 0,3; Standardabweichung: 0,75; Power &gt; 90%.</li> </ul>
		• DAS28-4(ESR): Differenz der Ansprechrate von mind. 15%; Rate unter Placebo: 10%; Power ca. 94%.

Item <sup>a</sup>	Charakteristikum	Studieninformation		
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es waren keine Zwischenanalysen geplant und keine Kriterien für einen vorzeitigen Studienabbruch festgelegt.		
8	Randomisierung, Erzeugung der Behandlungsfolge			
8a	Methode zur Generierung der zufälligen Zuteilung	Die Patienten wurden mittels eines zentralen Interactive Voice Response Systems (IVRS) den Behandlungsgruppen zugeteilt.		
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)  Das Zuteilungsverhältnis war 4:4:1:1:4 (Gruppe 1:Gruppe 2:Gruppe 5)			
		Gruppe 1: CP-690,550 5 mg BID Gruppe 2: CP-690,550 10 mg BID Gruppe 3: Placebo -> CP-690,550 5 mg BID (Wechsel nach spätestens 6 Monaten) Gruppe 4: Placebo -> CP-690,550 10 mg BID (Wechsel nach spätestens 6 Monaten) Gruppe 5: Adalimumab 40 mg jede 2. Woche		
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Zentrale Randomisierung mittels eines Interactive Voice Response Systems (IVRS), durch welches die Geheimhaltung gewährleistet war.		
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Zentrale Randomisierung mittels eines Interactive Voice Response Systems (IVRS)		
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) Patienten: verblindet b) Behandler: verblindet c) Beurteiler: verblindet Äußerliche identische Placebo-Tabletten und -Spritzen wurden zusätzlich (bzw. anstelle von) den aktiven Substanzen verabreicht. Jeder Patient bekam somit die gleiche Menge an Tabletten und Spritzen, um die doppelblinde Studiendurchführung zu gewährleisten.		
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	s.o. 11a  Placebo-Tabletten und –Spritzen waren äußerlich nicht von der aktiven Substanz zu unterscheiden.		
12	Statistische Methoden			

Itema	Charakteristikum	Studienini	ormat	ion								
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Generell wird für binomiale Endpunkte die Anzahl und Prozentzahl der Patienten berichtet, für kontinuierliche Endpunkte die statistischen Kenngrößen: Mittelwert, Standardabweichung, Minimum, Maximum, Median, 25%-Quantil, 75%-Quantil.										
		primäre Zielkriterien: Für ACR20 und DAS28-4(ESR) < 2,6 jeweils nach 6 Monaten wird die Differenz der Raten zwischen den CP-690,550-Gruppen und Placebo mittels Normalapproximation für binomiale Raten analysiert.										
		Für die mit des HAQ-l gemischtes gerechnet.										
		primären binomia Kontinuierliche se					ndpunkte werden analog wie die omialen Endpunkte analysiert. Ehe sekundäre Endpunkte wurden ne Weise wie HAQ-DI nach 3					
		Die Dauerhaftigkeit ACR20, ACR50, AC anhand der Raten zu bestehenden Anspred Studienzeitpunkte au										
		Daten zur Sicherheit/Verträglichkeit werden deskriptiv nach den allgemeinen Standards ausgewertet.										
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Es waren keine Subgruppenanalysen geplant. Als Sensitivitätsanalysen für die primären Zielkriterien wurden die Analysen für das Per Protokoll Kollektiv wiederholt (anstelle des Full Analysis Set).										
Resulta	nte											
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)											
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die	Gruppe:										
	a) randomisiert wurden,		1	2	3	4	5					
	b) tatsächlich die geplante Behandlung/Intervention erhalten haben,	a)	204	201	56	52	204					
	c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	b)	204	201	56	52	204					
	octaensienagt warden	c)	201	199	56	51	201					
		Gruppe 1:	CP-690	),550 5	5 mg BID							

Item <sup>a</sup>	Charakteristikum	Studieninformation	l				
		Gruppe 2: CP-690,550 10 mg BID					
		Gruppe 3: Placebo -> (Wechsel nach späte					g BID
		Gruppe 4: Placebo -> (Wechsel nach späte					ng BID
		Gruppe 5: Adalimun	nab 40	mg j	ede	2. V	Woche
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung						
	mit Angabe von Gründen			G	rupp	e:	ı
			1	2	3	4	5
		UE	24	24	2	5	22
		mangelnde Wirksamkeit	6	7	3	3	6
		Lost to follow-up	2	1	0	0	0
		fehlendes Einverständnis	4	2	0	1	1
		Tod	0	0	0	0	1
		andere	18	9	4	4	12
		Gruppe 1: CP-690,550 5 mg BID Gruppe 2: CP-690,550 10 mg BID Gruppe 3: Placebo -> CP-690,550 5 mg BID (Wechsel nach spätestens 6 Monaten)					
		Gruppe 4: Placebo -> CP-690,550 10 mg BID (Wechsel nach spätestens 6 Monaten)					
		Gruppe 5: Adalimun	nab 40	mg j	ede	2. V	Woche
14	Aufnahme / Rekrutierung						
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	erster Patientenbesud letzter Patientenbesu					
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde pla	nmäß	ig be	ende	et.	

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

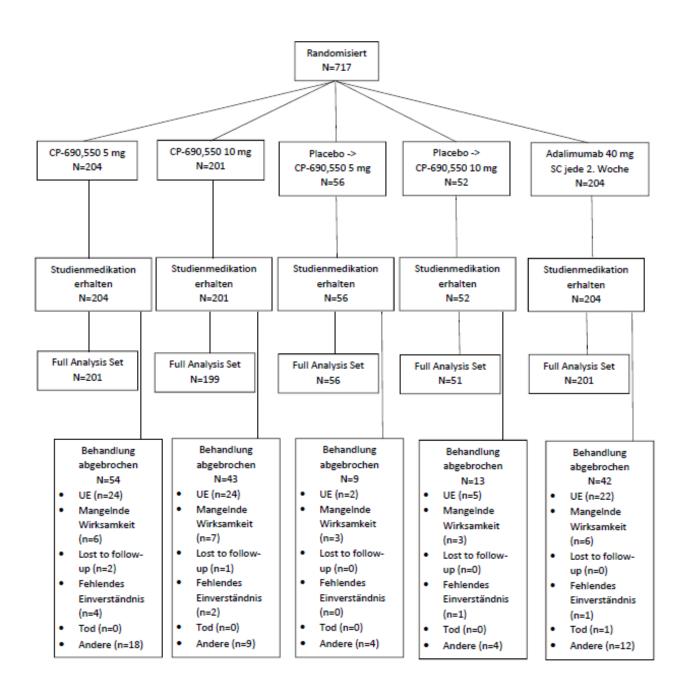


Abbildung 60: Flussdiagramm der Studie A3921064 (ORAL STANDARD)

## 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 4-195 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie A3921064 (ORAL STANDARD)

Kürzel

A

Studie: A3921064 (ORAL STANDARD)

Studienbericht der Studie A3921064 (ORAL

Genaue Benennung der Quelle

## Tabelle: Liste der für die Bewertung herangezogenen Quellen

STANDARD)
A Verzerrungsaspekte auf Studienebene:
Einstufung als randomisierte Studie
☐ <b>nein</b> → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien
Angaben zum Kriterium:
1,
für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz
Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen
☐ ja ☐ unklar ☐ nein
Angaben zum Kriterium; falls unklar oder nein, obligate 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, obligate Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Vergleichbarkeit der Grupp von prognostisch relevanten Faktoren	
☐ ja ☐ unklar ☐ nein	
Angaben zum Kriterium; falls unklar oder nein, obligate E	egründung für die Einstufung:
Verblindung von Patienten und behandelnden Personen	
Patient:	
⊠ ja □ unklar □ nein	
Angaben zum Kriterium; obligate Begründung für die Einstufun	ng:
Die Studie wurde doppelblind durchgeführt. Weder der behan Sponsor hatten Kenntnis über die tatsächlich verabreichte Medi	
behandelnde bzw. weiterbehandelnde Personen:	
⊠ ja □ unklar □ nein	
Angaben zum Kriterium; obligate Begründung für die Einstufung	ng:
Die Studie wurde doppelblind durchgeführt. Weder der behan Sponsor hatten Kenntnis über die tatsächlich verabreichte Medi	
Ergebnisunabhängige Berichterstattung aller relevanten End	punkte
⊠ ja □ unklar □ nein	
Angaben zum Kriterium; falls unklar oder nein, obligate Begrü	indung für die Einstufung:
Keine sonstigen (endpunktübergreifenden) Aspekte, die zu V	erzerrungen führen können
⊠ ja □ nein	
Angaben zum Kriterium; falls nein, obligate Begründung für d	ie Einstufung:
instufung des Verzerrungspotenzials der Ergebnisse auf Studier andomisierte Studien durchzuführen):	nebene (ausschließlich für

#### Begründung für die Einstufung:

Die Randomisierung erfolgte zentral, jedem Patienten wurde nach dem Screening-Besuch eine Radomisierungsnummer mit der dem Randomisierungsplan entsprechenden Behandlung zugeteilt. Die Verdeckung der Gruppenzuteilung wurde durch den Einsatz eines zentralen, unabhängigen *Interactive Voice Response System* (IVRS) gewährleistet.

Die Studie wurde doppelblind durchgeführt. Weder der behandelnde Arzt noch der Patient oder der Sponsor hatten Kenntnis über die tatsächlich verabreichte Medikation.

Die im Studienbericht präsentierten Ergebnisse folgen den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch sonstigen Aspekte, die das Verzerrungspotenzial der Studie beeinflussen könnten.

#### B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: alle im Dossier dargestellten Endpunkte zur Morbidität (außer für DAS28-4(ESR), EULAR-Response(ESR) und WLQ), zur Lebensqualität und zur Sicherheit/Verträglichkeit 1. Verblindung der Endpunkterheber ⊠ ja unklar nein Angaben zum Kriterium; obligate Begründung für die Einstufung: Die Studie wurde doppelblind durchgeführt. Weder der behandelnde Arzt noch der Patient oder der Sponsor hatten Kenntnis über die tatsächlich verabreichte Medikation. 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

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

□ niedrig □ hoch

☐ nein

Begründung für die Einstufung:

⊠ ja

Durch das doppelblinde Studiendesign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Bei der Auswertung wird das ITT-Prinzip adäquat umgesetzt. Die Auswertung der Endpunkte folgt jeweils den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren existieren auch sonstigen Aspekte, die das Verzerrungspotenzial

bezüglich dieser Endpunkte beeinflussen könnten.

## Endpunkt: DAS28-4(ESR) und EULAR-Response(ESR)

1.	Verblindung o	der Endpunkterh	eber
	⊠ ja	unklar	nein nein
	Angaben zum	n Kriterium; <u>obliga</u>	te Begründung für die Einstufung:
			durchgeführt. Weder der behandelnde Arzt noch der Patient oder der ie tatsächlich verabreichte Medikation.
2.	Adäquate Um	setzung des ITT-	Prinzips
	☐ ja	□ unklar	□ nein
	Angaben zur	m Kriterium; falls	unklar oder nein, obligate Begründung für die Einstufung:
	Weiterleitun Studie zu ge Ergebnisse z führt, dass ei	g des Testergebnis währleisten. Es ist zu diesem Endpunk ine adäquate Umse	tren erhoben, bei denen die Voraussetzungen für eine direkte ses an das Zentrallabor gegeben waren, um die Verblindung der anzunehmen, dass der hohe Anteil von Patienten, für die keine et vorliegen (> 30% des FAS in den jeweiligen Studienarmen) dazu etzung des ITT-Prinzips in Frage gestellt wird und dass eine sinnvolle es DAS28-4(ESR) bzw. der EULAR-Response(ESR) nicht möglich
3.	Ergebnisunab	hängige Berichte	rstattung dieses Endpunkts alleine
	-	_	_
	Angaben zur	iii Kriteriuiii; iaiis	unklar oder nein, <u>obligate</u> Begründung für die Einstufung:
4.	Keine sonstige	en (endpunktspez	ifischen) Aspekte, die zu Verzerrungen führen können
	⊠ ja	☐ nein	
	Angaben zur	m Kriterium; falls	nein, obligate Begründung für die Einstufung:
		erzerrungspotenz udien durchzufüh	ials der Ergebnisse des Endpunkts (ausschließlich für ren):
	niedrig	⊠ ho	ch
F	Begründung für o	die Einstufung:	
ŀ	Kenntnis über di	e tatsächlich veral	sign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei breichte Studienmedikation hat. Die Auswertung der Endpunkte folgt orab erstellten statistischen Analyseplans, die Darstellung der

Studienergebnisse erfolgte daher nicht ergebnisabhängig. Wegen des Aspekts zur fraglichen Umsetzung

des ITT-Prinzips wird das Verzerrungspotenzial als hoch eingestuft.

Endpu	Endpunkt: WLQ						
5.	Verblindung o	der Endpunkterh	eber				
	⊠ ja	unklar	nein nein				
	Angaben zum	Kriterium; obliga	ate Begründung für die Einstufung:				
			durchgeführt. Weder der behandelnde Arzt noch der Patient oder der ie tatsächlich verabreichte Medikation.				
6.	Adäquate Um	setzung des ITT-	Prinzips				
	☐ ja	unklar 🛮	☐ nein				
	Angaben zui	m Kriterium; falls	unklar oder nein, obligate Begründung für die Einstufung:				
	Anteil von d "Berufstätigl Ergebnisse d dass dies nic Teilpopulatie	eutlich weniger als keit" wurde nicht s lurch ein Verzerru ht auszuschließen on c)-Kombi und c	nach nur von Berufstätigen ausgefüllt werden. Es liegen nur für einen s 50% der Studienpatienten Werte zum WLQ vor. Nach dem Faktor stratifiziert randomisiert. Da es schwer ist, zu beurteilen, inwieweit die ngspotenzial verfälscht sein könnten, wird hier davon ausgegangen, ist. Auch der hohe Anteil von fehlenden Werten (ca. 40-60% für ca. 30-35% für Teilpopulation d)-Kombi in den jeweiligen s eine sinnvolle Interpretation der Ergebnisse des WLQ nicht möglich				
7.	⊠ ja	unklar	erstattung dieses Endpunkts alleine  nein  unklar oder nein, obligate Begründung für die Einstufung:				
8.	Keine sonstige	en (endpunktspez	zifischen) Aspekte, die zu Verzerrungen führen können				
	⊠ ja	nein nein					
	Angaben zui	m Kriterium; falls	nein, obligate Begründung für die Einstufung:				
		erzerrungspotenz ıdien durchzufüh	zials der Ergebnisse des Endpunkts (ausschließlich für nren):				
	niedrig	⊠ ho	och				
E	Begründung für d	lie Einstufung:					
Ι	Ourch das doppe	elblinde Studiende	sign ist die Voraussetzung erfüllt, dass der Endpunkterheber keinerlei				

Kenntnis über die tatsächlich verabreichte Studienmedikation hat. Die Auswertung der Endpunkte folgt jeweils den Vorgaben eines vorab erstellten statistischen Analyseplans, die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Wegen der Aspekte zur fraglichen Umsetzung des ITT-Prinzips und dem hohen Anteil fehlender Werte wird das Verzerrungspotenzial als hoch eingestuft.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

#### Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

weraen (A: enapunktubergreifena; B: enapunktspezifiscn).						
A Verzerrungsaspekte auf Studienebene:						
Einstufung als randomisierte Studie						
☐ ja → Bewertung der Punkte 1 und 2 für randomisierte Studien						
<ul> <li>□ 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.</li> <li>→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien</li> </ul>						
Angaben zum Kriterium:						
1. <u>für randomisierte Studien:</u>						
Adäquate Erzeugung der Randomisierungssequenz						
<b>ja</b> : 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						
<b>ja</b> : 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:						

	ner Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutz
<u>für</u>	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 Arzneimittelln/Arzneimittelbehältern
	Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags.
	die Gruppenzuteilung beinhaltet
	unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder si
	ungenügend genau.
	nein: Die Gruppenzuteilung erfolgte nicht verdeckt.
	Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
	-
<u>für</u>	nicht randomisierte Studien:
<u>füı</u>	nicht randomisierte Studien:  Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren
<u>füı</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren
<u>füı</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.  Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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).
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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.
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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
<u>füi</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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.
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevationen  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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.
<u>für</u>	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevationen  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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.
	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren  ja: Eines der folgenden Merkmale trifft zu:  Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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:
Ve	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren    ja: Eines der folgenden Merkmale trifft zu:   Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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:
Ve	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren    ja: Eines der folgenden Merkmale trifft zu:   Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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:    rblindung von Patienten und behandelnden Personen tient:
Ve	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch releva Faktoren    ja: Eines der folgenden Merkmale trifft zu:   Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen d 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:

## 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.

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.
<b>nein:</b> 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:
<ul> <li>Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können</li> <li>z. B.</li> <li>zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien</li> <li>intransparenter Patientenfluss</li> <li>Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:</li> <li>Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).</li> <li>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).</li> <li>Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.</li> <li>Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.</li> </ul>
□ ja
nein 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.

5.

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).

	<b>unklar</b> : 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
	<ul> <li>z. B.</li> <li>relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen</li> <li>unplausible Angaben</li> <li>Anwendung inadäquater statistischer Verfahren</li> </ul>
	☐ 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 <u>sowie</u> 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: