

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

Abrocitinib (CIBINQO®)

PFIZER PHARMA GmbH
als örtlicher Vertreter des Zulassungsinhabers
Pfizer Europe MA EEIG

Modul 4 A

*Behandlung von mittelschwerer bis schwerer atopischer
Dermatitis bei Erwachsenen, die für eine systemische
Therapie infrage kommen*

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	11
Abkürzungsverzeichnis.....	13
4 Modul 4 – allgemeine Informationen	17
4.1 Zusammenfassung der Inhalte von Modul 4.....	18
4.2 Methodik	28
4.2.1 Fragestellung	28
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	30
4.2.3 Informationsbeschaffung	32
4.2.3.1 Studien des pharmazeutischen Unternehmers	32
4.2.3.2 Bibliografische Literaturrecherche	32
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken	34
4.2.3.4 Suche auf der Internetseite des G-BA.....	35
4.2.3.5 Selektion relevanter Studien	36
4.2.4 Bewertung der Aussagekraft der Nachweise.....	37
4.2.5 Informationssynthese und -analyse	38
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	38
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	39
4.2.5.3 Meta-Analysen.....	56
4.2.5.4 Sensitivitätsanalysen.....	58
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	59
4.2.5.6 Indirekte Vergleiche	62
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	64
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	64
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	64
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	64
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	66
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken	68
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA	69
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	70
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	71
4.3.1.2.1 Studiendesign und Studienpopulationen	71
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	80
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	81
4.3.1.3.1 Endpunkte – RCT.....	82
4.3.1.3.1.1 Morbidität – RCT	84
4.3.1.3.1.1.1 Krankheitsschwere (EASI) – RCT	84

4.3.1.3.1.1.2	Krankheitsschwere (SCORAD) – RCT	92
4.3.1.3.1.1.3	Betroffene BSA – RCT	99
4.3.1.3.1.1.4	Patientenberichtete Symptomatik (POEM) – RCT	105
4.3.1.3.1.1.5	Juckreiz – RCT	112
4.3.1.3.1.1.6	Schlafstörungen – RCT	126
4.3.1.3.1.1.7	Hautschmerz (<i>Skin Pain</i> NRS) – RCT	141
4.3.1.3.1.1.8	Ängste und Depressionen (HADS) – RCT	148
4.3.1.3.1.1.9	Gesundheitszustand (EQ-5D VAS) – RCT	155
4.3.1.3.1.2	Gesundheitsbezogene Lebensqualität	161
4.3.1.3.1.2.1	Krankheitsspezifische Lebensqualität (DLQI) – RCT	161
4.3.1.3.1.3	Verträglichkeit – RCT	168
4.3.1.3.2	Subgruppenanalysen – RCT	181
4.3.1.4	Liste der eingeschlossenen Studien - RCT	207
4.3.2	Weitere Unterlagen	207
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	207
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	208
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche	208
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	208
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	209
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	211
4.3.2.1.4	Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT ..	211
4.3.2.2	Nicht randomisierte vergleichende Studien	211
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	211
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien	212
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	213
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien	213
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	214
4.3.2.2.4	Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien	214
4.3.2.3	Weitere Untersuchungen	215
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	215
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	215
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	216
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	216
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	217
4.3.2.3.4	Liste der eingeschlossenen Studien – weitere Untersuchungen	217
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens	217
4.4.1	Beurteilung der Aussagekraft der Nachweise	217
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß	218
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	228
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	229
4.5.1	Begründung für die Vorlage indirekter Vergleiche	229

4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	229
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	229
4.5.4	Verwendung von Surrogatendpunkten.....	229
4.6	Referenzliste.....	231
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		237
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken.....		241
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		243
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken).....		244
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		253
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		275

Tabellenverzeichnis

	Seite
Tabelle 4-1: Prädefinierte Ein-/ Ausschlusskriterien für die Selektion von RCT für die Nutzenbewertung – Suche nach RCT	19
Tabelle 4-2: Wahrscheinlichkeit und Ausmaß des Zusatznutzens	22
Tabelle 4-3: Prädefinierte Ein-/ Ausschlusskriterien für die Selektion von RCT für die Nutzenbewertung – Suche nach RCT	31
Tabelle 4-4: Patientencharakteristika der Studie JADE DARE	40
Tabelle 4-5: Übersicht der im Nutzendossier dargestellten patientenrelevanten Endpunkte... 42	42
Tabelle 4-6: Übersicht der Analysen für die betrachteten patientenrelevanten Endpunkte	55
Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	65
Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel	66
Tabelle 4-9: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel	69
Tabelle 4-10: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel.....	70
Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	71
Tabelle 4-12: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	72
Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	73
Tabelle 4-14: Charakterisierung der Studienpopulationen für JADE DARE – RCT mit dem zu bewertenden Arzneimittel.....	74
Tabelle 4-15: Hintergrundtherapie der Studie JADE DARE anhand der <i>Safety</i> -Population... 76	76
Tabelle 4-16: <i>Rescue</i> -Therapie der Studie JADE DARE anhand der <i>Safety</i> -Population	77
Tabelle 4-17: Behandlungsdauer der Studie JADE DARE anhand der <i>Safety</i> -Population.....	80
Tabelle 4-18: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	81
Tabelle 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	82
Tabelle 4-20: Operationalisierung des Endpunkts EASI	85
Tabelle 4-21: Bewertung des Verzerrungspotenzials für EASI in RCT mit dem zu bewertenden Arzneimittel	86
Tabelle 4-22: Beobachtungsdauer für EASI in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	86
Tabelle 4-23: Rücklaufquote für EASI aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	87

Tabelle 4-24: Ergebnisse für Krankheitsschwere (EASI) aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	88
Tabelle 4-25: Ergebnisse für mittlere prozentuale Veränderung des EASI; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	89
Tabelle 4-26: Operationalisierung des Endpunkts SCORAD	92
Tabelle 4-27: Bewertung des Verzerrungspotenzials für SCORAD in RCT mit dem zu bewertenden Arzneimittel	93
Tabelle 4-28: Beobachtungsdauer für SCORAD in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	94
Tabelle 4-29: Rücklaufquote für SCORAD in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	94
Tabelle 4-30: Ergebnisse für Krankheitsschwere (SCORAD) aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	95
Tabelle 4-31: Ergebnisse für mittlere prozentuale Veränderung im SCORAD; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	96
Tabelle 4-32: Operationalisierung des supportiven Endpunkts betroffene BSA	99
Tabelle 4-33: Bewertung des Verzerrungspotenzials für betroffene BSA in RCT mit dem zu bewertenden Arzneimittel	100
Tabelle 4-34: Beobachtungsdauer für betroffene BSA in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	100
Tabelle 4-35: Rücklaufquote für betroffene BSA in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	101
Tabelle 4-36: Ergebnisse für mittlere prozentuale Veränderung der betroffenen BSA; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	102
Tabelle 4-37: Operationalisierung des Endpunkts patientenberichtete Symptomatik	105
Tabelle 4-38: Bewertung des Verzerrungspotenzials für POEM in RCT mit dem zu bewertenden Arzneimittel	106
Tabelle 4-39: Beobachtungsdauer für POEM in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	107
Tabelle 4-40: Rücklaufquote für POEM in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	107
Tabelle 4-41: Ergebnisse für patientenberichtete Symptomatik (POEM) aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	108
Tabelle 4-42: Ergebnisse für mittlere Veränderung im POEM-Gesamtscore; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	110
Tabelle 4-43: Operationalisierung des Endpunkts Juckreiz	112
Tabelle 4-44: Bewertung des Verzerrungspotenzials für Juckreiz in RCT mit dem zu bewertenden Arzneimittel	113
Tabelle 4-45: Beobachtungsdauer für <i>Peak Pruritus</i> NRS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	114

Tabelle 4-46: Rücklaufquote für <i>Peak Pruritus</i> NRS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	114
Tabelle 4-47: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber <i>Baseline</i> in der <i>Peak Pruritus</i> NRS aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	115
Tabelle 4-48: Ergebnisse für mittlere Veränderung der <i>Peak Pruritus</i> NRS; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	116
Tabelle 4-49: Beobachtungsdauer für SCORAD VAS für Juckreiz in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	119
Tabelle 4-50: Rücklaufquote für SCORAD VAS für Juckreiz in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	119
Tabelle 4-51: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 2 Punkte gegenüber <i>Baseline</i> in der SCORAD VAS für Juckreiz aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	120
Tabelle 4-52: Ergebnisse für mittlere Veränderung in der SCORAD VAS für Juckreiz; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	121
Tabelle 4-53: Beobachtungsdauer für POEM-Einzelitem für Juckreiz in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	124
Tabelle 4-54: Rücklaufquote für POEM-Einzelitem für Juckreiz in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	124
Tabelle 4-55: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für Juckreiz aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	125
Tabelle 4-56: Operationalisierung des Endpunkts Schlafstörungen	126
Tabelle 4-57: Bewertung des Verzerrungspotenzials für Schlafstörungen in RCT mit dem zu bewertenden Arzneimittel	127
Tabelle 4-58: Beobachtungsdauer für MOS <i>Sleep Scale</i> in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	128
Tabelle 4-59: Rücklaufquote für MOS <i>Sleep Scale Score</i> Index-1 und MOS <i>Sleep Scale Score</i> Index-2 in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	128
Tabelle 4-60: Ergebnisse für MOS <i>Sleep Scale</i> aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	129
Tabelle 4-61: Ergebnisse für MOS <i>Sleep Scale</i> ; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	130
Tabelle 4-62: Beobachtungsdauer für SCORAD VAS für Schlafstörungen in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	134
Tabelle 4-63: Rücklaufquote für SCORAD VAS für Schlafstörungen in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	134
Tabelle 4-64: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 2 Punkte gegenüber <i>Baseline</i> in der SCORAD VAS für Schlafstörungen aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	135

Tabelle 4-65: Ergebnisse für mittlere Veränderung in der SCORAD VAS für Schlafstörungen; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	136
Tabelle 4-66: Beobachtungsdauer für POEM-Einzelitem für Schlafstörungen in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	139
Tabelle 4-67: Rücklaufquote für POEM-Einzelitem für Schlafstörung in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	139
Tabelle 4-68: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für Schlafstörungen aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	140
Tabelle 4-69: Operationalisierung des Endpunkts Hautschmerz.....	141
Tabelle 4-70: Bewertung des Verzerrungspotenzials für Hautschmerz in RCT mit dem zu bewertenden Arzneimittel.....	142
Tabelle 4-71: Beobachtungsdauer für <i>Skin Pain</i> NRS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	142
Tabelle 4-72: Rücklaufquote für <i>Skin Pain</i> NRS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	143
Tabelle 4-73: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber <i>Baseline</i> in der <i>Skin Pain</i> NRS aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	144
Tabelle 4-74: Ergebnisse für mittlere Veränderung in der <i>Skin Pain</i> NRS; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	145
Tabelle 4-75: Operationalisierung des Endpunkts Ängste und Depressionen.....	148
Tabelle 4-76: Bewertung des Verzerrungspotenzials für Ängste und Depressionen in RCT mit dem zu bewertenden Arzneimittel.....	149
Tabelle 4-77: Beobachtungsdauer für HADS in RCT JADE DARE mit dem zu bewertenden Arzneimittel.....	149
Tabelle 4-78: Rücklaufquote für HADS Subscore für Angst und HADS Subscore für Depression in RCT JADE DARE mit dem zu bewertenden Arzneimittel.....	150
Tabelle 4-79: Ergebnisse für Ängste und Depressionen (HADS); MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	151
Tabelle 4-80: Operationalisierung des Endpunkts Gesundheitszustand.....	155
Tabelle 4-81: Bewertung des Verzerrungspotenzials für Gesundheitszustand in RCT mit dem zu bewertenden Arzneimittel.....	156
Tabelle 4-82: Beobachtungsdauer für EQ-5D VAS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	156
Tabelle 4-83: Rücklaufquote für EQ-5D VAS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	157
Tabelle 4-84: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 15 Punkte gegenüber <i>Baseline</i> in der EQ-5D VAS aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	158

Tabelle 4-85: Ergebnisse für mittlere Veränderung der EQ-5D VAS; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	159
Tabelle 4-86: Operationalisierung des Endpunkts krankheitsspezifische Lebensqualität	161
Tabelle 4-87: Bewertung des Verzerrungspotenzials für krankheitsspezifische Lebensqualität in RCT mit dem zu bewertenden Arzneimittel.....	162
Tabelle 4-88: Beobachtungsdauer für DLQI in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	163
Tabelle 4-89: Rücklaufquote für DLQI in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	163
Tabelle 4-90: Ergebnisse für DLQI aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	164
Tabelle 4-91: Ergebnisse für mittlere Veränderung des DLQI; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	165
Tabelle 4-92: Operationalisierung der Endpunkte zu Verträglichkeit	168
Tabelle 4-93: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel	169
Tabelle 4-94: Beobachtungsdauer für Verträglichkeit in RCT JADE DARE mit dem zu bewertenden Arzneimittel	170
Tabelle 4-95: Ergebnisse für Verträglichkeit aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	171
Tabelle 4-96: Häufige UE nach SOC und PT aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	172
Tabelle 4-97: Ergebnisse für UE, die zum Therapieabbruch geführt haben, nach SOC und PT aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	178
Tabelle 4-98: Ergebnisse für spezifische UE aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	180
Tabelle 4-99: Matrix der durchgeführten Subgruppenanalysen.....	183
Tabelle 4-100: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für JADE DARE	186
Tabelle 4-101: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: EASI-75.....	192
Tabelle 4-102: Detaillierte Ergebnisse der Subgruppenanalyse für den Anteil der Patienten mit einer Verbesserung um ≥ 5 Punkte gegenüber <i>Baseline</i> im POEM-Gesamtscore mit statistisch signifikantem Interaktionsterm ($p < 0,05$).....	193
Tabelle 4-103: Detaillierte Ergebnisse der Subgruppenanalyse für Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber <i>Baseline</i> in der <i>Skin Pain</i> NRS mit statistisch signifikantem Interaktionsterm ($p < 0,05$).....	194
Tabelle 4-104: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach SOC Erkrankungen des Gastrointestinaltrakts.....	195

Tabelle 4-105: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach SOC Erkrankungen des Nervensystems	197
Tabelle 4-106: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Konjunktivitis	198
Tabelle 4-107: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Follikulitis.....	199
Tabelle 4-108: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Übelkeit.....	200
Tabelle 4-109: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Erbrechen.....	202
Tabelle 4-110: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Kopfschmerzen.....	203
Tabelle 4-111: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Akne.....	204
Tabelle 4-112: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Spezifische UE Follikulitis	205
Tabelle 4-113: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Spezifische UE Konjunktivitis.....	206
Tabelle 4-114: Liste der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	207
Tabelle 4-115: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	208
Tabelle 4-116: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	209
Tabelle 4-117: Operationalisierung von <Endpunkt xxx>.....	209
Tabelle 4-118: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	210
Tabelle 4-119: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	210
Tabelle 4-120: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	212
Tabelle 4-121: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien	213
Tabelle 4-122: Operationalisierung von <Endpunkt xxx>.....	213
Tabelle 4-123: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	214
Tabelle 4-124: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen ..	216
Tabelle 4-125: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	216
Tabelle 4-126: Wahrscheinlichkeit und Ausmaß des Zusatznutzens.....	220

Tabelle 4-127: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	228
Tabelle 4-128 (Anhang): Studiendesign und -methodik für Studie JADE DARE	254
Tabelle 4-129 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie JADE DARE	276

Abbildungsverzeichnis

	Seite
Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	67
Abbildung 4-2: Studienschema der Studie JADE DARE	78
Abbildung 4-3: Verlaufskurve für die mittlere prozentuale Veränderung des EASI von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	91
Abbildung 4-4: Verlaufskurve für die mittlere prozentuale Veränderung des SCORAD von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	98
Abbildung 4-5: Verlaufskurve für die mittlere prozentuale Veränderung der BSA von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	104
Abbildung 4-6: Verlaufskurve für die mittlere Veränderung des POEM-Gesamtscores von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	111
Abbildung 4-7: Verlaufskurve für die mittlere Veränderung der <i>Peak Pruritus</i> NRS von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	118
Abbildung 4-8: Verlaufskurve für die mittlere Veränderung der SCORAD VAS für Juckreiz von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	123
Abbildung 4-9: Verlaufskurve für die mittlere Veränderung des MOS <i>Sleep Scale</i> Index-1 von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	132
Abbildung 4-10: Verlaufskurve für die mittlere Veränderung des MOS <i>Sleep Scale</i> Index-2 von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	133
Abbildung 4-11: Verlaufskurve für die mittlere Veränderung der SCORAD VAS für Schlafstörungen von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	138
Abbildung 4-12: Verlaufskurve für die mittlere Veränderung der <i>Skin Pain</i> NRS von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	147
Abbildung 4-13: Verlaufskurve für die mittlere Veränderung des HADS- <i>Subscores</i> für Angst von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	153
Abbildung 4-14: Verlaufskurve für die mittlere Veränderung des HADS- <i>Subscores</i> für Depression von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	154
Abbildung 4-15: Verlaufskurve für die mittlere Veränderung der EQ-5D VAS von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	160
Abbildung 4-16: Verlaufskurve für die mittlere Veränderung des DLQI von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	167

Abbildung 4-17: Verlaufskurve für den Anteil der Patienten mit SCORAD-90 Ansprechen von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	225
Abbildung 4-18: Verlaufskurve für den Anteil der Patienten mit EASI-90 Ansprechen von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	226
Abbildung 4-19: Verlaufskurve für den Anteil der Patienten mit einem POEM-Gesamtscore von 0-2 von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel	226
Abbildung 4-20: Verlaufskurve für den Anteil der Patienten mit einem DLQI von 0 oder 1 von <i>Baseline</i> bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel.....	227
Abbildung 4-21: Flow-Chart der Studie JADE DARE.....	274

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AD	Atopische Dermatitis
ALT	Alanin-Aminotransferase
AMIce	Arzneimittel-Informationssystem
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	<i>Analysis of Covariance</i>
AST	Aspartat-Aminotransferaset
BSA	Körperoberfläche (<i>Body Surface Area</i>)
C-SSRS	<i>Columbia Suicide Severity Rating Scale</i>
CD	<i>Cluster of Differentiation</i>
CONSORT	<i>Consolidated Standards of Reporting Trials</i>
CoV	Coronavirus
CRF	<i>Case Report Form</i>
CTCAE	<i>Common Terminology Criteria for Adverse Events</i>
d. h.	Das heißt
DIMDI	Deutsches Institut für Medizinische Dokumentation
DLQI	<i>Dermatology Life Quality Index</i>
DNA	Desoxyribonukleinsäure (<i>Deoxyribonucleic Acid</i>)
EASI	<i>Eczema Area and Severity Index</i>
EG	Europäische Gemeinschaft
EMA	Europäischen Arzneimittel-Agentur (<i>European Medicines Agency</i>)
EQ-5D	<i>European Quality of Life Group 5 Dimensions</i>
EQ-5D-5L	<i>European Quality of Life Group 5 Dimensions 5 Level</i>
etc.	<i>Et cetera</i>
EU-CTR	<i>European Union Clinical Trials Register</i>
FAS	<i>Full-Analysis-Set</i>
ggf.	Gegebenenfalls
G-BA	Gemeinsamer Bundesausschuss
GCP	Gute klinische Praxis (<i>Good Clinical Practice</i>)
HADS	<i>Hospital Anxiety and Depression Scale</i>
HBc	Hepatitis-B-Core

Abkürzung	Bedeutung
HBs	Hepatitis-B-Surface
HBV	Hepatitis-B-Virus
HCRU	Inanspruchnahme von Gesundheitsressourcen (<i>Health Care Resource Utilization</i>)
ICH	<i>International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use</i>
ICTRP	<i>International Clinical Trials Registry Platform</i>
IGA	<i>Investigator's Global Assessment</i>
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	<i>Interactive Response Technology</i>
ITT	<i>Intention-to-treat</i>
IWR	<i>Interactive Web-based Response</i>
JAK	Januskinase
kg	Kilogramm
KI	Konfidenzintervall
LLQ	Untere Quantifizierungsgrenze (<i>Lower Limit of Quantification</i>)
LS	Methode der kleinsten Quadrate (<i>Least Squares</i>)
LTE	Long-term extension
max	Maximal
MedDRA	<i>Medical Dictionary for Regulatory Activities</i>
mg	Milligramm
min	Minimal
mITT	Modifiziertes <i>Intention-to-treat-Analysis-Set</i>
mIU	Tausend internationale Units
ml	Milliliter
mm ³	Kubikmillimeter
MMRM	<i>Mixed effect Model Repeat Measurement</i>
MOS	<i>Medical Outcomes Study</i>
Ms	Millisekunden
MTC	<i>Mixed Treatment Comparison</i>
MW	Mittelwert
N	Anzahl der (in der Analyse eingeschlossenen) Patienten

Abkürzung	Bedeutung
n	Anzahl der Patienten in der Kategorie/ mit Ereignis
N*	Anzahl der Patienten im FAS mit Baseline-Wert und Wert zum jeweiligen Beobachtungszeitpunkt
NA	Nicht anwendbar
NRS	<i>Numerical Rating Scale</i>
OR	Odds Ratio
oSCORAD	<i>Objective Scoring Atopic Dermatitis</i>
PGA	<i>Physician Global Assessment</i>
p.o.	Peroral
POEM	<i>Patient-Oriented Eczema Measure</i>
PPD	<i>Purified Protein Derivative Test</i>
PT	Bevorzugte Bezeichnung (<i>Preferred Term</i>)
Q2W	Alle 2 Wochen
QD	Täglich
QFT-G	<i>QuantiFERON®-TB Gold In-Tube Test</i>
RCT	Randomisierte kontrollierte Studie (<i>Randomized Controlled Trial</i>)
RD	Risikodifferenz
RR	Relatives Risiko
SAP	Statistischer Analyseplan
SARS	<i>Severe Acute Respiratory Syndrome</i>
SAS	<i>Safety-Analysis-Set</i>
s.c.	Subkutan
SCORAD	<i>Scoring Atopic Dermatitis</i>
Scr/WO	<i>Screening/ Washout</i>
SD	Standardabweichung (<i>Standard Deviation</i>)
SE	Standardfehler (<i>Standard Error</i>)
SGB	Sozialgesetzbuch
SMD	Standardisierte Mittelwertdifferenz
SMQs	<i>Standardised MedDRA Queries</i>
SOC	Systemorganklasse (<i>System Organ Class</i>)
STE	Surrogate Threshold Effects
STROBE	<i>Strengthening the Reporting of Observational Studies in</i>

Abkürzung	Bedeutung
	<i>Epidemiology</i>
SUE	Schwerwiegendes UE
TCI	Topische Calcineurin-Inhibitoren
TCS	Topische Glukokortikosteroide (<i>Topical corticosteroids</i>)
TREND	<i>Transparent Reporting of Evaluations with Non-Randomized Design</i>
u. a.	Unter anderem
UE	Unerwünschtes Ereignis
ULN	Obere Grenze der Norm (<i>Upper limit of normal</i>)
USA	<i>United States of America</i>
usw.	Und so weiter
VAS	Visuelle Analogskala
VerfO	Verfahrensordnung
vs.	Versus
WHO	Weltgesundheitsorganisation (<i>World Health Organization</i>)
z. B.	Zum Beispiel
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

Patientenpopulation

Abrocitinib (CIBINQO®) wird angewendet zur Behandlung von mittelschwerer bis schwerer atopischer Dermatitis (AD) bei Erwachsenen, die für eine systemische Therapie infrage kommen.

Die Fragestellung für das vorliegende Modul 4 A ist die Bewertung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens hinsichtlich patientenrelevanter Endpunkte für die Behandlung der mittelschweren bis schweren atopischen Dermatitis durch Abrocitinib im Vergleich zur zweckmäßigen Vergleichstherapie (zVT) Dupilumab bei Erwachsenen, die für eine systemische Therapie infrage kommen.

Intervention

Die zu bewertende Intervention ist Abrocitinib gemäß den Angaben in der Fachinformation. Diese sieht eine 1-mal tägliche orale Gabe von 200 mg Abrocitinib vor. Abrocitinib kann alleine oder mit anderen Arzneimitteln zur topischen Anwendung bei AD angewendet werden (1). Die begleitenden topischen Therapien umfassen das Auftragen von Emollientien und/ oder von leichten bis mittelstark wirksamen topischen Glukokortikosteroiden (*Topical corticosteroids*, TCS) oder topischen Calcineurin-Inhibitoren (TCI).

Vergleichstherapie

Im vorliegenden Modul 4 A wird der Zusatznutzen gegenüber der folgenden zVT dargestellt (siehe Modul 3 A, Abschnitt 3.1):

- Dupilumab (ggf. in Kombination mit TCS und/ oder TCI)

Die Vergleichstherapie besteht aus der subkutanen Gabe von Dupilumab gemäß der Fachinformation (2). Die Vergleichstherapie besteht somit für Erwachsene aus einer Initialdosis von 600 mg Dupilumab (2 aufeinanderfolgende Injektionen zu je 300 mg) an Tag 1 und der anschließenden Gabe von 300 mg Dupilumab 1-mal alle 2 Wochen. Dupilumab kann mit oder ohne einer begleitenden topischen Therapie mit TCS oder TCI angewendet werden. Die Anwendung der TCI sollte dabei auf Problemzonen wie Gesicht, Hals, intertriginöse Bereiche und den Genitalbereich beschränkt werden.

Endpunkte

Unter Berücksichtigung der Nutzendimensionen Mortalität, Morbidität, Lebensqualität und Verträglichkeit erfolgt die Bewertung des Ausmaßes des medizinischen Nutzens und Zusatznutzens anhand von patientenrelevanten Endpunkten. Es werden randomisierte kontrollierte Studien (*Randomized Controlled Trial*, RCT) eingeschlossen.

Datenquellen

Zur Identifizierung relevanter Studien und Publikationen zur Beurteilung des medizinischen Nutzens und Zusatznutzens von Abrocitinib wurde eine umfassende systematische Suche in Studienregistern (clinicaltrials.gov, *European Union Clinical Trials Register* (EU-CTR) und *International Clinical Trials Registry Platform* (ICTRP) Suchportal der Weltgesundheitsorganisation (*World Health Organization*, WHO)), sowie in bibliografischen Datenbanken (MEDLINE®, EMBASE® und *Cochrane Central Register of Controlled Trials*) durchgeführt. Die Selektion der relevanten Studien wurde von zwei Reviewern unabhängig voneinander vorgenommen und Diskrepanzen bei der Beurteilung im Konsens geklärt. Für alle im Rahmen dieser systematischen Recherche identifizierten Studien wurden die Website des Gemeinsamen Bundesausschusses (G-BA, <https://www.g-ba.de/sys/suche/?suchbegriff=>), das *Clinical Data* Suchportal der Europäischen Arzneimittel-Agentur (*European Medicines Agency*, EMA) sowie das Arzneimittel-Informationssystem (AMIce) nach zusätzlicher Evidenz durchsucht.

Im Rahmen dieser Suche wurde die Studie JADE DARE als einzige nutzenbewertungsrelevante Studie im zugelassenen Anwendungsgebiet von Abrocitinib bestätigt. Der Studienbericht dieser Studie diente als Hauptquelle, für die im vorliegenden Dokument berichteten Resultate. Darüber hinaus wurden Zusatzanalysen durchgeführt.

Ein-/ Ausschlusskriterien für Studien

Die für die Identifizierung von im Rahmen der vorliegenden Nutzenbewertung relevanten Studien verwendeten Ein- und Ausschlusskriterien sind in Tabelle 4-1 zusammengefasst und begründet.

Tabelle 4-1: Prädefinierte Ein-/ Ausschlusskriterien für die Selektion von RCT für die Nutzenbewertung – Suche nach RCT

#	Kriterium	Einschlusskriterien (E)	Ausschlusskriterien (A)	Begründung
1	Studienpopulation	Erwachsene Patienten mit mittelschwerer bis schwerer atopischer Dermatitis, die für eine systemische Therapie infrage kommen	Studienpopulation nicht der Zulassung entsprechend	Konsistenz mit der Fachinformation
2	Intervention	Therapie mit Abrocitinib (CIBINQO®, PF-04965842) gemäß Zulassung	Keine Therapie mit Abrocitinib (CIBINQO®, PF-04965842) gemäß der Zulassung	Gegenstand der Nutzenbewertung und Konsistenz mit der Fachinformation
3	Vergleichstherapie	Dupilumab (ggf. in Kombination mit TCS und/ oder TCI) gemäß Zulassung	Keine Therapie mit der vom G-BA festgelegten zVT: Dupilumab (ggf. in Kombination mit TCS und/ oder TCI) gemäß der Zulassung	Festlegung der zVT durch den G-BA (3)

#	Kriterium	Einschlusskriterien (E)	Ausschlusskriterien (A)	Begründung
4	Endpunkte ^a	Patientenrelevante Endpunkte aus den Nutzendimensionen: Mortalität, Morbidität, Lebensqualität, Verträglichkeit	Keine patientenrelevanten Endpunkte, die zur Bewertung von Mortalität, Morbidität, Lebensqualität oder Verträglichkeit verwendet werden können	Auswahl der Nutzendimensionen gemäß § 5 Absatz 2 VerfO (4)
5	Studientyp	Randomisierte kontrollierte klinische Studien	Nicht randomisierte, unkontrollierte Studien; nicht-klinische Studien; Fallstudien; Fallserien	Studientyp mit höchster Ergebnissicherheit
6	Studiendauer ^b	≥ 24 Wochen	< 24 Wochen	Angemessene Mindest-Studien-dauer im Anwendungsgebiet laut G-BA (5)
7	Publikationstyp ^c	Berichterstattung liefert ausreichende Informationen zur Beurteilung von Methodik/ Ergebnissen (z. B. Vollpublikation, Ergebnisbericht aus einem Studienregister oder Studienbericht, der den Kriterien des CONSORT Statements genügt)	Berichterstattung liefert keine ausreichenden Informationen zur Beurteilung von Methodik/ Ergebnissen (z. B. narrativer <i>Review</i> , <i>Letter</i> , <i>Editorial</i> , <i>Errata</i> , <i>Note</i> , Konferenz- <i>Abstract</i> oder <i>-Paper</i> , <i>Short Survey</i> , <i>Case Report</i> oder unsystematische Übersichtsarbeit)	Anforderung an eine ausreichende Datenverfügbarkeit zur Bewertung von Methodik und Ergebnissen

a: Auf eine Aufzählung aller möglichen oder wahrscheinlich verwendeten Messinstrumente und Operationalisierungen von Endpunkten wird an dieser Stelle verzichtet. Endpunkte zur Mortalität, Morbidität, Lebensqualität und Sicherheit werden in den Tabellen (Tabelle 4-20, Tabelle 4-26, Tabelle 4-32, Tabelle 4-43, Tabelle 4-56, Tabelle 4-69, Tabelle 4-80, Tabelle 4-86 und Tabelle 4-92) operationalisiert.

b: Eine solche Mindestbehandlungsdauer wird vom G-BA für Studien zur Nutzenbewertung für Arzneimittel zur Behandlung chronischer Erkrankungen regelhaft gefordert. Zur Rechtfertigung eines Evidenztransfers können in der vorliegenden Nutzenbewertung auch Studien berücksichtigt werden, deren Studiendauer von der vom G-BA geforderten Mindestdauer abweichen.

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

CONSORT: *Consolidated Standards of Reporting Trials*; EMA: Europäische Arzneimittel-Agentur (*European Medicines Agency*); G-BA: Gemeinsamer Bundesausschuss; ggf.: Gegebenenfalls; TCI: Topische Calcineurin-Inhibitoren; TCS: Topische Glukokortikoide (*Topical corticosteroids*); VerfO: Verfahrensordnung; z.B.: Zum Beispiel; zVT: zweckmäßige Vergleichstherapie

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

Im vorliegenden Modul 4 A wurden die verwendeten Nachweise hinsichtlich ihrer Aussagekraft sowohl auf Studienebene als auch auf Ebene der betrachteten Endpunkte gemäß Vorgaben in Anhang 4-F untersucht.

Für die Darstellung des medizinischen Nutzens und Zusatznutzens wurden die Ergebnisse der multizentrischen, randomisierten, doppelblinden, *double-dummy*, aktivkontrollierten Phase-III-Studie JADE DARE herangezogen, welche nach § 5 des 5. Kapitels der Verfahrensordnung (VerfO) des G-BA der Evidenzstufe Ib (randomisierte klinische Studie) zuzuordnen ist. Die Studie JADE DARE war nicht Teil der initialen Einreichung bei der EMA und wurde erst nachträglich in das Studienprogramm für Abrocitinib aufgenommen, um den Anforderungen des G-BA an die Dauer von Studien zur Nutzenbewertung von Arzneimitteln zur Behandlung chronischer Erkrankungen zu entsprechen. Die Studie schloss erwachsene Patienten mit mittelschwerer bis schwerer AD ein, die gemäß Leitlinienkriterien für eine systemische Therapie infrage kamen und beinhaltet einen direkten Vergleich mit der vom G-BA festgelegten zVT Dupilumab (Vorgangsnummer 2021-B-183-z) (3). In beiden Behandlungsarmen der Studie konnten die Patienten eine Hintergrundtherapie mit einer nicht-medikamentösen AD-Therapie (z. B. Emollentien) sowie mit einer medikamentösen AD-Therapie (schwach bis mittelstark wirksame TCS/ TCI) verwenden. Mit einer für die Nutzenbewertung bei chronischen Erkrankungen angemessenen Behandlungsdauer von 26 Wochen erlaubt die Studie eine Bewertung patientenrelevanter Endpunkte im vorliegenden Anwendungsgebiet. Das Verzerrungspotenzial der Studie JADE DARE wird als niedrig eingestuft.

Alle in der vorliegenden Nutzenbewertung zur Quantifizierung des Zusatznutzens herangezogenen Endpunkte sind patientenrelevant. Sie wurden mit validen Instrumenten erhoben und unter Verwendung adäquater Methoden analysiert.

Die Aussagesicherheit der für die Nutzenbewertung vorliegenden Studie JADE DARE ist sowohl auf Studien- als auch auf Endpunktebene durchweg als hoch einzuschätzen, sodass Aussagen zum medizinischen Zusatznutzen als **Hinweis** zu bewerten sind.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

In der folgenden Tabelle 4-2 wird die Wahrscheinlichkeit und das Ausmaß des Zusatznutzens von Abrocitinib gegenüber der zVT Dupilumab bei erwachsenen Patienten mit mittelschwerer bis schwerer AD zu Woche 26 für die betrachteten Endpunkte der Studie JADE DARE aufgezeigt.

Tabelle 4-2: Wahrscheinlichkeit und Ausmaß des Zusatznutzens

Endpunkt	Abrocitinib 200 mg vs. Dupilumab 300 mg Ereignisanteile % vs. % Effektschätzer [95 %-KI]; p-Wert ^a	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
Morbidität		
EASI		Hinweis auf einen geringen Zusatznutzen
Verbesserung um ≥ 75 %	70,2 % vs. 71,5 % RR = 0,98 [0,89; 1,08]; p = 0,6980	
Verbesserung um ≥ 90 %	52,5 % vs. 47,1 % RR = 1,11 [0,96; 1,29]; p = 0,1472	
SCORAD		
Verbesserung um ≥ 75 %	42,0 % vs. 36,4 % RR = 1,15 [0,96; 1,38]; p = 0,1282	
Verbesserung um ≥ 90 %	22,1 % vs. 14,3 % RR = 1,55 [1,13; 2,13]; p = 0,0067	
POEM		
Gesamtscore = 0-2	29,6 % vs. 19,0 % RR = 1,56 [1,19; 2,03]; p = 0,0011	
Verbesserung um ≥ 5 Punkte	76,1 % vs. 79,6 % RR = 0,96 [0,88; 1,03]; p = 0,2519	
Peak Pruritus NRS		
Verbesserung um ≥ 4 Punkte	67,5 % vs. 62,9 % RR = 1,07 [0,96; 1,19]; p = 0,1978	
SCORAD VAS für Juckreiz		
Verbesserung um ≥ 2 Punkte	75,4 % vs. 80,5 % RR = 0,94 [0,87; 1,01]; p = 0,0926	
MOS Sleep Scale Score Index-1		
Verbesserung um ≥ 15 Punkte	36,2 % vs. 32,2 % RR = 1,12 [0,92; 1,37]; p = 0,2642	
MOS Sleep Scale Score Index-2		
Verbesserung um ≥ 15 Punkte	38,4 % vs. 38,5 % RR = 1,00 [0,83; 1,20]; p = 0,9724	
SCORAD VAS für Schlafstörung		
Verbesserung um ≥ 2 Punkte	76,9 % vs. 79,6 % RR = 0,97 [0,89; 1,05]; p = 0,3960	
Skin Pain NRS		
Verbesserung um ≥ 4 Punkte	64,9 % vs. 62,2 % RR = 1,04 [0,93; 1,17]; p = 0,4745	
EQ-5D VAS		
Verbesserung um ≥ 15 Punkte	3,6 % vs. 1,7 % RR = 2,20 [0,85; 5,71]; p = 0,1040	

Endpunkt	Abrocitinib 200 mg vs. Dupilumab 300 mg Ereignisanteile % vs. % Effektschätzer [95 %-KI]; p-Wert ^a	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
Lebensqualität		
DLQI		Zusatznutzen nicht belegt
Gesamtscore = 0 oder 1	38,3 % vs. 31,6 % RR = 1,21 [0,99; 1,48]; p = 0,0603	
Verbesserung um ≥ 5 Punkte	72,3 % vs. 78,3 % RR = 0,92 [0,85; 1,01]; p = 0,0707	
Verträglichkeit		
Gesamtraten		Zusatznutzen nicht belegt
UE jeglichen Schweregrades (ergänzend)	74,0 % vs. 65,5 % RR = 1,13 [1,03; 1,24]; p = 0,0124	
Schwere UE	3,0 % vs. 2,2 % RR = 1,39 [0,56; 3,41]; p = 0,4763	
SUE	1,7 % vs. 1,6 % RR = 1,01 [0,33; 3,10]; p = 0,9885	
UE, die zum Therapieabbruch führten	2,5 % vs. 2,5 % RR = 1,01 [0,40; 2,51]; p = 0,9859	
UE jeglichen Schweregrades nach SOC und PT		
SOC: Infektionen und parasitäre Erkrankungen	18,8 % vs. 20,8 % RR = 0,90 [0,67; 1,21]; p = 0,4911	
PT: COVID-19	4,1 % vs. 3,3 % RR = 1,26 [0,60; 2,65]; p = 0,5427	
PT: Konjunktivitis	2,2 % vs. 9,6 % RR = 0,23 [0,11; 0,49]; p = 0,0001	
PT: Follikulitis	3,3 % vs. 0,8 % RR = 4,03 [1,15; 14,17]; p = 0,0296	
PT: Herpes simplex	3,3 % vs. 1,4 % RR = 2,42 [0,86; 6,80]; p = 0,0936	
PT: Nasopharyngitis	3,9 % vs. 3,3 % RR = 1,18 [0,55; 2,51]; p = 0,6742	
PT: Oraler Herpes	2,5 % vs. 4,1 % RR = 0,60 [0,27; 1,36]; p = 0,2260	
PT: Infektion der oberen Atemwege	2,8 % vs. 2,5 % RR = 1,12 [0,46; 2,72]; p = 0,8022	
SOC: Erkrankungen des Gastrointestinaltrakts	19,3 % vs. 3,0 % RR = 6,42 [3,46; 11,91]; p < 0,0001	
PT: Übelkeit	19,3 % vs. 2,2 % RR = 8,82 [4,31; 18,07]; p < 0,0001	
PT: Erbrechen	3,0 % vs. 1,6 % RR = 1,85 [0,69; 4,95]; p = 0,2211	

Endpunkt	Abrocitinib 200 mg vs. Dupilumab 300 mg Ereignisanteile % vs. % Effektschätzer [95 %-KI]; p-Wert ^a	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort PT: Erschöpfung	2,8 % vs. 1,4 % RR = 2,02 [0,70; 5,84]; p = 0,1962 2,8 % vs. 1,4 % RR = 2,02 [0,70; 5,84]; p = 0,1962	
SOC: Untersuchungen PT: Kreatinphosphokinase im Blut erhöht PT: Abnahme der natürlichen Killerzellen PT: SARS-CoV-2 positiv	10,5 % vs. 7,1 % RR = 1,47 [0,91; 2,37]; p = 0,1112 3,9 % vs. 3,6 % RR = 1,09 [0,52; 2,28]; p = 0,8275 2,8 % vs. 0,0 % RR = 20,19 [1,18; 344,43]; p = 0,0378 4,1 % vs. 3,6 % RR = 1,16 [0,56; 2,41]; p = 0,6838	
SOC: Erkrankungen des Nervensystems PT: Schwindelgefühl PT: Kopfschmerzen	15,2 % vs. 7,4 % RR = 2,05 [1,33; 3,18]; p = 0,0012 2,8 % vs. 1,1 % RR = 2,52 [0,80; 7,96]; p = 0,1152 13,0 % vs. 6,6 % RR = 1,97 [1,23; 3,16]; p = 0,0045	
SOC: Erkrankungen der Haut und des Unterhautgewebes PT: Akne PT: Dermatitis atopisch	16,9 % vs. 6,3 % RR = 2,67 [1,69; 4,22]; p < 0,0001 12,7 % vs. 2,7 % RR = 4,64 [2,38; 9,05]; p < 0,0001 4,7 % vs. 3,8 % RR = 1,22 [0,61; 2,45]; p = 0,5666	
Spezifische UE jeglichen Schweregrades		
Superinfektionen	8,6 % vs. 6,0 % RR = 1,42 [0,84; 2,41]; p = 0,1912	
Herpes zoster	2,5 % vs. 0,6 % RR = 4,54 [0,99; 20,85]; p = 0,0520	
Konjunktivitis	2,8 % vs. 10,7 % RR = 0,26 [0,13; 0,51]; p < 0,0001	
Akne	13,3 % vs. 3,0 % RR = 4,40 [2,32; 8,33]; p < 0,0001	
Follikulitis	3,0 % vs. 0,8 % RR = 3,70 [1,04; 13,14]; p = 0,0433	
Analysen mit signifikantem Gruppenunterschied wurden fett markiert. a: In den Analysen zur Morbidität und Lebensqualität wurde der Effektschätzer inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4). In den Analysen zur Verträglichkeit wurde der Effektschätzer inklusive 95 %-KI und p-Wert berechnet anhand der unstratifizierten Approximation der Normalverteilung nach der Wald-Methode.		

Endpunkt	Abrocitinib 200 mg vs. Dupilumab 300 mg Ereignisanteile % vs. % Effektschätzer [95 %-KI]; p-Wert ^a	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
CoV: Coronavirus; DLQI: <i>Dermatology Life Quality Index</i> ; EASI: <i>Eczema Area and Severity Index</i> ; EQ-5D: <i>European Quality of Life Group 5 Dimensions</i> ; IGA: <i>Investigator's Global Assessment</i> ; KI: Konfidenzintervall; mg: Milligramm; MOS: <i>Medical Outcomes Study</i> ; NRS: <i>Numerical Rating Scale</i> ; POEM: <i>Patient-Oriented Eczema Measures</i> ; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SARS: <i>Severe Acute Respiratory Syndrome</i> ; SCORAD: <i>Scoring Atopic Dermatitis</i> ; SOC: Systemorganklasse (<i>System Organ Class</i>); UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala; vs.: Versus		

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Morbidität

Die Kategorie Morbidität umfasst AD-relevante Endpunkte, wie *Eczema Area and Severity Index* (EASI), *Scoring Atopic Dermatitis* (SCORAD), patientenberichtete Symptomatik, Juckreiz, Schlafstörungen, Hautschmerz, Ängste und Depressionen sowie den Gesundheitszustand.

Eine deutliche Verbesserung des therapierelevanten Nutzens gegenüber der zVT zeigte sich in der Studie JADE DARE zu Woche 26 in einem statistisch signifikant höheren Anteil an Patienten mit einer 90%igen Verbesserung des Schweregrades der AD gemessen anhand des SCORAD (SCORAD-90) und anhand eines statistisch signifikant höheren Anteils an Patienten, die zu Woche 26 nahezu beschwerdefrei waren, gemessen mit dem patientenberichteten Symptomfragebogen *Patient-Oriented Eczema Measure* (POEM)-Gesamtscore von 0-2. Die dargestellten Sensitivitätsanalysen zum SCORAD-90 und POEM-Gesamtscore von 0-2 unterstreichen die Robustheit der Ergebnisse der Hauptanalyse und zeigten ebenfalls statistisch signifikante Unterschiede zugunsten von Abrocitinib gegenüber Dupilumab. Auch wurde eine vergleichbare Verbesserung des Gesundheitszustandes, sowie eine vergleichbare Verminderung des Juckreizes, der Schlafstörungen, des Hautschmerzes und der Reduzierung von Ängsten und Depressionen zwischen den Behandlungsarmen festgestellt.

Schneller Wirkeintritt

Auch wenn die langfristige Krankheitskontrolle im Mittelpunkt der AD-Therapie steht, sind auch kurzfristige Effekte geeignet einen patientenrelevanten Zusatznutzen bei Patienten mit moderater bis schwerer AD zu zeigen. Erste Behandlungseffekte, wie z. B. die Reduktion des Juckreizes und die Verbesserung der Hautscores (EASI und SCORAD), traten mit Abrocitinib früher auf als mit Dupilumab. Dies wird durch statistisch signifikante und klinisch relevante Gruppenunterschiede bei der Veränderung der Hautscores zugunsten von Abrocitinib zwischen Woche 2 und Woche 16 belegt, die sich auch in einer Verbesserung der Lebensqualität widerspiegeln (siehe Abschnitt 4.3.1.3.1, Abbildung 4-3, Abbildung 4-4, Abbildung 4-6, Abbildung 4-7, Abbildung 4-8 und Abbildung 4-16).

Insgesamt ergibt sich für die Behandlung mit Abrocitinib gegenüber Dupilumab aufgrund des statistisch signifikant größeren Anteils an Patienten mit einer deutlichen Reduktion der Krankheitschwere und des signifikant höheren Anteils an Patienten, die eine Symptombefreiheit oder annähernde Symptombefreiheit erreichten, in der Kategorie Morbidität ein **Hinweis auf einen geringen Zusatznutzen**.

Gesundheitsbezogene Lebensqualität

Abrocitinib zeigte im *Dermatology Life Quality Index* (DLQI) zu Woche 26 eine vergleichbare Verbesserung in der krankheitsspezifischen gesundheitsbezogenen Lebensqualität gegenüber der hoch wirksamen zVT Dupilumab. Demnach ergibt sich in der Kategorie gesundheitsbezogene Lebensqualität **kein Beleg für einen Zusatznutzen** von Abrocitinib bei erwachsenen Patienten, die für eine systemische Therapie infrage kommen, gegenüber Dupilumab.

Verträglichkeit

In der Studie JADE DARE traten schwere oder schwerwiegende unerwünschte Ereignisse (SUE) in beiden Behandlungsarmen selten auf; die entsprechenden Raten waren zwischen den Behandlungsgruppen vergleichbar. Ebenso niedrig und zwischen den Behandlungsgruppen vergleichbar war der Anteil der Patienten, die die Therapie aufgrund von unerwünschten Ereignissen (UE) abbrachen. Die Gesamtrate der UE jeglichen Schweregrades wurde zur Bewertung des Zusatznutzens nicht herangezogen, sondern lediglich ergänzend dargestellt. Bei den beobachteten Unterschieden wird von einem nicht mehr als geringfügigen Effekt ausgegangen.

Bei den spezifischen UE zeigte sich ein statistisch signifikanter Unterschied zuungunsten von Abrocitinib bei den Endpunkten Übelkeit, Kopfschmerzen, Akne, Follikulitis und Abnahme der natürlichen Killerzellen sowie bei den entsprechenden Systemorganklassen (*System Organ Class*, SOC) Erkrankungen des Gastrointestinaltrakts, Erkrankungen des Nervensystems und Erkrankungen der Haut und des Unterhautgewebes. Bei den UE handelte es sich mehrheitlich um nicht schwere, nicht schwerwiegende Ereignisse, die nicht zum Abbruch der Therapie führten, gut beherrschbar waren und somit keine Auswirkungen auf die Lebensqualität der Patienten hatten. Für den Endpunkt Konjunktivitis zeigte sich ein statistisch signifikanter Unterschied zugunsten von Abrocitinib.

In der Kategorie Verträglichkeit lässt sich somit aufgrund vergleichbarer, in beiden Behandlungsarmen niedriger Gesamtraten schwerer UE, SUE) und der geringen und vergleichbaren Zahl an Therapieabbrüchen wegen UE, **kein Beleg für einen Zusatznutzen** von Abrocitinib gegenüber Dupilumab ableiten.

Gesamtschau des medizinischen Zusatznutzens

Mit Abrocitinib, als einem oral einzunehmenden Wirkstoff, steht eine wirksame und gut verträgliche Therapiemöglichkeit zur Behandlung von mittelschwerer bis schwerer AD bei erwachsenen Patienten, die für eine systemische Therapie infrage kommen, zur Verfügung.

Gemäß der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) § 35a Absatz 1 Sozialgesetzbuch (SGB) V für Erstattungsvereinbarungen nach § 130b SGB V konnte in den Wirksamkeitsergebnissen durch die Behandlung mit Abrocitinib eine bisher nicht erreichte deutliche Verbesserung des therapierlevanten Nutzens festgestellt werden. Dieser besteht vor allem in einer Reduzierung des Schweregrades der AD und der damit einhergehenden spürbaren Linderung der Erkrankung im Sinne der patientenberichteten Symptomatik bei einer allgemein guten Verträglichkeit. Es wird daher ein **Hinweis auf einen geringen Zusatznutzen** abgeleitet.

4.2 Methodik

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

4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

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

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

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

Die Fragestellung für das vorliegende Modul 4 A ist die Bewertung der Wahrscheinlichkeit und des Ausmaßes des Zusatznutzens hinsichtlich patientenrelevanter Endpunkte für die Behandlung der mittelschweren bis schweren AD durch Abrocitinib (CIBINQO®) im Vergleich zur zVT Dupilumab bei Erwachsenen, die für eine systemische Therapie infrage kommen.

Patientenpopulation

Abrocitinib wird angewendet zur Behandlung von mittelschwerer bis schwerer AD bei Erwachsenen, die für eine systemische Therapie infrage kommen (1).

Schweregrad der Erkrankung

Gemäß der Aktualisierung der „Systemtherapie bei Neurodermitis“ zur S2k-Leitlinie Neurodermitis der AWMF, wird die Schweregradbeurteilung durch klinische objektive Schweregrad-Scores wie den EASI, die von der AD betroffene Körperoberfläche (*Body Surface Area*, BSA), den SCORAD bzw. *Objective Scoring Atopic Dermatitis* (oSCORAD) oder den globalen Schweregrad anhand des *Physician Global Assessment* (PGA) empfohlen (6).

Dementsprechend liegt eine mittelschwere bis schwere AD vor, wenn mindestens eines der folgenden Kriterien erfüllt ist:

- EASI-Score von > 15
- SCORAD-Score von $> 40/$ oSCORAD > 20
- PGA-Score von ≥ 3
- BSA von $> 10 \%$
- das Vorliegen von therapierefraktären Ekzemen an sensitiven und/ oder sichtbaren Arealen oder einer hohen Frequenz von Schüben (> 10 Schübe pro Jahr) unter der derzeitigen Therapie

Des Weiteren wird die Erfassung der subjektiven relevanten Belastung empfohlen, die vorliegt, wenn mindestens eines der folgenden Kriterien erfüllt ist:

- DLQI-Score von > 10
- Pruritus-Score von > 6 (auf einer Skala von 0 bis 10 Punkten auf der visuellen Analogskala (VAS) oder der *Numerical Rating Scale* (NRS))
- eine relevante Störung des Nachtschlafs aufgrund von Pruritus oder eines Ekzems

Eignung für eine systemische Therapie

Die Eignung für eine systemische Therapie basiert auf einer Kombination aus dem relevanten objektiven Schweregrad (1. Kriterium), der relevanten subjektiven Belastung (2. Kriterium) sowie dem fehlenden oder unzureichendem Ansprechen (3. Kriterium) auf eine Lokal- oder Phototherapie, eine bereits eingeleitete Systemtherapie ohne Erfolg oder die fehlende Aussicht auf Erfolg mit lokalen bzw. topischen Interventionen. Die Indikation für eine systemische Therapie ist gegeben, wenn aus jedem der 3 Bereiche mindestens ein Kriterium zutrifft (6).

Intervention

Die zu bewertende Intervention ist Abrocitinib gemäß den Angaben in der Fachinformation. Diese sieht eine 1-mal tägliche orale Gabe von 200 mg Abrocitinib vor. Abrocitinib kann alleine oder mit anderen Arzneimitteln zur topischen Anwendung bei AD angewendet werden (1). Die begleitenden topischen Therapien umfassen das Auftragen von Emollientien und/ oder von leichten bis mittelstark wirksamen TCS oder TCI.

Vergleichstherapie

Der G-BA hat im Anwendungsgebiet von Abrocitinib die folgende zVT bestimmt (siehe Modul 3 A, Abschnitt 3.1):

- Dupilumab (ggf. in Kombination mit TCS und/ oder TCI)

Die Vergleichstherapie besteht aus der subkutanen Gabe von Dupilumab gemäß der Fachinformation (2). Die Vergleichstherapie besteht somit für Erwachsene aus einer Initialdosis von 600 mg Dupilumab (2 aufeinanderfolgende Injektionen zu je 300 mg) an Tag 1 und der

anschließenden Gabe von 300 mg Dupilumab 1-mal alle 2 Wochen. Dupilumab kann mit oder ohne einer begleitenden topischen Therapie mit TCS oder TCI angewendet werden. Die Anwendung der TCI sollte dabei auf Problemzonen wie Gesicht, Hals, intertriginöse Bereiche und den Genitalbereich beschränkt werden.

Endpunkte

Unter Berücksichtigung der Nutzendimensionen Mortalität, Morbidität, Lebensqualität und Verträglichkeit erfolgt die Bewertung des Ausmaßes des medizinischen Nutzens und Zusatznutzens anhand der folgenden patientenrelevanten Endpunkte:

Nutzendimension	Endpunkte
Mortalität	<ul style="list-style-type: none"> • Gesamtüberleben
Morbidität	<ul style="list-style-type: none"> • Krankheitsschwere • Patientenberichtete Symptomatik • Juckreiz • Schlafstörungen • Hautschmerz • Ängste und Depressionen • Gesundheitszustand <p>Supportive Endpunkte:</p> <ul style="list-style-type: none"> • Betroffene Körperoberfläche
Gesundheitsbezogene Lebensqualität	<ul style="list-style-type: none"> • Krankheitsspezifische Lebensqualität
Verträglichkeit	<ul style="list-style-type: none"> • Unerwünschte Ereignisse (UE) jeglichen Schweregrades • Schwere UE • Schwerwiegende UE (SUE) • UE, die zum Therapieabbruch führten • Spezifische UE
SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis	

Studientypen

Es werden RCT eingeschlossen.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann

beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister/ einer Studienergebnisdatenbank erfolgen, während ein Kongressabstrakt allein in der Regel nicht für eine Studienbewertung ausreicht.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar. Erstellen Sie dabei für unterschiedliche Themen der Recherche (z. B. unterschiedliche Fragestellungen) jeweils eine separate Übersicht.

Tabelle 4-3: Prädefinierte Ein-/ Ausschlusskriterien für die Selektion von RCT für die Nutzenbewertung – Suche nach RCT

#	Kriterium	Einschlusskriterien (E)	Ausschlusskriterien (A)	Begründung
1	Studienpopulation	Erwachsene Patienten mit mittelschwerer bis schwerer atopischer Dermatitis, die für eine systemische Therapie infrage kommen	Studienpopulation nicht der Zulassung entsprechend	Konsistenz mit der Fachinformation
2	Intervention	Therapie mit Abrocitinib (CIBINQO®, PF-04965842) gemäß Zulassung	Keine Therapie mit Abrocitinib (CIBINQO®, PF-04965842) gemäß der Zulassung	Gegenstand der Nutzenbewertung und Konsistenz mit der Fachinformation
3	Vergleichstherapie	Dupilumab (ggf. in Kombination mit TCS und/ oder TCI) gemäß Zulassung	Keine Therapie mit der vom G-BA festgelegten zVT: Dupilumab (ggf. in Kombination mit TCS und/ oder TCI) gemäß der Zulassung	Festlegung der zVT durch den G-BA (3)
4	Endpunkte ^a	Patientenrelevante Endpunkte aus den Nutzendimensionen: Mortalität, Morbidität, Lebensqualität, Verträglichkeit	Keine patientenrelevanten Endpunkte, die zur Bewertung von Mortalität, Morbidität, Lebensqualität oder Verträglichkeit verwendet werden können	Auswahl der Nutzendimensionen gemäß § 5 Absatz 2 Verfo (4)
5	Studientyp	Randomisierte kontrollierte klinische Studien	Nicht randomisierte, unkontrollierte Studien; nicht-klinische Studien; Fallstudien; Fallserien	Studientyp mit höchster Ergebnissicherheit
6	Studiendauer ^b	≥ 24 Wochen	< 24 Wochen	Angemessene Mindest-Studiendauer im Anwendungsgebiet laut G-BA (5)

#	Kriterium	Einschlusskriterien (E)	Ausschlusskriterien (A)	Begründung
7	Publikationstyp ^c	Berichterstattung liefert ausreichende Informationen zur Beurteilung von Methodik/ Ergebnissen (z. B. Vollpublikation, Ergebnisbericht aus einem Studienregister oder Studienbericht, der den Kriterien des CONSORT Statements genügt)	Berichterstattung liefert keine ausreichenden Informationen zur Beurteilung von Methodik/ Ergebnissen (z. B. narrativer <i>Review, Letter, Editorial, Errata, Note</i> , Konferenz- <i>Abstract</i> oder <i>-Paper, Short Survey, Case Report</i> oder unsystematische Übersichtsarbeit)	Anforderung an eine ausreichende Datenverfügbarkeit zur Bewertung von Methodik und Ergebnissen
<p>a: Auf eine Aufzählung aller möglichen oder wahrscheinlich verwendeten Messinstrumente und Operationalisierungen von Endpunkten wird an dieser Stelle verzichtet. Endpunkte zur Mortalität, Morbidität, Lebensqualität und Sicherheit werden in den Tabellen (Tabelle 4-20, Tabelle 4-26, Tabelle 4-32, Tabelle 4-43, Tabelle 4-56, Tabelle 4-69, Tabelle 4-80, Tabelle 4-86 und Tabelle 4-92) operationalisiert.</p> <p>b: Eine solche Mindestbehandlungsdauer wird vom G-BA für Studien zur Nutzenbewertung für Arzneimittel zur Behandlung chronischer Erkrankungen regelhaft gefordert. Zur Rechtfertigung eines Evidenztransfers können in der vorliegenden Nutzenbewertung auch Studien berücksichtigt werden, deren Studiendauer von der vom G-BA geforderten Mindestdauer abweichen.</p> <p>c: Studien, deren Ergebnisse auf der Webseite offizieller Studienregister publiziert wurden (z. B. clinicaltrials.gov), können ebenfalls berücksichtigt werden.</p> <p>CONSORT: <i>Consolidated Standards of Reporting Trials</i>; EMA: Europäische Arzneimittel-Agentur (<i>European Medicines Agency</i>); G-BA: Gemeinsamer Bundesausschuss; ggf.: Gegebenenfalls; TCI: Topische Calcineurin-Inhibitoren; TCS: Topische Glukokortikoide (<i>Topical corticosteroids</i>); VerfO: Verfahrensordnung; z.B.: Zum Beispiel; zVT: zweckmäßige Vergleichstherapie</p>				

4.2.3 Informationsbeschaffung

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

4.2.3.1 Studien des pharmazeutischen Unternehmers

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

4.2.3.2 Bibliografische Literaturrecherche

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

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

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

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

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

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

Es wurde eine umfassende, systematische bibliografische Literaturrecherche nach RCT mit Abrocitinib im vorliegenden Anwendungsgebiet durchgeführt.

Die Suche nach relevanten Publikationen erfolgte am 16.11.2021 in den folgenden Literaturdatenbanken:

- EMBASE[®]
- MEDLINE[®]
- *Cochrane Central Register of Controlled Trials*

Die Suchstrategie wurde nach Intervention und Studientypen an die technischen Vorgaben der jeweiligen Datenbank angepasst. Für die Einschränkung der Suche hinsichtlich des Studientyps auf RCT wurde in den Datenbanken MEDLINE[®] und EMBASE[®] der validierte und in den allgemeinen Methoden des Instituts für Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (7) empfohlene Wong-Filter verwendet (8). In MEDLINE[®] wurde zusätzlich der Cochrane-Filter für RCT zu diesem Zweck eingesetzt. Eine zeitliche Einschränkung erfolgte nicht. Für die

Identifikation der relevanten Publikationen wurden die in Abschnitt 4.2.2 dargestellten Ein- und Ausschlusskriterien herangezogen. In die Suche wurde englisch- und deutschsprachige Literatur eingeschlossen.

Eine detaillierte Auflistung aller Suchstrategien bzw. Suchbegriffe samt Trefferzahlen findet sich in Anhang 4-A. Die Ergebnisse der Suche in den Studienregistern sind in Abschnitt 4.3.1.1.2 beschrieben.

4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken

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

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

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

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

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

Für Clinical Data (Suchportal der European Medicines Agency) und das Arzneimittel-Informationssystem (AMIS) genügt hingegen die Suche nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

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

Die Suche nach RCT mit dem zu bewertenden Arzneimittel sowie für indirekte Vergleiche erfolgte am 16.11.2021 in den folgenden Studienregistern und Studienergebnisdatenbanken:

- ClinicalTrials.gov (www.clinicaltrials.gov)
- ICTRP Suchportal der WHO (<https://trialsearch.who.int/Default.aspx>)
- EU-CTR (<http://www.clinicaltrialsregister.eu/ctrsearch>)

Die Suchstrategien wurden entsprechend den Eingabemöglichkeiten der durchsuchten Plattformen angepasst. Weitere Einschränkungen der Suche wurden nicht vorgenommen. Die Suchstrategien sind in Anhang 4-B dargestellt. Für die Identifikation der relevanten Studien wurden die in Abschnitt 4.2.2 dargestellten Ein- und Ausschlusskriterien herangezogen.

Darüber hinaus erfolgte am 16.11.2021 in folgenden Datenbanken eine Suche nach Einträgen mit Ergebnisberichten zu dem zu bewertenden Arzneimittel:

- Clinical Data Suchportal der EMA (<https://clinicaldata.ema.europa.eu>)
- Arzneimittel-Informationssystem (AMIS öffentlicher Teil, <https://www.dimdi.de/dynamic/de/medizinisch/medizinisch-recherchieren/amis/>)

4.2.3.4 Suche auf der Internetseite des G-BA

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

Auf der Internetseite des G-BA werden Dokumente zur frühen Nutzenbewertung nach §35a SGB V veröffentlicht. Diese enthalten teilweise anderweitig nicht veröffentlichte Daten zu Studienmethodik und –ergebnissen¹. Solche Daten sind dabei insbesondere in den Modulen 4

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

der Dossiers pharmazeutischer Unternehmer, in IQWiG-Nutzenbewertungen sowie dem Beschluss des G-BA einschließlich der Tragenden Gründe und der Zusammenfassenden Dokumentation zu erwarten.

Die Suche auf der Internetseite des G-BA muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche auf der G-BA Internetseite immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird. Die Suche ist dann sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie durchzuführen. Es genügt die Suche nach Einträgen zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

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

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

Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Die Suche auf der Internetseite des G-BA nach RCT mit dem zu bewertenden Arzneimittel Abrocitinib erfolgte am 16.11.2021 nach Einträgen zu RCT, die bereits anderweitig identifiziert wurden. Es wurde sowohl nach dem Wirkstoff als auch nach den relevanten Studiennummern gesucht.

4.2.3.5 Selektion relevanter Studien

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

Die in der bibliografischen Literaturrecherche und der Studienregistersuche identifizierten Publikationen bzw. Studien wurden zunächst auf Titel-/ Abstract-Ebene und anschließend auf Volltext-Ebene anhand der in Abschnitt 4.2.2 in Tabelle 4-3 erläuterten Ein- und Ausschlusskriterien bewertet. Die Bewertung wurde von zwei Reviewern unabhängig voneinander vorgenommen, wobei eventuelle Abweichungen in der Bewertung bis zu einer Konsensfindung diskutiert wurden. Gesichtete und ausgeschlossene Studien auf Volltext-Ebene aus der bibliografischen Literaturrecherche werden in Anhang 4-C präsentiert. Die

ausgeschlossenen Studien aus der Suche in Studienregistern sind mit Ausschlussgrund in Anhang 4-D dokumentiert.

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.

Im vorliegenden Modul 4 A (Abschnitt 4.3.1.2.2 und Abschnitte 4.3.1.3.1.1 bis 4.3.1.3.1.3) werden die verwendeten Nachweise hinsichtlich ihrer Aussagekraft sowohl auf Studienebene als auch auf Ebene der betrachteten Endpunkte gemäß den Vorgaben in Anhang 4-Funtersucht. Die Bewertung von Verzerrungsaspekten erfolgt anhand der zur Verfügung stehenden Informationen aus den Studienprotokollen und Studienberichten.

Die Bewertung der Aussagekraft der Nachweise zum Zusatznutzen von Abrocitinib gegenüber der zVT wird in Abschnitt 4.4.1 zusammenfassend beschrieben.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)². Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-³ bzw. STROBE-Statements⁴ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

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

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

⁴ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.

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

Die zusammenfassende Beschreibung von Design und Methodik der eingeschlossenen Studie erfolgte anhand des *Consolidated Standards of Reporting Trials (CONSORT)*-Statements. Es wurde gemäß Modulvorlage das CONSORT-Statement in Anhang 4-E ausgefüllt und ein Flow-Chart erstellt. Darüber hinaus wurden wichtige Studiencharakteristika in Abschnitt 4.3.1.2.1 dargestellt.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

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

Gegenstand der Bewertung ist die Studie JADE DARE mit dem zu bewertenden Arzneimittel Abrocitinib.

Patientencharakteristika

Anhand folgender demografischer und krankheitsspezifischer Charakteristika, die zum Zeitpunkt der Randomisierung erhoben wurden, werden die Patienten in der Studie JADE DARE (siehe Tabelle 4-4) beschrieben.

Tabelle 4-4: Patientencharakteristika der Studie JADE DARE

Patientencharakteristika	Statistische Maße und Strata
Demografische Charakteristika	
Geschlecht	n (%) Männlich Weiblich
Alter (Jahre)	n MW (SD) Median (min-max)
Altersgruppe	n (%) < 40 Jahre ≥ 40 Jahre < 65 Jahre ≥ 65 Jahre
Abstammung	n (%) Kaukasisch Schwarz oder afroamerikanisch Asiatisch Andere
Region	n (%) Vereinigte Staaten/ Kanada/ Australien Europa Asien Lateinamerika
Gewicht (kg)	n MW (SD) Median (min-max)
Gewichtsguppe	n (%) < 70 kg 70 bis 100 kg > 100 kg
Krankheitsspezifische Charakteristika	
Erkrankungsdauer (Jahre)	n MW (SD) Median (min-max)
Erkrankungsdauer-Gruppe	n (%) < 26 Jahre ≥ 26 Jahre
IGA-Score	n (%) Moderat (IGA=3) Schwer (IGA=4)

Patientencharakteristika	Statistische Maße und Strata
EASI-Score	n MW (SD) Median (min-max)
EASI-Gruppe	n (%) < 16 Punkte 16-25 Punkte > 25 Punkte
Betroffene BSA (%)	n MW (SD) Median (min-max)
Betroffene BSA (%)-Gruppe	n (%) 10-30 % > 30-50 % > 50 %
Peak Pruritus NRS-Score	n MW (SD) Median (min-max)
Peak Pruritus NRS-Gruppe	n (%) > 4 Punkte 4-6 Punkte > 7 Punkte
SCORAD-Score	n MW (SD) Median (min-max)
Vorbehandlung mit Ciclosporin	n (%) Therapienaiv Vorbehandelt
Vorherige Medikation für AD	n (%) Nur topische Medikation Systemische Medikation
AD: Atopische Dermatitis; BMI: <i>Body Mass Index</i> ; BSA: Körperoberfläche (<i>Body Surface Area</i>); EASI: <i>Eczema Area and Severity Index</i> ; IGA: <i>Investigator's Global Assessment</i> ; kg: Kilogramm; max: Maximal; min: Minimal; MW: Mittelwert; n: Anzahl der Patienten; NRS: <i>Numerical Rating Scale</i> ; SCORAD: <i>Scoring Atopic Dermatitis</i> ; SD: Standardabweichung (<i>Standard Deviation</i>)	

Patientenrelevante Endpunkte

Zur Ableitung des medizinischen Nutzens und Bewertung des Ausmaßes des medizinischen Zusatznutzens wurden gemäß § 3 Absatz 1 und 2, 5. Kapitel der Verfo des G-BA Ergebnisse zu patientenrelevanten Endpunkten aus den Nutzendimensionen Morbidität, gesundheitsbezogene Lebensqualität und Verträglichkeit herangezogen.

Die patientenrelevanten Endpunkte sind in Tabelle 4-5 dargestellt. Die Argumentation zur Patientenrelevanz, Validität und Operationalisierung dieser Endpunkte im Anwendungsgebiet erfolgt im Anschluss.

Tabelle 4-5: Übersicht der im Nutzendossier dargestellten patientenrelevanten Endpunkte

Nutzendimension	Endpunkte
Morbidität	<ul style="list-style-type: none"> • Krankheitsschwere (EASI, SCORAD) • Patientenberichtete Symptomatik (POEM) • Juckreiz (<i>Peak Pruritus</i> NRS, SCORAD VAS für Juckreiz, Einzelitem des POEM) • Schlafstörungen (MOS <i>Sleep Scale</i>, SCORAD VAS für Schlafstörung, Einzelitem des POEM) • Hautschmerz (<i>Skin Pain</i> NRS) • Ängste und Depressionen (HADS) • Gesundheitszustand (EQ-5D VAS) <p>Supportive Endpunkte:</p> <ul style="list-style-type: none"> • Betroffene BSA
Gesundheitsbezogene Lebensqualität	<ul style="list-style-type: none"> • Krankheitsspezifische Lebensqualität (DLQI)
Verträglichkeit	<ul style="list-style-type: none"> • UE jeglichen Schweregrades <ul style="list-style-type: none"> ○ Gesamtrate ○ Häufige UE nach SOC und PT • Schwere UE <ul style="list-style-type: none"> ○ Gesamtrate ○ Häufige schwere UE nach SOC und PT • SUE <ul style="list-style-type: none"> ○ Gesamtrate ○ Häufige SUE nach SOC und PT • UE, die zum Therapieabbruch führten <ul style="list-style-type: none"> ○ Gesamtrate ○ Darstellung nach SOC und PT (deskriptiv) • Spezifische UE, jeweils als Gesamtrate und differenziert nach Schweregrad (schwer, schwerwiegend): <ul style="list-style-type: none"> ○ Superinfektionen ○ Herpes zoster ○ Konjunktivitis ○ Akne ○ Follikulitis
<p>BSA: Körperoberfläche (<i>Body Surface Area</i>); DLQI: <i>Dermatology Life Quality Index</i>; EASI: <i>Eczema Area and Severity Index</i>; EQ-5D: <i>European Quality of Life Group 5 Dimensions</i>; HADS: <i>Hospital Anxiety and Depression Scale</i>; MOS: <i>Medical Outcomes Study</i>; NRS: <i>Numerical Rating Scale</i>; POEM: <i>Patient-Oriented Eczema Measure</i>; PT: <i>Bevorzugte Bezeichnung (Preferred Term)</i>; SCORAD: <i>Scoring Atopic Dermatitis</i>; SOC: <i>Systemorganklasse (System Organ Class)</i>; SUE: <i>Schwerwiegendes unerwünschtes Ereignis</i>; UE: <i>Unerwünschtes Ereignis</i>; VAS: <i>Visuelle Analogskala</i></p>	

Morbidität

EASI

Patientenrelevanz

Der EASI ist ein Standardinstrument zur Einstufung des Schweregrades der AD durch den Arzt. Der EASI zählt als einer der besten und etabliertesten Instrumente zur Beurteilung der klinischen Anzeichen von AD. Er ist sowohl für die Diagnosestellung als auch Verlaufskontrolle der Krankheitsschwere in der Versorgung von Relevanz (9). Der EASI wurde vom G-BA in bisherigen Verfahren im Anwendungsgebiet als patientenrelevant anerkannt (10, 11), da er die Veränderungen der für die AD relevanten Symptome Erythem, Papulae/Induration, Exkoration sowie Lichenifikation der Haut untersucht. Laut G-BA sind im Anwendungsgebiet vor allem ein EASI 75-Ansprechen und ein EASI 90-Ansprechen als patientenrelevante Therapieziele einzustufen; ein EASI 100-Ansprechen kann – z. B. aufgrund verbleibender Lichenifikation – oft nicht erreicht werden (12).

Validität

Der EASI hat eine adäquate Validität, Ansprechbarkeit, interne Konsistenz, Intraobserver-Reliabilität und eine mittlere Interobserver-Reliabilität, jedoch eine unklare Interpretierbarkeit und Durchführbarkeit (13). In der Aktualisierung „Systemtherapie bei Neurodermitis“ zur deutschen S2k-Leitlinie Neurodermitis wird der EASI als validiertes Messinstrument zur Erfassung der Krankheitsschwere empfohlen (6).

Operationalisierung

Im EASI werden das flächenhafte Ausmaß (1-9 %, 10-29 %, 30-49 %, 50-69 %, 70-89 % und 90-100 %) und die Intensität (keine (0), leichte (1), mäßige (2) und starke (3)) der Hautveränderungen in den 4 Körperbereichen Kopf und Hals, Rumpf, obere Extremität und untere Extremität bestimmt. Betrachtet werden dabei die Symptome Erythem, Papulae/Induration, Exkoration und Lichenifikation. Basierend auf der Bewertung der Symptome und der Einschätzung der betroffenen BSA wird ein Gesamtscore gebildet. Der EASI-Gesamtscore kann Werte von 0-72 annehmen, wobei ein höherer Wert einem höheren Schweregrad entspricht (14).

Die Operationalisierung des EASI erfolgte anhand der Anzahl der Patienten, die von Studienbeginn bis zu Woche 26 jeweils eine Verbesserung des EASI-Scores um 75 % (EASI-75) bzw. 90 % (EASI-90) erreichten. Darüber hinaus wurde die prozentuale Änderung des EASI von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (siehe Abschnitt 4.2.5.4). In der Studie JADE DARE wurde der EASI zu Woche 2, 4, 8, 12, 16, 20 und 26 erfasst.

Im vorliegenden Modul 4 A wurde der Endpunkt EASI für die Studie JADE DARE in den Hauptanalysen somit operationalisiert als:

- Anteil der Patienten, die eine 75 %ige Verbesserung (Reduktion) gegenüber dem *Baseline*-Wert erreichen (EASI-75)
- Anteil der Patienten, die eine 90 %ige Verbesserung (Reduktion) gegenüber dem *Baseline*-Wert erreichen (EASI-90)

SCORAD

Patientenrelevanz

Der SCORAD ist ein etabliertes und validiertes Instrument zur Erfassung des flächenhaften Ausmaßes und der Intensität der Hautveränderungen durch den Arzt. Es erfolgt die Beurteilung der Veränderungen der für die AD relevanten Symptome Erythem, Papulae/ Induration, nässende Areale/ Krustenbildung, Exkoration, Lichenifikation und Trockenheit nicht-betroffener Hautareale sowie Juckreiz und Schlafstörungen auf einer patientenberichteten VAS. Eine Verbesserung im SCORAD stellt somit ein patientenrelevantes Therapieziel im Anwendungsgebiet dar. Der G-BA stuft eine 75 %ige Verbesserung und eine 90 %ige Verbesserung gegenüber dem *Baseline*-Wert in den tragenden Gründen zum G-BA-Beschluss zu Dupilumab in der Indikation AD als patientenrelevant ein (10, 11), da eine 75 %ige bzw. 90 %ige Verbesserung bereits einem Therapieerfolg der Patienten entspricht.

Validität

Der SCORAD hat eine adäquate Validität, Ansprechbarkeit, Interobserver-Reliabilität und Interpretierbarkeit, jedoch eine unklare Intraobserver-Reliabilität (13). In der Aktualisierung „Systemtherapie bei Neurodermitis“ zur deutschen S2k-Leitlinie Neurodermitis wird der SCORAD als validiertes Messinstrument zur Erfassung des flächenhaften Ausmaßes und der Intensität der Hautveränderungen empfohlen (6).

Operationalisierung

Im SCORAD wird der Schweregrad (keine (0), leichte (1), mäßige (2) und starke (3)) von 6 Symptomen (Erythem, Ödem/ Knötchenbildung, nässende Areale/ Krustenbildung, Exkoration, Lichenifikation und Trockenheit nicht-betroffener Hautareale) durch den Arzt bestimmt (15). Vom Patienten selbst werden die Symptome Schlafstörungen und Juckreiz in den letzten 3 Tagen bzw. Nächten jeweils auf einer VAS von 0 (keine Symptome) bis 10 (schwerste Symptome) erhoben. Aus den 3 Komponenten des SCORAD wird ein Gesamtscore errechnet. Der SCORAD kann Werte zwischen 0 und 103 annehmen, wobei ein höherer Wert einem höheren Schweregrad entspricht (14, 16).

Die Operationalisierung des SCORAD erfolgte anhand der Anzahl der Patienten, die von Studienbeginn bis zu Woche 26 jeweils eine Verbesserung des SCORAD-Scores um 75 % (SCORAD-75) bzw. 90 % (SCORAD-90) erreichten. Darüber hinaus wurde die prozentuale Änderung des SCORAD vom Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu

Woche 26 analysiert und grafisch dargestellt (siehe Abschnitt 4.2.5.4). In der Studie JADE DARE wurde der SCORAD zu Woche 2, 4, 8, 12, 16, 20 und 26 erfasst.

Im vorliegenden Modul 4 A wurde der Endpunkt SCORAD für die Studie JADE DARE in den Hauptanalysen somit operationalisiert als:

- Anteil der Patienten, die eine 75 %ige Verbesserung (Reduktion) gegenüber dem *Baseline*-Wert erreichen (SCORAD-75)
- Anteil der Patienten, die eine 90 %ige Verbesserung (Reduktion) gegenüber dem *Baseline*-Wert erreichen (SCORAD-90)

Patientenberichtete Symptomatik (POEM)

Patientenrelevanz

Der POEM ist ein patientenberichtetes und validiertes Instrument zur Beurteilung der Beeinträchtigungen durch die AD. Anhand des POEM wird die Häufigkeit der Symptomatik vom Patienten selbst berichtet. Der Endpunkt ist daher als patientenrelevant anzusehen. Die Patientenrelevanz des POEM wurde zudem bereits vom G-BA im Anwendungsgebiet akzeptiert (10, 11).

Validität

In einem systematischen Review aus dem Jahr 2017 wurden u. a. die Messeigenschaften des POEM untersucht. Beim POEM wurde eine begrenzte Evidenz für eine gute interne Konsistenz, moderate Evidenz für eine gute Konstruktvalidität und für eine gute Responsivität festgestellt. Aufgrund schlechter methodischer Studienqualität bleibt die Test-Retest-Reliabilität unbekannt. Die gute Inhaltsvalidität wird durch moderate Evidenz unterstützt (17). In einer bevölkerungsbezogenen Querschnittsstudie mit 2893 Erwachsenen, darunter 602 Erwachsene, die die *United Kingdom*-Diagnosekriterien für AD erfüllten, wurde der Schweregrad der AD mit dem POEM bewertet. Die Ergebnisse zeigten, dass der POEM eine gute interne Konsistenz (Cronbachs Alpha = 0,86), gute konvergente Validität und gute Kriteriumsvalidität hatte trotz der Berücksichtigung von Bodeneffekten (18). In der Aktualisierung „Systemtherapie bei Neurodermitis“ zur deutschen S2k-Leitlinie Neurodermitis wird der POEM als validiertes Messinstrument zur Beurteilung der Beeinträchtigungen durch eine AD empfohlen (6).

Operationalisierung

Die für die AD relevante Häufigkeit des Auftretens 7 verschiedener Symptome (Juckreiz, Schlafstörungen, Blutungen der Haut, nässende Haut, rissige Haut, schuppige Haut, trockene/raue Haut) wird in dem Fragebogen für die vorangegangenen Woche erfasst. Die Häufigkeit wird wie folgt mit einem *Score* bewertet: 0 = an keinem Tag, 1 = an 1-2 Tagen, 2 = an 3-4 Tagen, 3 = an 5-6 Tagen, 4 = jeden Tag. Der Gesamtscore kann Werte zwischen 0 und 28 annehmen, wobei ein höherer Wert einem häufigeren Auftreten der Symptome entspricht (16). Die Einzelitems (Fragen) können Werte von 0-4 annehmen. Ein Gesamtscore im POEM von 0-2 entspricht einem normalen bis fast normalem Hautbild und wird daher als patientenrelevantes Therapieziel erachtet. Eine Veränderung um $\geq 15\%$ der Skalenspannweite

entspricht im POEM-Gesamtscore einer Veränderung um ≥ 5 Punkte sowie in den POEM-Einzelitems einer Veränderung um ≥ 1 Punkt (siehe Abschnitt 4.2.5.4) und gilt somit als relevante Verbesserung (7). Darüber hinaus wurde die Änderung des POEM-Gesamtscores und der jeweiligen POEM-Einzelitems von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (siehe Abschnitt 4.2.5.4). In der Studie JADE DARE wurde der POEM zu Woche 12, 16 und 26 erfasst.

Im vorliegenden Modul 4 A wurde der Endpunkt POEM für die Studie JADE DARE in den Hauptanalysen somit operationalisiert als:

- Anteil der Patienten mit POEM-Gesamtscore = 0-2
- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 5 Punkte im POEM-Gesamtscore gegenüber dem *Baseline*-Wert erreichen

Juckreiz (Peak Pruritus NRS, SCORAD VAS für Juckreiz, Einzelitem des POEM)

Patientenrelevanz

Juckreiz zählt zu dem häufigsten, am meisten belastenden und einschränkenden Symptom bei Hauterkrankungen. Die Intensität des Juckreizes ist mit Beeinträchtigungen in vielen Lebensbereichen, wie reduzierter Schlafqualität und Arbeitsproduktivität, sowie Depressionen und Suizidalität verbunden (19). Juckreiz wird aufgrund der starken Einschränkung der Lebensqualität der Patienten als patientenrelevanter Endpunkt betrachtet (10, 11).

Validität

Das für die AD relevante Symptom Juckreiz wurde mittels der *Peak Pruritus NRS*, der *SCORAD VAS* für Juckreiz und des Einzelitems des POEM für Juckreiz erfasst, welche vom Patienten selbst berichtet wurden. In einer *Mixed-Methods*-Studie mit Erwachsenen (≥ 18 Jahre) und Jugendlichen (12-17 Jahre) mit klinisch bestätigter AD wurde die Validität des *Peak Pruritus NRS* untersucht. Die *Peak Pruritus NRS* erwies sich dabei als valides, zuverlässiges und sensitives Instrument zur Erfassung von Juckreiz bei AD (20). Weitere Untersuchungen zeigten zudem eine sehr gute Test-Retest-Reliabilität, Konstruktvalidität und Änderungssensitivität sowie eine genaue Erfassung der Intensität des schlimmsten Juckreizes (21).

Die Validität der *SCORAD VAS* und des POEM wurden bereits zuvor in diesem Abschnitt (Abschnitt 4.2.5.2) dargelegt.

Operationalisierung

Die *Peak Pruritus NRS* ist eine Skala, auf der der Patient die maximale Stärke des Juckreizes innerhalb der letzten 24 Stunden bewertet, wobei ein Wert von 0 keinem Juckreiz und ein Wert von 10 dem schlimmsten vorstellbaren Juckreiz entspricht (21). Eine Verbesserung um ≥ 4 Punkte gilt als relevante Veränderung (21). Darüber hinaus wurde die Änderung der *Peak Pruritus NRS* von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (siehe Abschnitt 4.2.5.4). In der Studie JADE DARE wurde der *Peak Pruritus NRS* zu Woche 2, 4, 8, 12, 16, 20 und 26 erfasst.

Auf der SCORAD VAS für Juckreiz beurteilen die Patienten den Juckreiz in den letzten 3 Tagen bzw. Nächten auf einer Skala von 0-10, wobei ein Wert von 0 keinem Juckreiz und ein Wert von 10 dem schlimmsten vorstellbaren Juckreiz entspricht (15). Eine Veränderung um $\geq 15\%$ der Skalenspannweite der SCORAD VAS für Juckreiz entspricht einer Veränderung um ≥ 2 Punkte und gilt somit als relevante Verbesserung (7). Darüber hinaus wurde die Änderung der SCORAD VAS von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (siehe Abschnitt 4.2.5.4). In der Studie JADE DARE wurde die SCORAD VAS zu Woche 2, 4, 8, 12, 16, 20 und 26 erfasst.

Im ersten Item (Frage) des POEM („An wie vielen Tagen in der letzten Woche juckte Ihre Haut aufgrund Ihres Ekzems?“) berichtet der Patient die Häufigkeit des Auftretens von Juckreiz auf einer 5-Punkte-Skala (0 = an keinem Tag, 1 = an 1-2 Tagen, 2 = an 3-4 Tagen, 3 = an 5-6 Tagen, 4 = an jedem Tag). Das Item kann einen Wert von 0-4 annehmen (16). Eine Veränderung um $\geq 15\%$ der Skalenspannweite des POEM-Einzelitems für Juckreiz entspricht einer Veränderung um ≥ 1 Punkt und gilt somit als relevante Verbesserung (7). Die Auswertung des POEM-Einzelitems für Juckreiz wurde als Sensitivitätsanalyse betrachtet und im Abschnitt 4.2.5.4 beschrieben.

Im vorliegenden Modul 4 A wurde der Endpunkt Juckreiz anhand der *Peak Pruritus* NRS und SCORAD VAS für die Studie JADE DARE in den Hauptanalysen somit operationalisiert als:

Peak Pruritus NRS

- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 4 Punkte in der *Peak Pruritus* NRS gegenüber dem *Baseline*-Wert erreichen

SCORAD VAS für Juckreiz

- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 2 Punkte in der SCORAD VAS für Juckreiz gegenüber dem *Baseline*-Wert erreichen

Schlafstörungen (MOS Sleep Scale, SCORAD VAS für Schlafstörung, Einzelitem des POEM)

Patientenrelevanz

Schlafstörungen sind bei Patienten mit dermatologischen Erkrankungen weit verbreitet und stellen eine erhebliche Belastung dar, die die Lebensqualität stark einschränkt (22–24). Kinder mit AD weisen eine verminderte Schlafqualität auf, mit verminderter Schlafdauer, geringer Schlaffeffizienz und vermehrtem Aufwachen, und es ist klinisch bekannt, dass sich die Krankheit in der Nacht zirkadian verschlimmert (25–27). Folgen einer geringeren Schlafqualität sind Müdigkeit, die die motorischen und mentalen Funktionen sowie die psychosoziale Leistungsfähigkeit (Stresserleben) einschränkt (28). Das für die AD relevante Symptom Schlafstörungen wurde mittels der *Medical Outcomes Study (MOS) Sleep Scale*, der SCORAD VAS für Schlafstörungen und des Einzelitems des POEM für Schlafstörungen erfasst, welche vom Patienten selbst berichtet wurden und demnach als patientenrelevant anzusehen sind.

Validität

Die MOS *Sleep Scale* ist ein validiertes Instrument zur Beurteilung der Schlafqualität. Sie hat gute psychometrische Eigenschaften (29, 30), einschließlich Reliabilität und Konstruktvalidität (31).

Die Validität der SCORAD VAS und des POEM wurden bereits zuvor in diesem Abschnitt (Abschnitt 4.2.5.2) dargelegt.

Operationalisierung

Die MOS *Sleep Scale* ist ein vom Patienten selbst dokumentiertes und validiertes Instrument zur Beurteilung der Schlafqualität in den letzten 4 Wochen. Sie umfasst insgesamt 12 Items in den 6 Domänen Schlafstörungen, Schlafadäquanz, Schlafquantität, Somnolenz, Schnarchen, Kurzatmigkeit oder Kopfschmerzen. Die Items können zu 2 Indices zusammengefasst werden. Der MOS *Sleep Scale Score Index-1* umfasst 6 Items (Schlafprobleme ohne Schlafquantität) und der MOS *Sleep Scale Score Index-2* umfasst 9 Items (Gesamtübersicht der Schlafprobleme). Der Gesamtscore der MOS *Sleep Scale Score Indices* kann jeweils Werte von 0-100 annehmen (29). Eine Veränderung um $\geq 15\%$ der jeweiligen Skalenspannweite der MOS *Sleep Scale Score Indices* entspricht einer Veränderung um ≥ 15 Punkte und gilt somit als relevante Verbesserung (7). Darüber hinaus wurde die jeweilige Änderung der MOS *Sleep Scale Score Indices* von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (Abschnitt 4.2.5.2). In der Studie JADE DARE wurde die MOS *Sleep Scale* zu Woche 12, 16 und 26 erfasst.

Auf der SCORAD VAS für Schlafstörungen beurteilen die Patienten die Schlafstörungen der letzten 3 Tage bzw. Nächte auf einer Skala von 0-10, wobei ein Wert von 0 keinen Schlafstörungen und ein Wert von 10 den schlimmsten vorstellbaren Schlafstörungen entspricht (15). Eine Veränderung um $\geq 15\%$ der Skalenspannweite der SCORAD VAS für Schlafstörungen entspricht einer Veränderung um ≥ 2 Punkte und gilt somit als relevante Verbesserung (7). Darüber hinaus wurde die Änderung der SCORAD VAS für Schlafstörungen von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (Abschnitt 4.2.5.2). In der Studie JADE DARE wurde die SCORAD VAS zu Woche 2, 4, 8, 12, 16, 20 und 26 erfasst.

Im zweiten Item (Frage) des POEM („In wie vielen Nächten in der letzten Woche war Ihre Nachtruhe aufgrund Ihres Ekzems gestört?“) berichtet der Patient die Häufigkeit des Auftretens von Schlafstörungen auf einer 5-Punkte-Skala (0 = in keiner Nacht, 1 = in 1-2 Nächten, 2 = in 3-4 Nächten, 3 = in 5-6 Nächten, 4 = jede Nacht). Das Item kann einen Wert von 0-4 annehmen (16). Eine Veränderung um $\geq 15\%$ der Skalenspannweite des POEM-Einzelitems für Schlafstörungen entspricht einer Veränderung um ≥ 1 Punkt und gilt somit als relevante Verbesserung (7). Die Auswertung des POEM-Einzelitems für Schlafstörungen wurde als Sensitivitätsanalyse betrachtet und im Abschnitt 4.2.5.4 beschrieben.

Im vorliegenden Modul 4 A wurde somit der Endpunkt Schlafstörungen anhand der MOS *Sleep Scale* und SCORAD VAS für Schlafstörungen für die Studie JADE DARE in den Hauptanalysen operationalisiert als:

MOS *Sleep Scale*

- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 15 Punkte in den jeweiligen MOS *Sleep Scale Score* Indices gegenüber dem *Baseline*-Wert erreichen

SCORAD VAS für Schlafstörung

- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 2 Punkte in der SCORAD VAS für Schlafstörung gegenüber dem *Baseline*-Wert erreichen

Hautschmerz (*Skin Pain NRS*)

Patientenrelevanz

AD geht einher mit Symptomen wie Juckreiz, Hautentzündungen, Barrierestörungen und Kratzen, die alle mit Hautschmerzen assoziiert sein können. Schmerz wird von Patienten mit AD als stark belastend und einschränkend empfunden (32, 33). Der Endpunkt Hautschmerz wurde mittels der *Skin Pain NRS* erfasst. Die *Skin Pain NRS* wurde vom Patienten selbst berichtet. Aufgrund der starken Einschränkung der Lebensqualität der Patienten wird der Hautschmerz daher als patientenrelevanter Endpunkt betrachtet.

Validität

Silverberg et al. untersuchten die Messeigenschaften der NRS für Schmerzen bei AD. Die *Skin Pain NRS* wies eine ausreichende Validität, Reliabilität, Ansprechbarkeit und Interpretierbarkeit bei Erwachsenen mit AD auf und gilt als grundsätzlich geeignet für den Einsatz in klinischen Studien und in der Praxis (34).

Operationalisierung

Die *Skin Pain NRS* fragt den schlimmsten Hautschmerz in den letzten 24 Stunden auf einer 11-Punkte-Skala von 0 (kein Schmerz) bis 10 (schlimmster vorstellbarer Schmerz) ab (35). Eine Verbesserung um ≥ 4 Punkte gilt als relevante Veränderung (35). Darüber hinaus wurde die Änderung der *Skin Pain NRS* von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (siehe Abschnitt 4.2.5.4). In der Studie JADE DARE wurde der *Skin Pain NRS* zu Woche 2, 12, 16, 20 und 26 erfasst.

Im vorliegenden Modul 4 A wurde somit der Endpunkt Hautschmerz anhand der *Skin Pain NRS* für die Studie JADE DARE in den Hauptanalysen operationalisiert als:

- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 4 Punkte in der *Skin Pain NRS* gegenüber dem *Baseline*-Wert erreichen

Ängste und Depressionen (HADS)

Patientenrelevanz

Die *Hospital Anxiety and Depression Scale* (HADS) ist ein validiertes und patientenberichtetes Instrument zur Beurteilung von Angstzuständen und Depression während der vergangenen Woche und berücksichtigt die psychosozialen Belastungen durch die AD. In einer systematischen Übersichtsarbeit und Meta-Analyse wurde festgestellt, dass die Prävalenz jeglicher Depression bei Personen mit AD höher war als bei Personen ohne AD (20,1 % versus 14,8 %). AD war mit signifikant höheren Werten auf Depressionsskalen, elterlicher Depression, Einnahme von Antidepressiva und Suizidalität assoziiert (36). AD steht neben der Assoziation mit Depression auch mit signifikant erhöhten Angststörungen in Zusammenhang (37). Endpunkte zu Angstzuständen und Depression sind patientenrelevant, da Patienten mit AD eine erhöhte Vulnerabilität für psychische Erkrankungen haben.

Validität

Silverberg et al. untersuchten in einer Querschnittsstudie mit 602 Erwachsenen die Messeigenschaften der HADS zur Beurteilung der psychischen Auswirkungen der AD bei Erwachsenen. Die HADS zeigte gute Konstruktvalidität, interne Konsistenz (Cronbachs Alpha = 0,89) und insgesamt keine Decken- und Bodeneffekte (38).

Operationalisierung

Die HADS setzt sich aus 2 Subskalen (Angst und Depression) mit jeweils 7 Items zusammen, die auf einer 4-Punkte-Skala (0-3) bewertet werden, wobei ein Wert von 3 die höchste Angst- oder Depressionsstufe bezeichnet. Die HADS-Subskalen können jeweils Werte von 0-21 annehmen, wobei ein höherer Wert einer größeren psychosozialen Belastung entspricht (39, 40). Ab einem Wert von ≥ 8 Punkten liegt eine klinisch relevante Symptomatik in der jeweiligen Subskala vor (41, 42). Die Änderung der jeweiligen HADS-Subscores von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 wurde analysiert und grafisch dargestellt. In der Studie JADE DARE wurde die HADS zu Woche 12, 16 und 26 erfasst.

Im vorliegenden Modul 4 A wurde somit der Endpunkt Ängste und Depressionen anhand der HADS für die Studie JADE DARE in der Hauptanalyse operationalisiert als:

- Mittlere Veränderung im HADS-Subscore Angst bzw. Depression gegenüber dem Baseline-Wert anhand eines *Mixed effect Model Repeat Measurement* (MMRM)

Gesundheitszustand (EQ-5D VAS)

Patientenrelevanz

Die VAS des Fragebogens *European Quality of Live Group 5 Dimensions* (EQ-5D) ist ein etabliertes, validiertes und patientenberichtetes Instrument zur Erhebung des Gesundheitszustands. Anhand der EQ-5D VAS wird der Gesundheitszustand vom Patienten selbst berichtet. Der Endpunkt ist daher als patientenrelevant anzusehen. Die Patientenrelevanz des EQ-5D VAS wurde zudem bereits vom G-BA im Anwendungsgebiet akzeptiert (10, 11).

Validität

In einer systematischen Literaturübersicht wurde die Reliabilität, Validität und Responsivität des EQ-5D bei Patienten mit Hauterkrankungen untersucht. Insgesamt wurde die Validität und Responsivität des EQ-5D bei Patienten mit Hauterkrankungen als gut befunden (43).

Operationalisierung

Den für die AD relevanten Gesundheitszustand schätzt der Patient auf einer Skala von 0-100 ein, wobei 0 den denkbar schlechtesten Gesundheitszustand und 100 den denkbar besten Gesundheitszustand widerspiegelt (44). Eine Veränderung um $\geq 15\%$ der Skalenspannweite der EQ-5D VAS entspricht einer Veränderung um ≥ 15 Punkte und gilt somit als relevante Verbesserung (7). Darüber hinaus wurde die Änderung der EQ-5D VAS von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (Abschnitt 4.2.5.2). In der Studie JADE DARE wurde die EQ-5D VAS zu Woche 12, 16 und 26 erfasst.

Im vorliegenden Modul 4 A wurde somit der Endpunkt Gesundheitszustand anhand der EQ-5D VAS für die Studie JADE DARE operationalisiert als:

- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 15 Punkte in der EQ-5D VAS gegenüber dem *Baseline*-Wert erreichen

Endpunkt betroffene BSA (supportiver Endpunkt)

Operationalisierung

Bei der Beurteilung der von der AD betroffenen BSA werden die 4 Körperbereiche Kopf und Hals, Rumpf, obere Extremitäten und untere Extremitäten betrachtet. Die Bewertung der BSA erfolgt auf einer Skala von 0-100 % (45). Für den Endpunkt betroffene BSA wurde die prozentuale Änderung von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (siehe Abschnitt 4.2.5.4). In der Studie JADE DARE wurde die betroffene BSA zu Woche 2, 4, 8, 12, 16, 20 und 26 erfasst.

Die Analyse der betroffenen BSA wird im Rahmen des Nutzendossiers anhand der präspezifizierten Operationalisierung in der Studie JADE DARE supportiv dargestellt:

- Mittlere prozentuale Änderung der betroffenen BSA gegenüber dem *Baseline*-Wert anhand eines MMRM

Gesundheitsbezogene Lebensqualität

Krankheitsspezifische Lebensqualität (DLQI)

Patientenrelevanz

Beim DLQI handelt es sich um einen validierten und patientenberichteten Fragebogen zur Ermittlung der krankheitsspezifischen gesundheitsbezogenen Lebensqualität bei erwachsenen Patienten mit dermatologischen Erkrankungen. Der DLQI wird vom Patienten selbst berichtet und wurde vom G-BA im Anwendungsgebiet als patientenrelevant betrachtet (4, 10, 11).

Validität

In einer prospektiven Studie, die in einer dermatologischen Praxis mit 340 Erwachsenen durchgeführt wurde, wurde die gesundheitsbezogene Lebensqualität anhand des DLQI bewertet. Der DLQI zeigte gute Inhaltsvalidität, gute interne Konsistenz (Cronbachs Alpha = 0,89) und Konstruktvalidität sowie Responsivität bei der Beurteilung von AD bei Erwachsenen und war für den Einsatz in klinischen Studien und der Praxis geeignet. Es gab keine Boden- oder Deckeneffekte für die Gesamtscores (46). In der Aktualisierung „Systemtherapie bei Neurodermitis“ zur deutschen S2k-Leitlinie Neurodermitis wird der DLQI als validiertes Messinstrument zur Ermittlung der Lebensqualität bei Patienten mit AD empfohlen (6).

Operationalisierung

Der DLQI umfasst 10 Items (Fragen), die sich auf Beeinträchtigungen der befragten Personen in den vergangenen 7 Tagen beziehen und vom Patienten selbst beantwortet werden. Die Items werden den folgenden 6 Subskalen zugeordnet: Symptome und Gefühle, tägliche Aktivitäten, Freizeit, Arbeit/ Schule, persönliche Beziehungen und Auswirkungen der Behandlung auf den Alltag. Dabei werden die Items auf einer 4-Punkte-Skala beantwortet (0: gar nicht, 1: etwas, 2: stark, 3: sehr stark). Der DLQI-Gesamtscore berechnet sich durch die Aufsummierung der in jeder Domäne erzielten Punkte und kann Werte von 0-30 annehmen, wobei ein höherer Wert einer größeren Beeinträchtigung der Lebensqualität entspricht (47–50). Ein Gesamtscore im DLQI von 0 oder 1 entspricht keiner Beeinträchtigung der Lebensqualität und wird daher als patientenrelevantes Therapieziel erachtet (10, 11). Eine Veränderung um $\geq 15\%$ der Skalenspannweite des DLQI entspricht einer Veränderung um ≥ 5 Punkte und gilt somit als relevante Verbesserung (7). Darüber hinaus wurde die Änderung des DLQI von Studienbeginn zu allen Beobachtungszeitpunkten bis hin zu Woche 26 analysiert und grafisch dargestellt (siehe Abschnitt 4.2.5.4). In der Studie JADE DARE wurde der DLQI zu Woche 2, 12, 16, 20 und 26 erfasst.

Im vorliegenden Modul 4 A wurde der Endpunkt krankheitsspezifische Lebensqualität anhand des DLQI für die Studie JADE DARE in den Hauptanalysen somit operationalisiert als:

- Anteil der Patienten mit DLQI-Gesamtscore = 0 oder 1
- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 5 Punkte gegenüber dem *Baseline*-Wert erreichen

Verträglichkeit

Im Rahmen der Zulassung eines Arzneimittels muss dessen Unbedenklichkeit und Verträglichkeit anhand der Dokumentation von unerwünschten Ereignissen (UE) nachgewiesen werden.

Unerwünschte Ereignisse

Patientenrelevanz

Eine detaillierte Erfassung der UE erlaubt die Erstellung des Toxizitätsprofils einer Medikation und ist daher für eine Nutzen-Risiko-Abwägung unerlässlich. Der Nachweis der Verringerung von Nebenwirkungen zählt zu den Aspekten einer Verbesserung der therapiebedingten Morbidität sowie der Lebensqualität und ist ein durch das IQWiG anerkannter patientenrelevanter Endpunkt (7). Ein Wirkstoff kann im Rahmen einer Therapie neben der erwünschten Wirkung auf die behandelte Erkrankung auch unerwünschte Wirkungen hervorrufen, die die Patienten neben der eigentlichen Erkrankung zusätzlich physisch und psychisch belasten können und die der Patient unmittelbar erlebt und berichtet. Dies kann unter anderem zu einer reduzierten Therapieadhärenz führen. Die Patientenrelevanz ist als allgemein anerkannt in der Nutzenbewertung zu betrachten. Die VerfO des G-BA § 3 Absatz 1, 5. Kapitel nennt die Verringerung der Nebenwirkungen explizit als patientenrelevanten Effekt (4).

Validität

Die Erfassung des Sicherheitsprofil ist bei interventionellen, klinischen Studien zwingend erforderlich und wird nach verbindlichen internationalen Standards (gute klinische Praxis, *Good Clinical Practice*, [GCP], *International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use* [ICH] E6) berichtet (51). Hierfür werden alle vom Patienten spontan berichteten UE in freier Textform („verbatim“) erfasst. Anschließend gewährleistet eine Kodierung in einheitlichem und international verbindlichem Vokabular (*Medical Dictionary for Regulatory Activities*, MedDRA) die wissenschaftliche Auswertung. Die UE wurden in der Studie JADE DARE standardmäßig nach der GCP erhoben sowie gemäß Version 24.0 des MedDRA kodiert.

Operationalisierung

Die Erhebung der UE erfolgte in der Studie JADE DARE ab Studienbeginn und für Patienten, die nicht in die Erweiterungsstudie eingeschlossen wurden, bis hin zu 28 Tagen nach der letzten Dosis der Studienmedikation (*Safety Follow-up*). Zur Evaluierung der für die Bewertung relevanten Aspekte wurden für die Studie JADE DARE die folgenden Endpunkte dargestellt:

- UE jeglichen Schweregrades
 - Gesamtrate
 - Häufige UE nach SOC und bevorzugter Bezeichnung (*Preferred Term*, PT), die bei mindestens 10 % der Patienten oder mindestens 10 Patienten und 1 % der Patienten in mindestens einem Studienarm aufgetreten sind.
- Schwere UE
 - Gesamtrate
 - Häufige schwere UE nach SOC und PT, die bei mindestens 5 % der Patienten oder mindestens 10 Patienten und 1 % der Patienten in mindestens einem Studienarm aufgetreten sind.

- SUE
 - Gesamtrate
 - Häufige SUE nach SOC und PT, die bei mindestens 5 % der Patienten oder mindestens 10 Patienten und 1 % der Patienten in mindestens einem Studienarm aufgetreten sind.
- UE, die zum Therapieabbruch führten
 - Gesamtrate
 - Darstellung nach SOC und PT (deskriptiv)
- Spezifische UE, jeweils als Gesamtrate und differenziert nach Schweregrad (schwer, schwerwiegend):
 - Superinfektionen (definiert als Infektionen, die > 2 Wochen mit antibiotischen, antiviralen oder fungiziden Wirkstoffen behandelt wurden)
 - Herpes zoster (definiert durch *Customized MedDRA-Queries*)
 - Konjunktivitis (definiert durch eine detaillierte Erfassung im *Case Report Form* (CRF))
 - Akne (definiert durch eine detaillierte Erfassung im CRF)
 - Follikulitis (definiert durch eine detaillierte Erfassung im CRF)

Ergebnisse zu UE jeglichen Schweregrades, schwere UE und SUE ohne Progressionsereignisse sind in Anhang 4-G dargestellt.

Die Bewertung der Intensität bzw. des Schweregrades der UE erfolgte in der Studie JADE DARE durch den Prüfarzt nach standardisierten Kriterien und war folgendermaßen definiert (52):

- Nicht schwere UE umfassten die Ereignisse, die als mild und moderat bewertet wurden:
 - Mild: Ein Ereignis, das vom Teilnehmer leicht toleriert werden konnte, minimales Unbehagen verursachte und die alltäglichen Aktivitäten nicht beeinträchtigte.
 - Moderat: Ein Ereignis, das hinreichende Beschwerden verursachte und die alltäglichen Aktivitäten beeinträchtigte.
- Schwere UE, waren Ereignisse, die die Ausübung von alltäglichen Aktivitäten verhinderten.

Auswertung der Endpunkte

Die Analyse der Wirksamkeitsendpunkte erfolgte in der Studie JADE DARE gemäß statistischem Analyseplan (SAP) basierend auf dem modifizierten *Intention-to-treat-Analysis-Set* (mITT) und die Auswertung der UE erfolgte auf Basis des *Safety-Analysis-Set* (SAS). Das

mITT beinhaltete alle randomisierten Patienten aus dem *Full-Analysis-Set* (FAS), die mindestens eine Dosis der Studienmedikation erhalten haben, und zu *Baseline* das entsprechende *Response*-Kriterium der Responderanalyse erreicht haben. Die Auswertung erfolgte gemäß der randomisierten Behandlungszuteilung. Das SAS umfasste alle Patienten, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben. Die Auswertung erfolgte gemäß der tatsächlich erhaltenen Therapie.

Die Responderanalysen der patientenberichteten Endpunkte (Morbidität und gesundheitsbezogene Lebensqualität) wurden auf Basis des mITT durchgeführt. Die Verlaufsbeobachtungen basierten auf dem FAS.

Im Ergebnisteil des vorliegenden Modul 4 A werden die in Tabelle 4-6 angegebenen Analysen für die patientenrelevanten Endpunkte dargestellt:

Tabelle 4-6: Übersicht der Analysen für die betrachteten patientenrelevanten Endpunkte

Endpunkt	Binär	Verlaufsbeobachtungen ^a
EASI	•	•
SCORAD	•	•
POEM	•	•
<i>Peak Pruritus</i> NRS	•	•
SCORAD VAS	•	•
MOS <i>Sleep Scale</i>	•	•
<i>Skin Pain</i> NRS	•	•
HADS	-	•
EQ-5D VAS	•	•
Betroffene BSA (supportiv)	-	•
DLQI	•	•
Verträglichkeit	•	-
<p>•: Durchgeführte Analyse. -: Nicht durchgeführte Analyse</p> <p>a: Die Verlaufsbeobachtungen werden für die Endpunkte EASI, SCORAD, POEM, <i>Peak Pruritus</i> NRS, SCORAD VAS, MOS <i>Sleep Scale</i>, <i>Skin Pain</i> NRS, HADS, EQ-5D VAS und DLQI als Sensitivitätsanalysen betrachtet (siehe Abschnitt 4.2.5.4).</p> <p>BSA: Körperoberfläche (<i>Body Surface Area</i>); DLQI: <i>Dermatology Life Quality Index</i>; EASI: <i>Eczema Area and Severity Index</i>; EQ-5D: <i>European Quality of Life Group 5 Dimensions</i>; HADS: <i>Hospital Anxiety and Depression Scale</i>; MOS: <i>Medical Outcomes Study</i>; NRS: <i>Numerical Rating Scale</i>; POEM: <i>Patient Oriented Eczema Measure</i>; SCORAD: <i>Scoring Atopic Dermatitis</i>; VAS: Visuelle Analogskala</p>		

Binäre Endpunkte

Für die binären Endpunkte EASI, SCORAD, POEM, *Peak Pruritus* NRS, SCORAD VAS, MOS *Sleep Scale*, *Skin Pain* NRS, EQ-5D VAS, DLQI und Verträglichkeit zeigt die tabellarische Ergebnisdarstellung die Anzahl der Studienteilnehmer, die in die Analyse eingegangen sind, sowie die Anzahl und den Anteil der Studienteilnehmer mit mindestens

einem Ereignis pro Behandlungsgruppe. Für die Responderanalysen wurde der *Estimand 1 (Composite Estimand)* angewendet. Beim *Estimand 1* wurden die beiden *intercurrent events* Therapieabbruch aus jeglichem Grund und Verwendung einer *Rescue*-Therapie als *Non-Response* betrachtet. Der Behandlungseffekt wird durch das relative Risiko (RR) mit jeweiligem 95 %-Konfidenzintervall (KI) und zweiseitigem p-Wert dargestellt (Cochran-Mantel-Haenszel-Test stratifiziert nach der Variable Krankheitsschwere zu *Baseline (Investigator's Global Assessment (IGA) = 3 versus IGA = 4)*). Zusätzlich wird das *Odds Ratio (OR)* und die Risikodifferenz (RD) mit dem jeweiligen 95 %-KI und zweiseitigem p-Wert dargestellt.

Verlaufsbeobachtungen

Für die Hauptanalysen der HADS und BSA sowie für die Sensitivitätsanalysen der Endpunkte EASI, SCORAD, POEM, *Peak Pruritus NRS*, SCORAD VAS, MOS *Sleep Scale*, *Skin Pain NRS*, EQ-5D VAS und DLQI (siehe Abschnitt 4.2.5.4) wurden MMRM-Analysen durchgeführt und tabellarisch die Anzahl der analysierten Studienteilnehmer, Mittelwert (MW) mit Standardabweichung (*Standard Deviation, SD*) und der Methode der kleinsten Quadrate (*Least-Squares, LS*)-*Mean* mit Standardfehler (*Standard Error, SE*) zu Studienbeginn (*Baseline*) und zu jedem Beobachtungszeitpunkt pro Behandlungsgruppe dargestellt. In den Analysen zur mittleren Änderung wurde der *Estimand 2 (Hypothetical Estimand)* betrachtet, wonach Werte nach dem Auftreten eines *intercurrent events* auf *missing* gesetzt und im Rahmen des MMRM berücksichtigt wurden. Der Behandlungseffekt wurde durch die *LS-Mean-Differenz* zwischen den Behandlungsgruppen mit einem 95 %-KI und p-Wert quantifiziert. Das MMRM mit unstrukturierter Kovarianz-Matrix beinhaltet gemäß SAP die festen Effekte Behandlung, Visite und die Interaktion *Visite*Behandlung* sowie den entsprechenden *Baseline*-Wert und Krankheitsschwere zu *Baseline* als Kovariaten. Um bezüglich der Signifikanzaussage Konsistenz zwischen der standardisierten Mittelwertdifferenz (SMD, Hedges'g) und der präspezifizierten Ausgangsanalyse im Studienbericht zu wahren, wurden auch bei der Berechnung der SMD die adjustierten Schätzer aus dem entsprechenden Modell verwendet. Fehlende Werte wurden für die maßgeblichen Analysen nicht ersetzt.

4.2.5.3 Meta-Analysen

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

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-

Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten nach der Knapp-Hartung-Methode mit der Paule-Mandel-Methode zur Heterogenitätsschätzung⁵ erfolgen. Im Fall von sehr wenigen Studien ist die Heterogenität nicht verlässlich schätzbar. Liegen daher weniger als 5 Studien vor, ist auch die Anwendung eines Modells mit festem Effekt oder eine qualitative Zusammenfassung in Betracht zu ziehen. Kontextabhängig können auch alternative Verfahren wie z. B. Bayes'sche Verfahren oder Methoden aus dem Bereich der generalisierten linearen Modelle in Erwägung gezogen werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁶ 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 geeigneter statistische Maße auf Vorliegen von Heterogenität^{7,5} erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik $\geq 0,05$), soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) 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.

Nicht zutreffend. Es wurden keine Meta-Analysen durchgeführt.

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

⁶ Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

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

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Sensitivitätsanalysen

Zur Überprüfung der Robustheit der Ergebnisse der Hauptanalysen der Studie JADE DARE werden im vorliegenden Modul 4 A folgende Sensitivitätsanalysen dargestellt:

- Mittlere prozentuale Änderung des EASI gegenüber dem *Baseline*-Wert anhand eines MMRM
- Mittlere prozentuale Änderung des SCORAD gegenüber dem *Baseline*-Wert anhand eines MMRM
- Mittlere Veränderung des POEM-Gesamtscores gegenüber dem *Baseline*-Wert anhand eines MMRM
- Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 1 Punkt im POEM-Einzelitem Juckreiz, Schlafstörungen, Blutungen der Haut, nässende Haut, rissige Haut, schuppige Haut und trockene/ raue Haut gegenüber dem *Baseline*-Wert erreichen
- Mittlere Veränderung der *Peak Pruritus* NRS gegenüber dem *Baseline*-Wert anhand eines MMRM
- Mittlere Veränderung der SCORAD VAS für Juckreiz gegenüber dem *Baseline*-Wert anhand eines MMRM
- Mittlere Veränderung der jeweiligen MOS *Sleep Scale Score* Indices gegenüber dem *Baseline*-Wert anhand eines MMRM
- Mittlere Veränderung der SCORAD VAS für Schlafstörungen gegenüber dem *Baseline*-Wert anhand eines MMRM
- Mittlere Veränderung der *Skin Pain* NRS gegenüber dem *Baseline*-Wert anhand eines MMRM

- Mittlere Veränderung der EQ-5D VAS gegenüber dem *Baseline*-Wert anhand eines MMRM
- Mittlere Veränderung des DLQI gegenüber dem *Baseline*-Wert anhand eines MMRM

Die folgenden zusätzlichen Sensitivitätsanalysen wurden für alle binären Endpunkte durchgeführt, um sicherzustellen, dass fehlende Werte keinen Einfluss auf die Ergebnisse haben:

- Sensitivitätsanalyse 1: Beobachtete Fälle zu Woche 26 bei gleichzeitiger Betrachtung der Werte nach *Rescue*-Therapie als fehlende Werte
- Sensitivitätsanalyse 2: Imputation von Patienten mit Therapieabbruch als *Non-Response* bei gleichzeitiger Berücksichtigung der Werte nach *Rescue*-Therapie
- Sensitivitätsanalyse 3: Multiple Imputation für Patienten mit Therapieabbruch bei gleichzeitiger Berücksichtigung der Werte nach *Rescue*-Therapie.

Die Verfahren zum Umgang mit fehlenden Werten waren im Studienprotokoll bzw. SAP präspezifiziert.

Für die Analyse der Gesamtraten der UE blieben in einer Sensitivitätsanalyse solche Ereignisse unberücksichtigt, die als Symptome bzw. Verschlechterung der Symptome (Progression) der Grunderkrankung zu sehen waren.

Die beschriebenen Sensitivitätsanalysen befinden sich in Anhang 4-G.

Supportive Analysen

Des Weiteren werden supportive Analysen dargestellt, mit dem Zweck die Ergebnisse zum Nutzen und Zusatznutzen von Abrocitinib zusätzlich zu stärken und die Konsistenz der Ergebnisse über verschiedene Analysen hinweg zu überprüfen. Hierbei handelt es sich um weitere Analysen, welche die Ableitung des Nutzens und Zusatznutzens von Abrocitinib unterstützen.

Die folgenden Analysen werden supportiv dargestellt:

- Mittlere prozentuale Änderung der betroffenen BSA gegenüber dem *Baseline*-Wert anhand eines MMRM
- Anteil der Patienten mit einem IGA-Gesamtscore = 0 oder 1 und einer Verbesserung (Reduktion) um ≥ 2 Punkte gegenüber dem *Baseline*-Wert zu Woche 26 (siehe Anhang 4-G)

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die

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

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

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen (zu ergänzenden Kriterien zur Darstellung siehe Abschnitt 4.3.1.3.2).

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

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

Um die Abhängigkeit der Ergebnisse von bestimmten Patientencharakteristika und weiteren Faktoren zu untersuchen, wurden Subgruppenanalysen durchgeführt. Im Einzelnen handelte es sich um folgende präspezifizierte Merkmale:

- Alter in Jahren (Kategorie 1: < 40 versus ≥ 40 ; Kategorie 2: < 65 versus ≥ 65)
- Geschlecht (Männlich versus Weiblich)
- Abstammung (kaukasisch versus schwarz oder afroamerikanisch versus asiatisch versus andere)

- Region der Studieneinschreibung (*United States of America (USA)*/ Kanada/ Australien versus Europa versus Asien)
- Gewicht in kg (< 70 versus 70-100 versus > 100 kg)
- Krankheitsdauer in Jahren (< 26 versus \geq 26)
- Krankheitsschwere zu *Baseline* (IGA = 3 versus IGA = 4)
- EASI-Score zu *Baseline* (16-25 versus > 25)
- Betroffene BSA (%) zu *Baseline* (10-30 % versus > 30-50 % versus > 50 %)
- Vorbehandlung mit Ciclosporin (ja versus nein)
- Vorbehandlung mit AD-Medikation (nur topische Mittel versus systemische Mittel)

Die Merkmale Alter, Geschlecht und Schweregrad sind laut VerFO des G-BA gefordert (4). Zusätzlich wurden potenzielle geografische Unterschiede durch den Faktor Region untersucht. Eine detailliertere Unterteilung der Studienpopulation auf Länder- oder Zentrumsebene wurde darüber hinaus nicht als sinnvoll erachtet, da die Anzahl der Patienten in den jeweiligen Subgruppen zu gering für eine Interpretation von Behandlungseffekten war. Die Auswahl der Subgruppenmerkmale richtete sich nach der Präspezifikation gemäß SAP der Studie JADE DARE. Zudem wurden Subgruppenanalysen mit dem Merkmal *Peak Pruritus* NRS (4-6 versus \geq 7) *posthoc* durchgeführt, um eventuelle Effektmodifikationen durch die Schwere des Juckreizes zu *Baseline* zu untersuchen.

Zur Berechnung des Interaktionstests wurde für alle Endpunkte das gleiche Modell verwendet wie zur Durchführung der Analyse auf Gesamtebene (Abschnitt 4.2.5.2).

Für binäre Endpunkte wurde die Subgruppenvariable als Zwei-Wege-Interaktion mit der Behandlung sowie als Haupteffekt in das Modell aufgenommen (subgruppenbildender Faktor + Behandlung \times subgruppenbildender Faktor). Der Interaktionstest war ein Wald-Test.

Es werden detaillierte Ergebnisse zu Subgruppen präsentiert, wenn es mindestens Belege auf eine Interaktion (zweiseitiger p-Wert des Interaktionstests < 0,05) zwischen Behandlungseffekt und Subgruppenmerkmal gab. Subgruppenergebnisse, bei denen der Interaktionstest nicht auf heterogene Ergebnisse hinwies, sind in Anhang 4-G dargestellt.

Gemäß IQWiG-Methodik wurden Subgruppenanalysen nur dann durchgeführt, falls jede Subgruppe mindestens 10 Personen umfasst und bei binären Daten zusätzlich mindestens 10 Ereignisse in einer der Subgruppen aufgetreten sind (7).

Hinsichtlich der Analysen der UE auf Ebene der SOC und PT wurden Subgruppenanalysen ausschließlich dann durchgeführt, falls das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant war.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Es besteht dabei internationaler Konsens, dass Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator (häufig als nicht adjustierte indirekte Vergleiche bezeichnet) regelhaft keine valide Analysemethoden darstellen⁸. Eine Ausnahme kann das Vorliegen von dramatischen Effekten sein. An Stelle von nicht adjustierten indirekten Vergleichen sollen je nach Datenlage einfache adjustierte indirekte Vergleiche⁹ oder komplexere Netzwerk-Meta-Analysen (auch als „Mixed Treatment Comparison [MTC] Meta-Analysen“ oder „Multiple Treatment Meta-Analysen“ bezeichnet) für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche berechnet werden. Aktuelle Verfahren wurden beispielsweise von Lu und Ades (2004)¹⁰ und Rücker (2012)¹¹ vorgestellt.

Alle Verfahren für indirekte Vergleiche gehen im Prinzip von den gleichen zentralen Annahmen aus. Hierbei handelt es sich um die Annahmen der Ähnlichkeit der eingeschlossenen Studien, der Homogenität der paarweisen Vergleiche und der Konsistenz zwischen direkter und indirekter Evidenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹².

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

⁸ Bender R, Schwenke C, Schmoor C, Hauschke D. Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR [online]. [Zugriff: 31.10.2016]. URL: http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf.

⁹ Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50(6): 683-691.

¹⁰ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

¹¹ Rücker G. Network meta-analysis, electrical networks and graph theory. *Res Synth Methods* 2012; 3(4): 312-324.

¹² Schöttker B, Lühmann D, Boukhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

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

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

- *Benennung aller potentiellen Brückenkomparatoren ersten Grades und ggf. Begründung für die Auswahl.*
- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayes'schen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten, die Art der Untersuchung der Konvergenz der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Ähnlichkeit der eingeschlossenen Studien.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenzannahme im Netzwerk.*
- *Bilden Sie den Code des Computerprogramms inklusive der einzulesenden Daten in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Nicht zutreffend.

¹³ Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

¹⁴ Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study *BMJ* 2011; 343 :d4909

¹⁵ Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. *Res Synth Methods* 2013; 4(4): 291-323.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

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

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

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

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

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

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer ggf. Datenschnitt	Therapiearme
B7451006 (NCT02780167)	nein	ja	abgeschlossen	Studienbeginn: April 2016 Studienende: April 2017	Abrocitinib 200 mg, Abrocitinib 100 mg, Abrocitinib 30 mg, Abrocitinib 10 mg, Placebo
JADE MONO-1 (NCT03349060)	ja	ja	abgeschlossen	Studienbeginn: Dezember 2017 Studienende: März 2019	Abrocitinib 200 mg, Abrocitinib 100 mg, Placebo
JADE MONO-2 (NCT03575871)	ja	ja	abgeschlossen	Studienbeginn: Juni 2018 Studienende: August 2019	Abrocitinib 200 mg, Abrocitinib 100 mg, Placebo
JADE REGIMEN (NCT03627767)	ja	ja	abgeschlossen	Studienbeginn: Juni 2018 Studienende: Oktober 2020	Abrocitinib 200 mg, Abrocitinib 100 mg, Placebo
JADE MOA (NCT03915496)	nein	ja	laufend	Studienbeginn: Juni 2020 Studienende: März 2022	Abrocitinib 200 mg, Abrocitinib 100 mg, Placebo
JADE COMPARE (NCT03720470)	ja	ja	abgeschlossen	Studienbeginn: Oktober 2018 Studienende: März 2020	Abrocitinib 200 mg, Abrocitinib 100 mg, Dupilumab
JADE EXTEND (NCT03422822)	nein	ja	laufend	Studienbeginn: März 2018 Geplantes Studienende: Februar 2024	Abrocitinib 200 mg, Abrocitinib 100 mg

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer ggf. Datenschnitt	Therapiearme
JADE DARE (NCT04345367)	nein	ja	abgeschlossen	Studienbeginn: Juni 2020 Studienende: Juli 2021	Abrocitinib 200 mg, Dupilumab
<p>In den Studien JADE COMPARE (NCT03720470), JADE EXTEND (NCT03422822) und JADE DARE (NCT04345367) wurde in den jeweiligen Therapiearmen zusätzlich Placebo verabreicht, um die Verblindung der Behandlungszuordnung zu gewährleisten. ggf.: Gegebenenfalls; mg: Milligramm</p>					

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 Dokuments maßgeblichen Zeitpunkt liegen.

Der Studienstatus bezieht sich auf die RCT im Anwendungsgebiet mit Stand vom 16.11.2021.

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.

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
B7451006 (NCT02780167)	A3 (Studienkomparator entspricht nicht der zVT)
JADE MONO-1 (NCT03349060)	A3 (Studienkomparator entspricht nicht der zVT)
JADE MONO-2 (NCT03575871)	A3 (Studienkomparator entspricht nicht der zVT)
JADE REGIMEN (NCT03627767)	A3 (Studienkomparator entspricht nicht der zVT)
JADE MOA (NCT03915496)	A3 (Studienkomparator entspricht nicht der zVT)
JADE COMPARE (NCT03720470)	A6 (Studiendauer liegt unter der Mindestbehandlungsdauer von 24 Wochen)
JADE EXTEND (NCT03422822)	A3 (Studienkomparator entspricht nicht zVT)
zVT: Zweckmäßige Vergleichstherapie	

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie

viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.

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

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

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

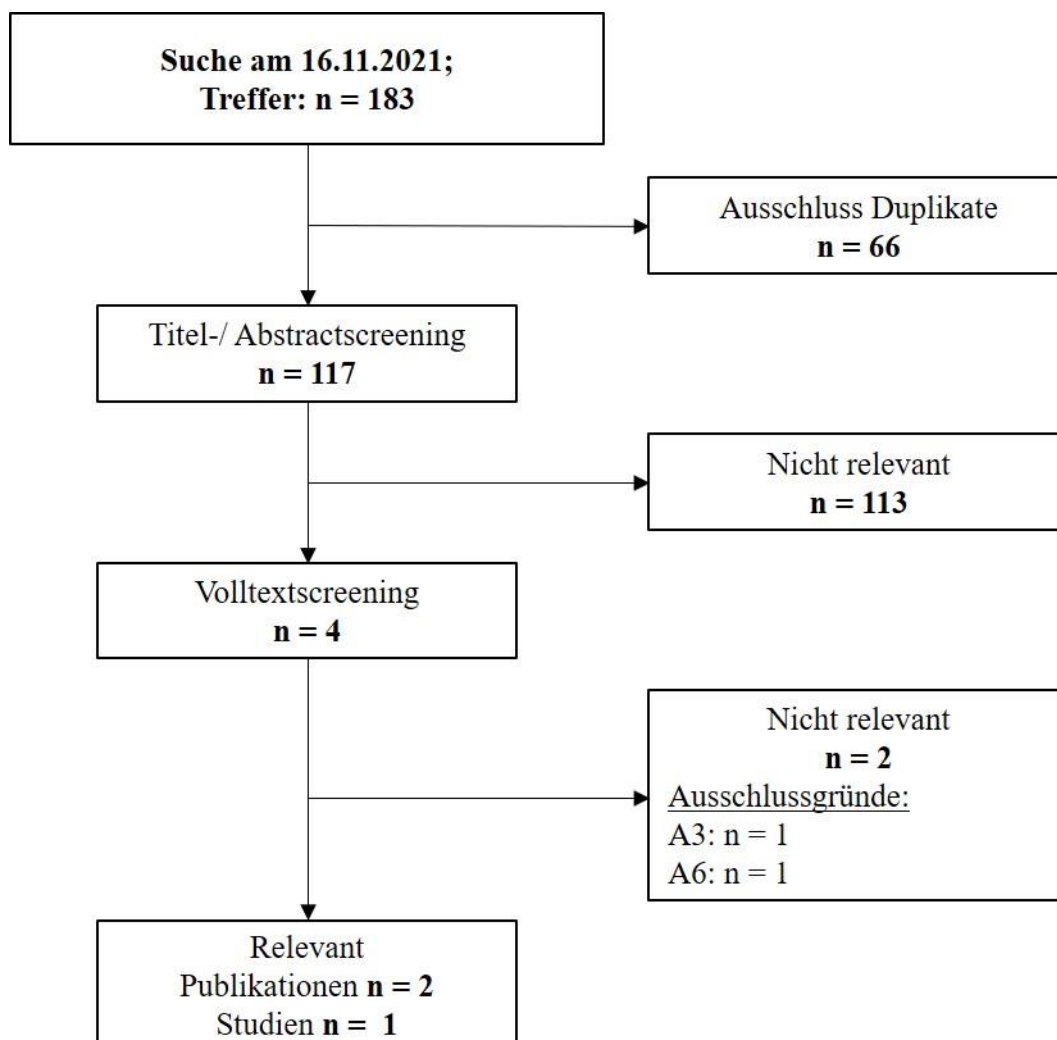


Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die systematische bibliografische Literaturrecherche nach RCT mit dem zu bewertenden Arzneimittel Abrocitinib wurde am 16.11.2021 wie in Abschnitt 4.2.3.2 beschrieben durchgeführt. Die angewendeten Suchstrategien finden sich in Anhang 4-A. Es konnten insgesamt 183 Treffer identifiziert werden. Die nach der Entfernung von Dubletten verbleibenden 117 Treffer wurden anhand des Titels und Abstracts anhand der in Abschnitt 4.2.2 spezifizierten Ein-/ Ausschlusskriterien selektiert. Es wurden 4 Treffer als potenziell relevant identifiziert und nachfolgend im Volltext evaluiert. Im Volltextscreening wurden 2 Publikationen ausgeschlossen, da sie entweder keinen Vergleich von Abrocitinib mit der vom G-BA festgelegten zVT Dupilumab enthielten (Ausschlussgrund: A3) oder die Studiendauer zu kurz war (Ausschlussgrund: A6). Bei den 2 als relevant eingeschlossenen Treffern handelt es sich um Registereinträge zu der Studie JADE DARE (53, 54).

4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-7) 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 / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Status (abgeschlossen/ abgebrochen/ laufend)
JADE DARE (NCT04345367)	<u>Clinicaltrial.gov</u> NCT04345367 (53) <u>EU-CTR</u> 2019-004013-13 (54) <u>WHO ICTRP</u> NCT04345367 / 2019-004013-13 (55)	ja	ja	abgeschlossen
<p>a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/ in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/ oder -ergebnisse.</p> <p>EU-CTR: <i>European Union Clinical Trials Register</i>; ICTRP: <i>International Clinical Trials Registry Platform</i>; WHO: <i>Weltgesundheitsorganisation (World Health Organization)</i></p>				

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 Angaben in Tabelle 4-9 bilden den Informationsstand vom 16.11.2021 ab.

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

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-7) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/ Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

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

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

Tabelle 4-10: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	Relevante Quellen ^a	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Studie durch Suche in Studienregistern / Studienergebnis datenbanken identifiziert (ja/nein)
-	-	-	-	-
a: Quellen aus der Suche auf der Internetseite des G-BA G-BA: Gemeinsamer Bundesausschuss				

Geben Sie an, welchen Stand die Information in Tabelle 4-10 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 am 16.11.2021 durchgeführte Suche nach relevanten Studien auf der Internetseite des G-BA ergab keine Ergebnisse.

4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2, 4.3.1.1.3 und 4.3.1.1.4) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich aller verfügbaren Quellen in Abschnitt 4.3.1.4 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

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

Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienberichte (ja/nein [Zitat])	Register-einträge ^c (ja/nein [Zitat])	Publikation und sonstige Quellen ^d (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebokontrolliert						
Keine						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
JADE DARE (NCT04 345367)	nein	ja	nein	ja (56)	ja (53–55)	nein
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (Liste der eingeschlossenen Studien) aufgeführt werden.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>d: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.</p> <p>G-BA: Gemeinsamer Bundesausschuss; ggf.: Gegebenenfalls; z. B.: Zum Beispiel</p>						

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

4.3.1.2.1 Studiendesign und Studienpopulationen

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

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

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

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

Studie	Studiendesign <RCT, doppelblind/einfach, verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
JADE DARE	RCT, Phase-IIIb, doppelblind, aktivkontrolliert, parallel, multizentrisch, international	Erwachsene (Alter ≥ 18 Jahre) mit mittelschwerer bis schwerer AD, die für eine systemische Therapie geeignet sind	Abrocitinib 200 mg + Hintergrundtherapie ^a (n = 365) Dupilumab 300 mg + Hintergrundtherapie ^a (n = 362)	Screening: 4 Wochen Behandlungsdauer: 26 Wochen Nachbeobachtung: 4 Wochen	Weltweit 220 Studienzentren (USA, Australien, Bulgarien, Kanada, Chile, Finnland, Deutschland, Ungarn, Italien, Korea, Lettland, Polen, Slowakei, Spanien, Taiwan) 6/2020 – 07/2021	<u>Primäre Endpunkte:</u> <i>Peak Pruritus</i> NRS (Woche 2) EASI-90 (Woche 4) <u>Sekundäre Endpunkte:</u> EASI-90 EASI-75 <i>Peak Pruritus</i> NRS DLQI Betroffene BSA SCORAD HADS EQ-5D POEM MOS <i>Sleep Scale</i> <i>Skin Pain</i> NRS Sicherheitsendpunkte
<p>Quelle: (52, 56)</p> <p>a: In beiden Behandlungsarmen war neben der nicht-medikamentösen und medikamentösen topischen Therapie auch eine <i>Rescue</i>-Therapie möglich. Die Verblindung erfolgte in Form eines Doppel-Dummy-Designs.</p> <p>AD: Atopische Dermatitis; BSA: Körperoberfläche (<i>Body Surface Area</i>); DLQI: <i>Dermatology Life Quality Index</i>; EASI: <i>Eczema Area and Severity Index</i>; EQ-5D: <i>European Quality of Life 5 Dimension</i>; etc.: <i>Et cetera</i>; ggf.: Gegebenenfalls; HADS: <i>Hospital Anxiety and Depression Scale</i>; mg: Milligramm; MOS: <i>Medical Outcomes Study</i>; NRS: <i>Numerical Rating Scale</i>; POEM: <i>Patient-Oriented Eczema Measure</i>; RCT: Randomisierte kontrollierte Studie (<i>Randomized Controlled Trial</i>); SCORAD: <i>Scoring Atopic Dermatitis</i>; USA: <i>United States of America</i>; z. B.: Zum Beispiel</p>						

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

Studie	Abrocitinib 200 mg	Dupilumab 300 mg	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in- Phase etc.</i>
JADE DARE	Abrocitinib 200 mg (2-mal 100 mg Tabletten), peroral (p.o.), 1-mal täglich + Placebo subkutan (s.c.) analog zu Dupilumab	Dupilumab 300 mg, s.c. Injektion, alle 2 Wochen (Anfangsdosis von 600 mg zu Studienbeginn) + Placebo p.o. analog zu Abrocitinib	<u>Hintergrundtherapie</u> <ul style="list-style-type: none"> • Nicht medikamentöse topische Therapie: Emollientien (z. B. Hyaluronsäure, Urea, Ceramide oder Filaggrin Abbauprodukte), 2-mal täglich • Medikamentöse topische Therapie: leichte bis mittelstark wirksame topische Glukokortikoide (TCS) und/ oder topische Calcineurin-Inhibitoren (TCI), 1-mal täglich <u>Rescue-Therapie</u> <ul style="list-style-type: none"> • Möglich nach 4 Wochen mit intolerablen AD-Symptomen • Stark wirksame TCS für bis zu 2 Wochen oder systemische Glukokortikoide für bis zu 10 Tage
Quelle: (56)			
AD: Atopische Dermatitis; ggf.: Gegebenenfalls; mg: Milligramm; p.o.: Peroral; s.c.: Subkutan; TCI: Topische Calcineurin-Inhibitoren; TCS: Topische Glukokortikosteroide (<i>Topical corticosteroids</i>); z. B.: Zum Beispiel			

Tabelle 4-14: Charakterisierung der Studienpopulationen für JADE DARE – RCT mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Demographische Charakteristika		
Geschlecht n (%)		
Männlich	193 (53,3)	204 (55,9)
Weiblich	169 (46,7)	161 (44,1)
Alter (Jahre)		
n	362	365
MW (SD)	36,6 (14,6)	35,5 (13,3)
Median (min-max)	33,0 (18,0-83,0)	32,0 (18,0-76,0)
Altersgruppe n (%)		
< 40 Jahre	230 (63,5)	247 (67,7)
≥ 40 Jahre	132 (36,5)	118 (32,3)
< 65 Jahre	341 (94,2)	354 (97,0)
≥ 65 Jahre	21 (5,8)	11 (3,0)
Abstammung n (%)		
Kaukasisch	269 (74,3)	248 (67,9)
Schwarz oder afroamerikanisch	25 (6,9)	26 (7,1)
Asiatisch	62 (17,1)	83 (22,7)
Andere	6 (1,7)	8 (2,2)
Region n (%)		
Vereinigte Staaten/ Kanada/ Australien	177 (48,9)	195 (53,4)
Europa	150 (41,4)	132 (36,2)
Asien	17 (4,7)	19 (5,2)
Lateinamerika	18 (5,0)	19 (5,2)
Gewicht (kg)		
n	362	365
MW (SD)	77,6 (19,1)	78,2 (21,5)
Median (min-max)	76,3 (37,8-208,7)	74,5 (40,4-188,5)
Gewichtsguppe n (%)		
< 70 kg	132 (36,5)	136 (37,3)
70 bis 100 kg	196 (54,1)	184 (50,4)
> 100 kg	34 (9,4)	45 (12,3)

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Krankheitsspezifische Charakteristika		
Erkrankungsdauer (Jahre)		
n	362	365
MW (SD)	24,2 (14,1)	24,1 (14,1)
Median (min-max)	22,9 (0,8-69,4)	22,8 (1,1-67-8)
Erkrankungsdauer-Gruppe n (%)		
< 26 Jahre	222 (61,3)	220 (60,3)
≥ 26 Jahre	140 (38,7)	145 (39,7)
IGA-Score n (%)		
Moderat (IGA=3)	216 (59,7)	220 (60,3)
Schwer (IGA=4)	146 (40,3)	145 (39,7)
EASI-Score		
n	362	365
MW (SD)	28,1 (11,5)	28,1 (11,9)
Median (min-max)	24,5 (7,2-68,4)	24,5 (13,3-72,0)
EASI-Gruppe n (%)		
16-25 Punkte	181 (50)	183 (50,1)
> 25 Punkte	172 (47,5)	174 (47,7)
Betroffene BSA (%)		
n	362	365
MW (SD)	42,5 (19,9)	42,6 (21,3)
Median (min-max)	39,0 (10,0-96,0)	36,0 (10,0-100,0)
Betroffene BSA (%) - Gruppe n (%)		
10-30 %	122 (33,7)	133 (36,4)
> 30-50 %	132 (36,5)	121 (33,2)
> 50 %	108 (29,8)	111 (30,4)
Peak Pruritus NRS		
n	362	365
MW (SD)	7,4 (1,6)	7,4 (1,6)
Median (min-max)	8,0 (2,0-10,0)	7,0 (3,0-10,0)
Peak Pruritus NRS-Gruppe n (%)		
4-6 Punkte	83 (22,9)	105 (28,8)
≥ 7 Punkte	274 (75,7)	259 (71,0)

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

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
SCORAD-Score		
n	362	365
MW (SD)	67,8 (12,8)	66,8 (12,7)
Median (min-max)	66,4 (37,4- 101,0)	65,2 (39,6- 102,2)
Vorbehandlung mit Ciclosporin n (%)		
Therapienaiv	323 (88,5)	315 (86,3)
Vorbehandelt	39 (10,8)	50 (13,7)
Vorherige Medikation für AD n (%)		
Nur topische Medikation	188 (51,9)	189 (51,8)
Systemische Medikation	172 (47,5)	176 (48,2)
Quelle: (56, 57) AD: Atopische Dermatitis; BSA: Körperoberfläche (<i>Body Surface Area</i>); EASI: <i>Eczema Area and Severity Index</i> ; IGA: <i>Investigator's Global Assessment</i> ; kg: Kilogramm; max: Maximal; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten in der Kategorie; NRS: <i>Numerical Rating Scale</i> ; SCORAD: Scoring Atopic Dermatitis, SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-15: Hintergrundtherapie der Studie JADE DARE anhand der *Safety*-Population

Studie JADE DARE	Abrocitinib 200 mg N = 362 n (%)	Dupilumab 300 mg N = 365 n (%)
Medikamentöse topische Therapie	353 (97,5)	362 (99,2)
TCS	338 (93,4)	353 (96,7)
mittelstark wirksame TCS	240 (66,3)	277 (75,9)
schwach wirksame TCS	220 (60,8)	220 (60,3)
TCI	70 (19,3)	73 (20,0)
Phosphodiesterase-4-Inhibitoren	5 (1,4)	3 (0,8)
Nicht-medikamentöse topische Therapie mit Emollentien	362 (100,0)	364 (99,7)
Quelle: (56) mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten in der Kategorie; TCI: Topische Calcineurin-Inhibitoren; TCS: Topische Glukokortikoide (<i>Topical corticosteroids</i>)		

Tabelle 4-16: *Rescue*-Therapie der Studie JADE DARE anhand der *Safety*-Population

Studie JADE DARE	Abrocitinib 200 mg	Dupilumab 300 mg
	N = 362 n (%)	N = 365 n (%)
Rescue-Therapie insgesamt	15 (4,1)	14 (3,8)
Topische Rescue-Therapie	12 (3,3)	12 (3,3)
stark wirksame TCS	12 (3,3)	12 (3,3)
Systemische Rescue-Therapie	4 (1,1)	2 (0,5)
nicht biologische Wirkstoffe	4 (1,1)	2 (0,5)
Kortikosteroide	4 (1,1)	2 (0,5)
Ciclosporin	0	1 (0,3)
biologische Wirkstoffe	0	0

Quelle: (56)
mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten in der Kategorie; TCS: Topische Glukokortikoide (*Topical corticosteroids*)

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

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

Studiendesign

Die Studie JADE DARE ist eine internationale, multizentrische, doppelblinde, *double-dummy*, randomisierte, aktivkontrollierte, Parallelgruppenstudie der Phase-IIIb. Das primäre Studienziel war der Vergleich der Wirksamkeit von Abrocitinib (PF-04965842) 200 mg 1-mal täglich mit Dupilumab 300 mg (bei einer Anfangsdosis von 600 mg zu Studienbeginn) bei erwachsenen Studienteilnehmern mit mittelschwerer bis schwerer AD. Beide Behandlungsgruppen erhielten eine begleitende Hintergrundtherapie und eine *Rescue*-Therapie nach Bedarf. Die sekundären Studienziele umfassten den Vergleich der Wirksamkeit und Verträglichkeit. In die Studie JADE DARE wurden erwachsene Patienten mit einer seit mindestens 1 Jahr bestehenden, bestätigten mittelschweren bis schweren AD (betroffene BSA ≥ 10 %, EASI-Score ≥ 16 und *Peak Pruritus* NRS-Score ≥ 4) eingeschlossen. Zudem sollten die Patienten ein dokumentiertes unzureichendes Ansprechen auf die Behandlung mit einer medikamentösen topischen Therapie für AD haben (angewendet über einen Zeitraum von

mindestens 4 Wochen) oder bereits eine systemische Therapie innerhalb des letzten Jahres erhalten haben.

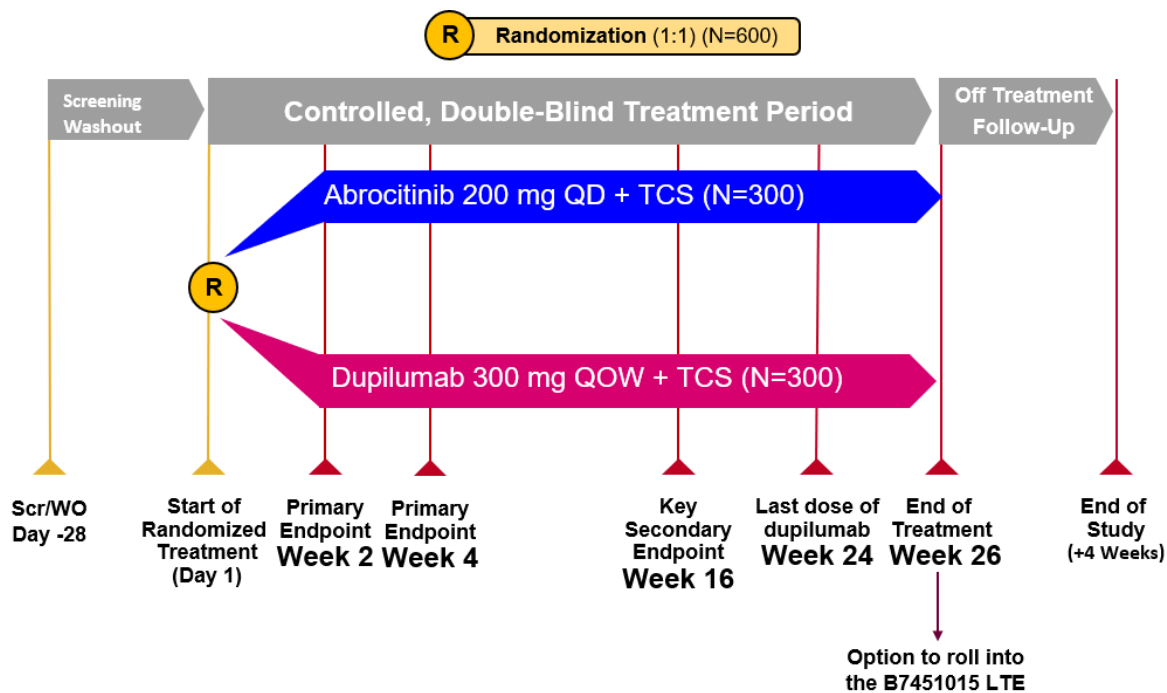


Abbildung 4-2: Studienschema der Studie JADE DARE

LTE: Long-term extension; mg: Milligramm; N: Anzahl der Patienten; TCS: Topische Glukokortikoide (*Topical corticosteroids*); QD: Täglich; QOW: Alle 2 Wochen; Scr/WO: *Screening/ Washout*

Insgesamt wurden 727 Patienten im Verhältnis 1:1 auf die beiden Behandlungsarme randomisiert (Abrocitinib: N = 362 versus Dupilumab: N = 365). Die Studie JADE DARE bestand aus einer Screeningphase von 4 Wochen, einer aktiven Behandlungsphase von 26 Wochen und einer Nachbeobachtungsphase von 4 Wochen (siehe Abbildung 4-2). Die Behandlung im Abrocitinib-Arm erfolgte mit einer oralen Einnahme von 200 mg Abrocitinib (2x 100 mg Tabletten) 1-mal täglich und einer s.c. Injektion von einem an Dupilumab angepassten Placebo alle 2 Wochen (2 Injektionen zu Studienbeginn; zur Vortäuschung der Anfangsdosis) von Tag 1 bis Woche 26 (die letzte Injektion des Placebos erfolgt zu Woche 24). Die Behandlung im Dupilumab-Vergleichsarm erfolgte als s.c. Injektion von 300 mg Dupilumab alle 2 Wochen (mit einer Anfangsdosis von 600 mg zu Studienbeginn) und einer oralen Einnahme von einem an Abrocitinib angepassten Placebo (2 Tabletten) 1-mal täglich von Tag 1 bis Woche 26 (die letzte Injektion von Dupilumab erfolgt zu Woche 24).

Alle Medikamente und Behandlungen, die die AD beeinflussen könnten, mussten vor Studienbeginn abgesetzt werden, mit Ausnahme von oralen Antihistaminika, nicht-medikamentösen topischen Therapien und medikamentösen topischen Therapien mit niedriger oder mittlerer Potenz. Zur Behandlung der AD konnten die Studienteilnehmer eine leichte oder mittelstark wirksame medikamentöse und nicht-medikamentöse topische Therapie anwenden, wobei das Ansprechen auf die Behandlung zu Studienbeginn unzureichend sein musste. Die Studienteilnehmer musste bereit und in der Lage sein, während der gesamten Dauer der Studie

die standardisierte topische Hintergrundtherapie einzuhalten. Die Hintergrundtherapie mit einer nicht medikamentösen topischen Therapie mit Emollientien ohne aktive Inhaltsstoffe oder Additiva (z. B. Hyaluronsäure, Urea, Ceramide oder Filaggrin-Abbauprodukte), die eine AD beeinflussen können, erfolgte begleitend zur Behandlungsphase 2-mal täglich beginnend mit Tag 1 (*Baseline*) und während der gesamten Studiendauer. Die Hintergrundtherapie einer medikamentösen topischen Therapie mit leicht bis mittelstark wirksamen TCS und/ oder TCI erfolgte 1-mal täglich auf die Bereiche mit aktiven Läsionen beginnend mit Tag 1 (*Baseline*) bis zum Ende der Studie oder bis die Läsionen unter Kontrolle waren. Nachdem die AD-Läsionen unter Kontrolle waren (erscheinungsfrei oder fast erscheinungsfrei), wurde die Anwendung von TCS nach weiteren 7 Tagen beendet. Die TCS-Behandlung wurde fortgesetzt, wenn die Läsionen zurückkehrten und das oben beschriebene Verfahren wurde nach dem Verschwinden der Läsionen angewendet. Eine *Rescue*-Therapie war nach 4 Wochen mit intolerablen AD-Symptomen möglich. Die *Rescue*-Therapie bestand aus stark wirksamen topischen TCS für bis zu 2 Wochen oder systemischen Glukokortikosteroiden für bis zu 10 Tage. Patienten, die eine systemische *Rescue*-Therapie erhalten hatten, mussten die Behandlung mit der Studienintervention vorübergehend unterbrechen. Eine topische *Rescue*-Therapie konnte weiterhin parallel zur Studienintervention angewendet werden. Patienten, die eine topische oder systemische *Rescue*-Therapie erhalten hatten, sollten die Studienvisiten und -bewertungen fortsetzen.

Patientencharakteristika

In die Studie JADE DARE wurden erwachsene Patienten ab 18 Jahren mit mittelschwerer bis schwerer AD eingeschlossen.

Insgesamt waren die Patienten in beiden Behandlungsarmen hinsichtlich der erfassten Patientencharakteristika (siehe Tabelle 4-14) vergleichbar. Die Patienten im Abrocitinib-Arm waren im Median 33,0 Jahre und im Dupilumab-Arm 32,0 Jahre alt. Die Patienten waren in beiden Behandlungsarmen mehrheitlich männlich (Abrocitinib: 53,3 % versus Dupilumab: 55,9 %). Die mediane Krankheitsdauer der AD lag in beiden Armen bei etwa 23 Jahren (Abrocitinib: 22,9 Jahre; Dupilumab: 22,8 Jahre).

Der mediane EASI-Score (Abrocitinib: 24,5 versus Dupilumab: 24,5), mediane *Peak Pruritus* NRS (Abrocitinib: 8,0 versus Dupilumab: 7,0), der mediane SCORAD-Score (Abrocitinib: 66,4 versus Dupilumab: 65,2) und die mediane BSA (Abrocitinib: 39,0 % versus Dupilumab: 36,0 %) waren zwischen den Behandlungsgruppen vergleichbar. Die Patientencharakteristika zu *Baseline* entsprachen somit den Kriterien der Checkliste der Aktualisierung „Systemtherapie bei Neurodermitis“ der S2k-Leitlinie Neurodermitis zur Beurteilung des objektiven Schweregrades und der subjektiven Belastung (6). Die Patienten kamen somit für eine systemische Therapie gemäß der Leitlinienkriterien in Frage.

Nahezu alle Patienten in beiden Behandlungsarmen erhielten im Studienverlauf kontinuierlich eine Hintergrundtherapie. Diese bestand aus einer nicht-medikamentösen Therapie (Abrocitinib: 100,0 % versus Dupilumab: 99,7 %) und den medikamentösen Therapien mit schwach wirksamen TCS (Abrocitinib: 60,8 % versus Dupilumab: 60,3 %), mittelstark

wirksamen TCS (Abrocitinib: 66,3 % versus Dupilumab: 75,9 %), TCI (Abrocitinib: 19,3 % versus Dupilumab: 20,0 %) oder Phosphodiesterase-4-Inhibitoren (Abrocitinib: 1,4 % versus Dupilumab: 0,8 %). Insgesamt waren die Patienten in beiden Behandlungsarmen hinsichtlich der Hintergrundtherapie vergleichbar.

Bis zu Woche 26 erhielten insgesamt 4,1 % der Patienten im Abrocitinib-Arm und 3,8 % der Patienten im Dupilumab-Arm mindestens eine *Rescue*-Therapie. Der Anteil der Patienten, die topische stark wirksame TCS für bis zu 2 Wochen erhielten, war in beiden Studienarmen identisch (3,3 %). Im Dupilumab-Arm erhielten weniger Patienten systemische nicht biologische Wirkstoffe für bis zu 10 Tage als im Abrocitinib-Arm (Abrocitinib: 1,1 % versus Dupilumab: 0,5 %).

In Tabelle 4-17 sind die Behandlungsdauern in beiden Armen dargestellt. Die Behandlungsdauern zwischen den beiden Behandlungsgruppen unterscheiden sich nicht.

Tabelle 4-17: Behandlungsdauer der Studie JADE DARE anhand der *Safety*-Population

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Behandlungsdauer (Tage)		
MW (SD)	166,6 (39,5)	171,2 (27,9)
Median (min-max)	180,0 (2,0-230,0)	180,0 (2,0-194,0)
Quelle: Table 14.4.1.1 (57)		
N: Anzahl der in der Analyse eingeschlossenen Patienten; max: Maximal; min: Minimal; MW: Mittelwert; SD: Standardabweichung (<i>Standard Deviation</i>)		

Übertragbarkeit auf den deutschen Versorgungskontext

Die Studie JADE DARE wurde unter anderem in Deutschland und primär westlichen Industrieländern (USA, Australien, Bulgarien, Kanada, Chile, Finnland, Ungarn, Italien, Korea, Lettland, Polen, Slowakei, Spanien, Taiwan) durchgeführt. Der Großteil der Patienten war kaukasischer Abstammung (Abrocitinib: 74,3 % versus Dupilumab: 67,9 %). Zudem zeigten entsprechende Subgruppenanalysen nach Abstammung (kaukasisch versus schwarz oder afroamerikanisch versus asiatisch versus andere) und Region (USA/ Kanada/ Australien versus Europa versus Asien versus Lateinamerika) keine fazitrelevanten Effektmodifikationen (siehe Abschnitt 4.3.1.3.2), sodass davon ausgegangen werden kann, dass die Ergebnisse der Studie JADE DARE auf den deutschen Versorgungskontext übertragbar sind.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

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

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

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

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
JADE DARE	ja	ja	ja	ja	ja	ja	niedrig
RCT: Randomisierte kontrollierte Studie (<i>Randomized Controlled Trial</i>)							

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei der Studie JADE DARE handelt es sich um eine multizentrische, randomisierte, doppelblinde, *double-dummy* und aktivkontrollierte Studie. Die Studienteilnahme setzte eine unterzeichnete Einwilligungserklärung und die Eignung des Patienten für die Studienteilnahme, wie vom Prüfarzt festgestellt (Erfüllung aller Einschlusskriterien und keines der Ausschlusskriterien), voraus. Die Randomisierungssequenz wurde computerbasiert erzeugt und die Gruppenzuteilung erfolgte verdeckt mittels *Interactive Response Technology* (IRT). Im Ergebnis lieferte die Randomisierung gut balancierte Behandlungsgruppen bezüglich der betrachteten Patientencharakteristika (siehe Tabelle 4-14). Die Randomisierung wird somit als adäquat eingeschätzt. Patienten und behandelnde Personen waren verblindet. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte oder Hinweise für eine ergebnisabhängige Berichterstattung identifiziert werden. Somit wird das Verzerrungspotenzial für die Studie JADE DARE 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 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Morbidität							Lebens- qualität	Verträglich- keit	
	Krankheitsschwere	Betroffene BSA ^a	Patientenberichtete Symptomatik	Juckreiz	Schlafstörungen	Hautschmerz	Ängste und Depressionen			Gesundheitszustand
JADE DARE	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja

a: Supportiver Endpunkt
BSA: Körperoberfläche (*Body Surface Area*)

4.3.1.3.1 Endpunkte – RCT

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

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

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

Unterschiedliche Beobachtungszeiten zwischen den Behandlungsgruppen sollen durch adäquate Analysen (z.B. Überlebenszeitanalysen) adressiert werden, und zwar für alle

Endpunkte (einschließlich UE nach den nachfolgend genannten Kriterien), für die eine solche Analyse aufgrund deutlich unterschiedlicher Beobachtungszeiten erforderlich ist.

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

Zu mit Skalen erhobenen patientenberichteten Endpunkten (z.B. zur gesundheitsbezogenen Lebensqualität oder zu Symptomen) sind immer auch die Werte im Studienverlauf anzugeben, auch als grafische Darstellung, sowie eine Auswertung, die die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt (z.B. als Symptomlast über die Zeit, geschätzt mittels MMRM-Analyse [falls aufgrund der Datenlage geeignet]).

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen:

1. Gesamtrate UE,
2. Gesamtrate schwerwiegender UE (SUE),
3. Gesamtrate der Abbrüche wegen UE,
4. Gesamtraten von UE differenziert nach Schweregrad, sofern dies in der/den relevante/n Studie/n erhoben wurde (z.B. gemäß CTCAE und/oder einer anderen etablierten bzw. validierten indikationsspezifischen Klassifikation) einschließlich einer Abgrenzung schwerer und nicht schwerer UE,
5. zu den unter 1, 2 und 4 genannten Kategorien (UE ohne weitere Differenzierung, SUE, UE differenziert nach Schweregrad) soll zusätzlich zu den Gesamtraten die Darstellung nach Organsystemen und Einzelereignissen (als System Organ Class [SOCs] und Preferred Terms [PT] nach MedDRA) jeweils nach folgenden Kriterien erfolgen:
 - UE (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10% der Patienten in einem Studienarm aufgetreten sind
 - Schwere UE (z.B. CTCAE-Grad ≥ 3) und SUE: Ereignisse, die bei mindestens 5% der Patienten in einem Studienarm aufgetreten sind
 - zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patienten UND bei mindestens 1 % der Patienten in einem Studienarm aufgetreten sind.
6. A priori definierte UE von besonderem Interesse [AESI]) sowie prädefinierte SOC-übergreifende UE-Auswertungen (z.B. als Standardised MedDRA Queries, SMQs) sollen unabhängig von der Ereignisrate dargestellt werden und zwar differenziert nach Schweregrad (dargestellt als Gesamtrate und differenziert nach Schweregrad, nicht schwer, schwer, schwerwiegend).

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

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

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

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

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

4.3.1.3.1.1 Morbidität – RCT

4.3.1.3.1.1.1 Krankheitsschwere (EASI) – RCT

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

Tabelle 4-20: Operationalisierung des Endpunkts EASI

Studie	Operationalisierung
JADE DARE	<p>Der <i>Eczema Area and Severity Index</i> (EASI) ist ein validiertes Standardinstrument zur Erfassung der objektiven Krankheitsschwere und des Therapieerfolgs bei atopischer Dermatitis. Im EASI werden das flächenhafte Ausmaß (1-9 %, 10-29 %, 30-49 %, 50-69 %, 70-89 % und 90-100 %) und die Intensität (keine (0), leichte (1), mäßige (2) und starke (3)) der Hautveränderungen in den 4 Körperbereichen Kopf und Hals, Rumpf, obere Extremität und untere Extremität bestimmt. Betrachtet werden dabei die Symptome Erythem, Papulae/ Induration, Exkoriation und Lichenifikation. Der EASI-Gesamtscore kann Werte von 0-72 annehmen, wobei ein höherer Wert einem höheren Schweregrad entspricht.</p> <p><u>Hauptanalysen:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine 75 %ige Verbesserung (Reduktion) gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen (EASI-75) • Anteil der Patienten, die eine 90 %ige Verbesserung (Reduktion) gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen (EASI-90) <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Mittlere prozentuale Änderung gegenüber dem <i>Baseline</i>-Wert anhand eines <i>Mixed Model Repeated Measures</i> (MMRM) zu Woche 2, 4, 8, 12, 16, 20 und 26 <p>Die Haupt- und Sensitivitätsanalysen erfolgten im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienmedikation erhalten haben.</p> <p>In den Responderanalysen wurden Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte mittels <i>Non-Response</i>-Imputation ersetzt.</p> <p>Die folgenden <i>Estimands</i> wurden betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 1 (Composite Estimand)</i> für Responderanalysen • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere prozentuale Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-Haenszel-Tests. Die Stratifizierung beinhaltete den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA=3 vs. IGA=4).</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares</i> (LS)-Mean-Differenzen und des Hedges'g (zur Beurteilung der klinischen Relevanz von Effekten) mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.</p>
<p>IGA: <i>Investigator's Global Assessment</i>; EASI: <i>Eczema Area and Severity Index</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); MMRM: <i>Mixed Model Repeated Measures</i>; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko</p>	

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

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

Tabelle 4-21: Bewertung des Verzerrungspotenzials für EASI 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
EASI: <i>Eczema Area and Severity Index</i> ; ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Verblindung der Patienten und Behandler stellte sicher, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das *Intention-to-treat* (ITT)-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

Tabelle 4-22: Beobachtungsdauer für EASI in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	175,6 (36,66)	179,4 (26,36)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-221,0)
Quelle: Table 14.2.2.6.3 (57) EASI: <i>Eczema Area and Severity Index</i> ; max: Maximal; mg: Milligramm; Min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-23: Rücklaufquote für EASI aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 2	350/362 (96,7)	350/365 (95,9)
Woche 4	342/362 (94,5)	351/365 (96,2)
Woche 8	336/362 (92,8)	348/365 (95,3)
Woche 12	329/362 (90,9)	341/365 (93,4)
Woche 16	324/362 (89,5)	336/365 (92,1)
Woche 20	320/362 (88,4)	333/365 (91,2)
Woche 26	301/362 (83,1)	324/365 (88,8)
Quelle: Table 14.2.2.2.1 (57) EASI: <i>Eczema Area and Severity Index</i> ; FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für den EASI zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

Tabelle 4-24: Ergebnisse für Krankheitsschwere (EASI) aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
EASI-75 zu Woche 26				
254/362 (70,2)	261/365 (71,5)	0,98 [0,89; 1,08] p=0,6980	0,94 [0,68; 1,29] p=0,6980	-1,3 [-7,9; 5,3] p=0,6983
EASI-90 zu Woche 26				
190/362 (52,5)	172/365 (47,1)	1,11 [0,96; 1,29] p=0,1472	1,24 [0,93; 1,66] p=0,1468	5,4 [-1,9; 12,6] p=0,1470
Quelle: Table 14.2.2.6.1.1.1, Table 14.2.2.6.2.1.1 (57)				
Werte nach Therapieabbruch und nach <i>Rescue</i> -Therapie sowie fehlende Werte wurden mittels <i>Non-Response</i> -Imputation ergänzt (<i>Composite Estimand</i> , siehe Abschnitt 4.2.5.2).				
a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i> -Kriterium erreicht haben.				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).				
EASI: <i>Eczema Area and Severity Index</i> ; IGA: <i>Investigator's Global Assessment</i> ; FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; mg: Milligramm; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i> ; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus				

Zu Woche 26 wiesen im Abrocitinib-Arm 70,2 % der Patienten eine Verbesserung des EASI-Scores um mindestens 75 % gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 71,5 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 0,98 [0,89; 1,08]; p-Wert = 0,6980).

Eine Verbesserung des EASI-Scores um mindestens 90 % zeigten zum selben Zeitpunkt im Abrocitinib-Arm 52,5 % der Patienten; unter Dupilumab lag der entsprechende Anteil der Patienten bei 47,1 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 1,11 [0,96; 1,29]; p-Wert = 0,1472).

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

Tabelle 4-25: Ergebnisse für mittlere prozentuale Veränderung des EASI; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean- Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean- Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere prozentuale Veränderung des EASI								
Baseline	362	28,1 (11,5)	NA	365	28,1 (11,9)	NA	NA	NA
Woche 2	350	12,7 (10,6)	-56,18 (1,53)	350	15,3 (11,3)	-45,72 (1,53)	-10,46 [-14,71; -6,22] p<0,0001	-0,37 [-0,52; -0,22]
Woche 4	342	7,6 (8,8)	-73,76 (1,36)	351	10,9 (9,9)	-61,76 (1,36)	-11,99 [-15,77; -8,22] p<0,0001	-0,47 [-0,62; -0,32]
Woche 8	336	5,1 (6,6)	-81,70 (1,25)	348	7,8 (8,2)	-72,46 (1,24)	-9,25 [-12,70; -5,79] p<0,0001	-0,40 [-0,55; -0,25]
Woche 12	329	4,1 (5,8)	-84,40 (1,05)	341	6,3 (7,3)	-78,05 (1,04)	-6,35 [-9,25; -3,44] p<0,0001	-0,33 [-0,48; -0,18]
Woche 16	324	3,8 (6,1)	-86,39 (1,03)	336	5,3 (6,7)	-81,14 (1,02)	-5,24 [-8,09; -2,40] p=0,0003	-0,28 [-0,44; -0,13]
Woche 20	320	3,5 (5,9)	-87,40 (0,98)	333	4,7 (6,6)	-83,42 (0,97)	-3,98 [-6,69; -1,27] p=0,0040	-0,23 [-0,38; -0,07]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean- Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean- Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Woche 26	301	3,3 (5,2)	-87,03 (1,00)	324	4,2 (5,8)	-84,96 (0,98)	-2,07 [-4,83; 0,69] p=0,1408	-0,12 [-0,28; 0,04]

Quelle: Table 14.2.2.2.1, Table 14.2.2.2.4 (57)

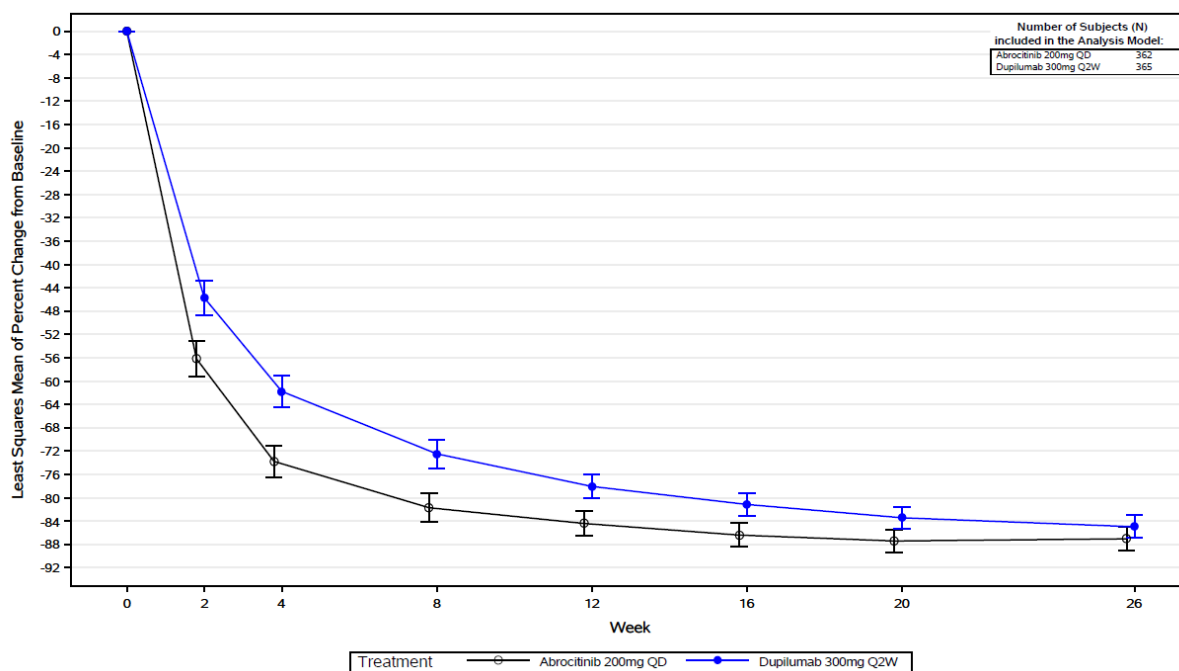
Werte nach Therapieabbruch und nach *Rescue*-Therapie wurden als fehlende Werte betrachtet (*Hypothetical Estimand*, siehe Abschnitt 4.2.5.2).

a: Deskriptive Analyse basierend auf den Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).

b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie *Baseline*-Wert + Krankheitsschwere zu *Baseline* als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem *Restricted Maximum Likelihood* Ansatz.

EASI: *Eczema Area and Severity Index*; FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (*Least Squares*); mg: Milligramm; MMRM: *Mixed Model Repeated Measures*; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SD: Standardabweichung (*Standard Deviation*); SE: Standardfehler (*Standard Error*); vs.: Versus

In der Analyse der mittleren prozentualen Veränderung des EASI-Scores gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 2, 4, 8, 12, 16 und 20 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind zu Woche 2, 4 und 8 als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges' g) außerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	2	4	8	12	16	20	26
N* Abrocitinib 200 mg	362	350	342	336	329	324	320	301
N* Dupilumab 300 mg	365	350	351	348	341	336	333	324

Abbildung 4-3: Verlaufskurve für die mittlere prozentuale Veränderung des EASI von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel
mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren prozentualen Veränderungen im EASI zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-3).

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

Es wurde keine Meta-Analyse durchgeführt, da die Voraussetzungen dafür nicht erfüllt sind.

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

4.3.1.3.1.2 Krankheitsschwere (SCORAD) – RCT

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

Tabelle 4-26: Operationalisierung des Endpunkts SCORAD

Studie	Operationalisierung
JADE DARE	<p>Der Scoring Atopic Dermatitis (SCORAD) ist ein etabliertes und validiertes Instrument zur Erfassung des flächenhaften Ausmaßes und der Intensität der Hautveränderungen. Es erfolgt die Beurteilung des Schweregrades (keine (0), leichte (1), mäßige (2) und starke (3)) von 6 Symptomen (Erythem, Ödem/ Knötchenbildung, nässende Areale/ Krustenbildung, Exkoration, Lichenifikation und Trockenheit nicht-betroffener Hautareale). Vom Patienten selbst werden die Symptome Schlafstörung und Juckreiz in den letzten 3 Tage bzw. Nächten jeweils auf einer VAS von 0 (keine Symptome) bis 10 (schwerste Symptome) erhoben. Aus den 3 Komponenten des SCORAD wird ein Gesamtscore errechnet. Der SCORAD-Gesamtscore kann Werte von 0-103 annehmen, wobei ein höherer Wert einem höheren Schweregrad entspricht.</p> <p><u>Hauptanalysen:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine 75 %ige Verbesserung (Reduktion) gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen (SCORAD-75) • Anteil der Patienten, die eine 90 %ige Verbesserung (Reduktion) gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen (SCORAD-90) <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Mittlere prozentuale Änderung gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 2, 4, 8, 12, 16, 20 und 26 <p>Die Haupt- und Sensitivitätsanalysen erfolgten im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben.</p> <p>In den Responderanalysen wurden Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte mittels <i>Non-Response</i>-Imputation ersetzt.</p> <p>Die folgenden <i>Estimands</i> wurden betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 1 (Composite Estimand)</i> für Responderanalysen • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-Haenszel-Tests. Die Stratifizierung beinhaltete den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares</i> (LS)-<i>Mean</i>-Differenzen und des Hedges'g mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.</p>
	<p>IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); MMRM: <i>Mixed Model Repeated Measures</i>; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; SCORAD: <i>Scoring Atopic Dermatitis</i>; vs.: Versus</p>

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 4-27: Bewertung des Verzerrungspotenzials für SCORAD 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie; SCORAD: <i>Scoring Atopic Dermatitis</i>						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

Tabelle 4-28: Beobachtungsdauer für SCORAD in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	175,1 (37,81)	179,2 (26,5)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-221,0)
Quelle: Table 14.2.13.9.10 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SCORAD: <i>Scoring Atopic Dermatitis</i> ; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-29: Rücklaufquote für SCORAD in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
Baseline	362/362 (100,0)	365/365 (100,0)
Woche 2	349/362 (96,4)	350/365 (95,9)
Woche 4	340/362 (93,9)	351/365 (96,2)
Woche 8	335/362 (92,5)	348/365 (95,3)
Woche 12	328/362 (90,6)	341/365 (93,4)
Woche 16	324/362 (89,5)	336/365 (92,1)
Woche 20	319/362 (88,1)	333/365 (91,2)
Woche 26	300/362 (82,9)	323/365 (88,5)
Quelle: Table 14.2.13.9.4.1 (57) FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS; SCORAD: <i>Scoring Atopic Dermatitis</i>		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für den SCORAD zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 82 %.

Tabelle 4-30: Ergebnisse für Krankheitsschwere (SCORAD) aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200mg	Dupilumab 300mg	Abrocitinib 200mg vs. Dupilumab 300mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
SCORAD-75 zu Woche 26				
152/362 (42,0)	133/365 (36,4)	1,15 [0,96; 1,38] p=0,1282	1,26 [0,94; 1,70] p=0,1278	5,5 [-1,5; 12,6] p=0,1273
SCORAD-90 zu Woche 26				
80/362 (22,1)	52/365 (14,2)	1,55 [1,13; 2,13] p=0,0067	1,71 [1,16; 2,51] p=0,0064	7,9 [2,3; 13,4] p=0,0061
Quelle: Table 14.2.13.9.1.1.1, Table 14.2.13.9.2.1.1 (57)				
Werte nach Therapieabbruch und nach <i>Rescue</i> -Therapie sowie fehlende Werte wurden mittels <i>Non-Response</i> -Imputation ergänzt (<i>Composite Estimand</i> , siehe Abschnitt 4.2.5.2).				
a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i> -Kriterium erreicht haben.				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).				
IGA: <i>Investigator's Global Assessment</i> ; FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; mg: Milligramm;				
n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten;				
OR: <i>Odds Ratio</i> ; RD: Risikodifferenz; RR: Relatives Risiko; SCORAD: <i>Scoring Atopic Dermatitis</i> ; vs.: Versus				

Zu Woche 26 wiesen im Abrocitinib-Arm 42,0 % der Patienten eine Verbesserung des SCORAD-Scores um mindestens 75 % gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 36,4 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 1,15 [0,96; 1,38]; p-Wert = 0,1282).

Es zeigte sich für das SCORAD-90-Ansprechen ein statistisch signifikanter Unterschied zugunsten von Abrocitinib gegenüber Dupilumab zum selben Zeitpunkt (RR [95 %-KI]: 1,55 [1,13; 2,13]; p-Wert = 0,0067). Eine Verbesserung des SCORAD-Scores um mindestens 90 % zeigten im Abrocitinib-Arm 22,1 % der Patienten; unter Dupilumab lag der entsprechende Anteil der Patienten bei 14,2 %.

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

Tabelle 4-31: Ergebnisse für mittlere prozentuale Veränderung im SCORAD; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere prozentuale Veränderung im SCORAD								
Baseline	362	67,8 (12,8)	NA	365	66,8 (12,7)	NA	NA	NA
Woche 2	349	38,0 (17,5)	-44,52 (1,24)	350	44,4 (18,0)	-33,47 (1,24)	-11,05 [-14,49; -7,61] p<0,0001	-0,48 [-0,63; -0,33]
Woche 4	340	27,5 (15,8)	-59,58 (1,18)	351	35,7 (17,1)	-46,83 (1,17)	-12,75 [-16,02; -9,49] p<0,0001	-0,58 [-0,74; -0,43]
Woche 8	335	22,9 (15,2)	-65,84 (1,22)	348	29,4 (15,6)	-55,64 (1,20)	-10,20 [-13,57; -6,84] p<0,0001	-0,46 [-0,61; -0,30]
Woche 12	328	20,9 (14,7)	-67,90 (1,17)	341	26,1 (14,8)	-60,59 (1,16)	-7,31 [-10,54; -4,08] p=0,0001	-0,34 [-0,50; -0,19]
Woche 16	324	19,4 (14,9)	-70,58 (1,16)	336	23,5 (14,7)	-64,44 (1,14)	-6,14 [-9,33; -2,96] p=0,0002	-0,29 [-0,45; -0,14]
Woche 20	319	18,3 (14,8)	-71,78 (1,19)	333	21,6 (14,7)	-66,84 (1,17)	-4,94 [-8,22; -1,67] p=0,0031	-0,23 [-0,39; -0,08]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Woche 26	300	18,0 (14,8)	-71,47 (1,23)	323	20,9 (14,6)	-68,21 (1,20)	-3,26 [-6,63; 0,11] p=0,0578	-0,15 [-0,31; 0,01]

Quelle: Table 14.2.13.1, Table 14.2.13.9.4.1 (57)

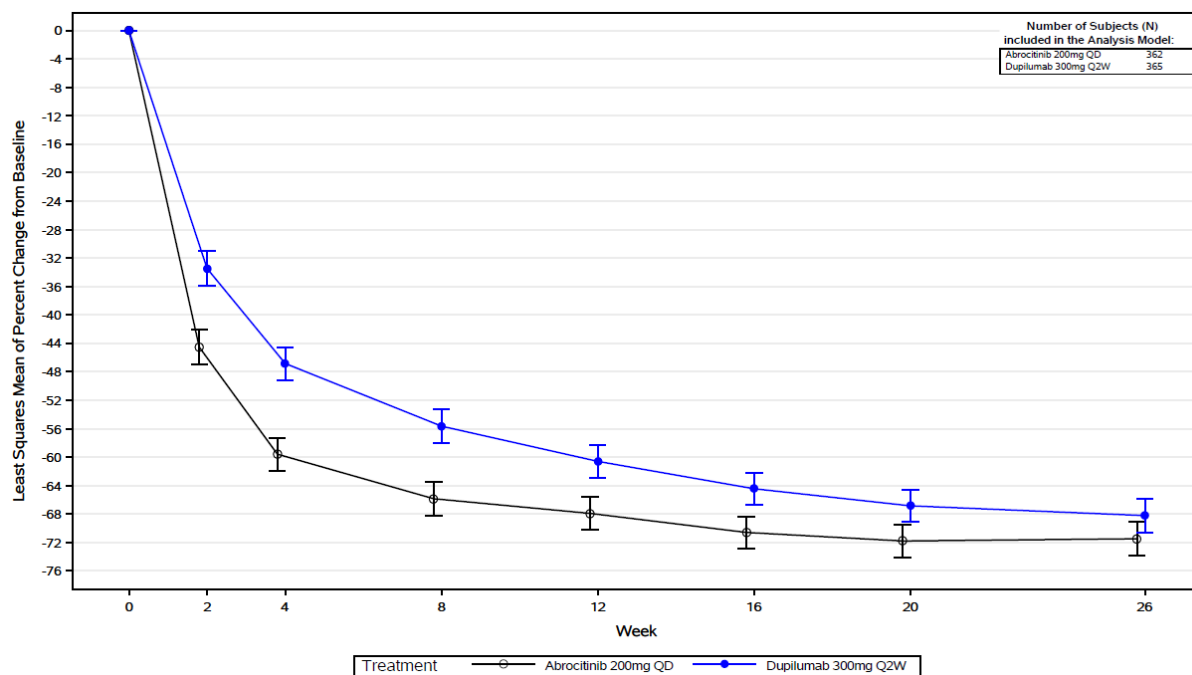
Werte nach Therapieabbruch und nach *Rescue*-Therapie wurden als fehlende Werte betrachtet (*Hypothetical Estimand*, siehe Abschnitt 4.2.5.2).

a: Deskriptive Analyse basierend auf den Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).

b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie *Baseline*-Wert + Krankheitsschwere zu *Baseline* als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem *Restricted Maximum Likelihood* Ansatz.

FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (*Least Squares*); mg: Milligramm; MMRM: *Mixed Model Repeated Measures*; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SCORAD: *Scoring Atopic Dermatitis*; SD: Standardabweichung (*Standard Deviation*); SE: Standardfehler (*Standard Error*); vs.: Versus

In der Analyse der mittleren prozentualen Veränderung des SCORAD-Scores gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 2, 4, 8, 12, 16 und 20 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind zu Woche 2, 4 und 8 als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges' g) außerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	2	4	8	12	16	20	26
N* Abrocitinib 200 mg	362	349	340	335	328	324	319	300
N* Dupilumab 300 mg	365	350	351	348	341	336	333	323

Abbildung 4-4: Verlaufskurve für die mittlere prozentuale Veränderung des SCORAD von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel
mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren prozentualen Veränderungen im SCORAD zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-4).

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 die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.1.3 Betroffene BSA – RCT

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

Tabelle 4-32: Operationalisierung des supportiven Endpunkts betroffene BSA

Studie	Operationalisierung
JADE DARE	<p>Bei der Beurteilung der von der atopischen Dermatitis betroffenen Körperoberfläche (<i>Body Surface Area, BSA</i>) werden die 4 Körperbereiche Kopf und Hals, Rumpf, obere Extremitäten und untere Extremitäten betrachtet. Die Bewertung der BSA erfolgt auf einer Skala von 0-100 %.</p> <p>Die Analyse der BSA wird im Rahmen des Nutzendossiers angelehnt an die präspezifizierte Operationalisierung in der Studie JADE DARE ergänzend dargestellt.</p> <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> Mittlere prozentuale Änderung gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 2, 4, 8, 12, 16, 20 und 26 <p>Die Analysen erfolgten im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben.</p> <p>Der folgende <i>Estimand</i> wurde betrachtet:</p> <ul style="list-style-type: none"> <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik des <i>Estimand</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares</i> (LS)-<i>Mean-Differenzen</i> und des <i>Hedges'g</i> mit den korrespondierenden 95 %-Konfidenzintervallen (KI) zu jedem Beobachtungszeitpunkt.</p>
BSA: Körperoberfläche (<i>Body Surface Area</i>); FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); MMRM: <i>Mixed Model Repeated Measures</i>	

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 4-33: Bewertung des Verzerrungspotenzials für betroffene BSA 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
BSA: Betroffene Körperoberfläche (<i>Body Surface Area</i>); ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft. Der Endpunkt BSA wird im Dossier ergänzend dargestellt.

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

Tabelle 4-34: Beobachtungsdauer für betroffene BSA in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	175,8 (36,62)	179,4 (26,36)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-221,0)
Quelle: Table 14.2.4.6.10 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-35: Rücklaufquote für betroffene BSA in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 2	350/362 (96,7)	350/365 (95,9)
Woche 4	342/362 (94,5)	351/365 (96,2)
Woche 8	336/362 (92,8)	348/365 (95,3)
Woche 12	329/362 (90,9)	342/365 (93,7)
Woche 16	324/362 (89,5)	336/365 (92,1)
Woche 20	320/362 (88,4)	333/365 (91,2)
Woche 26	302/362 (83,4)	324/365 (88,8)
Quelle: Table 14.2.4.6.3 (57) BSA: Betroffene Körperoberfläche (<i>Body Surface Area</i>); FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für die betroffene BSA zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

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

Tabelle 4-36: Ergebnisse für mittlere prozentuale Veränderung der betroffenen BSA; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere prozentuale Veränderung der BSA								
Baseline	362	42,5 (19,9)	NA	365	42,6 (21,3)	NA	NA	NA
Woche 2	350	25,4 (20,4)	-42,67 (1,69)	350	28,3 (20,7)	-33,36 (1,69)	-9,31 [-13,99; -4,62] p=0,0001	-0,29 [-0,44; -0,15]
Woche 4	342	17,1 (18,5)	-62,00 (1,78)	351	21,4 (19,0)	-49,53 (1,77)	-12,47 [-17,39; -7,55] p<0,0001	-0,38 [-0,53; -0,23]
Woche 8	336	11,4 (15,0)	-73,95 (1,63)	348	15,9 (17,3)	-62,84 (1,62)	-11,12 [-15,63; -6,60] p<0,0001	-0,37 [-0,52; -0,22]
Woche 12	329	9,0 (12,4)	-78,79 (1,57)	342	12,9 (15,6)	-69,41 (1,55)	-9,38 [-13,71; -5,06] p<0,0001	-0,33 [-0,48; -0,18]
Woche 16	324	8,3 (13,5)	-80,62 (1,45)	336	11,3 (14,3)	-73,67 (1,43)	-6,95 [-10,95; -2,95] p=0,0007	-0,27 [-0,42; -0,11]
Woche 20	320	7,8 (13,4)	-82,23 (1,33)	333	9,8 (13,8)	-76,91 (1,31)	-5,32 [-9,00; -1,65] p=0,0045	-0,22 [-0,38; -0,07]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Woche 26	302	7,3 (12,5)	-82,31 (1,36)	324	8,9 (12,8)	-78,95 (1,34)	-3,36 [-7,10; 0,39] p=0,0793	-0,14 [-0,30; 0,02]

Quelle: Table 14.2.4.6.1, Table 14.2.4.6.3 (57)

Werte nach Therapieabbruch und nach *Rescue*-Therapie wurden als fehlende Werte betrachtet (*Hypothetical Estimand*, siehe Abschnitt 4.2.5.2).

a: Deskriptive Analyse basierend auf den Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).

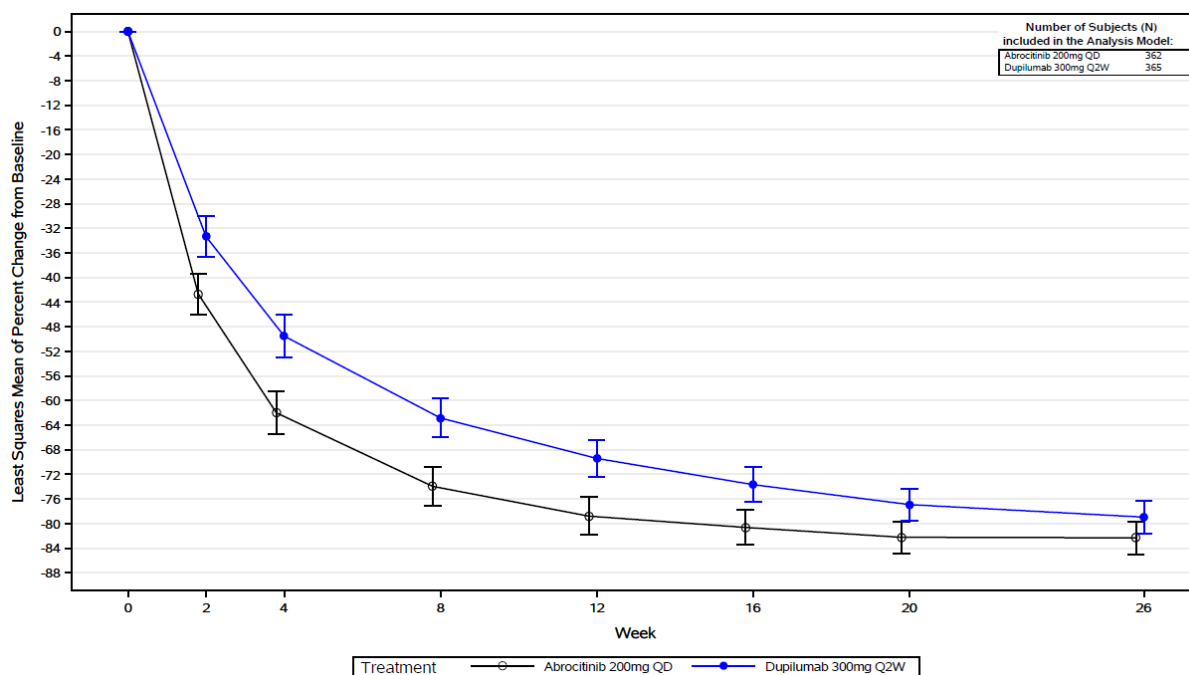
b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie *Baseline*-Wert + Krankheitsschwere zu *Baseline* als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem *Restricted Maximum Likelihood* Ansatz.

BSA: Körperoberfläche (*Body Surface Area*); FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (*Least Squares*);

mg: Milligramm; MMRM: *Mixed Model Repeated Measures*; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SD: Standardabweichung (*Standard Deviation*);

SE: Standardfehler (*Standard Error*); vs.: Versus

In der Analyse der mittleren prozentualen Veränderung der BSA gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 2, 4, 8, 12, 16 und 20 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind zu Woche 4 und 8 als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges'g) außerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	2	4	8	12	16	20	26
N* Abrocitinib 200 mg	362	350	342	336	329	324	320	302
N* Dupilumab 300 mg	365	350	351	348	342	336	333	324

Abbildung 4-5: Verlaufskurve für die mittlere prozentuale Veränderung der BSA von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel
mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren prozentualen Veränderungen der BSA zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-5).

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 die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.1.1.4 Patientenberichtete Symptomatik (POEM) – RCT

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

Tabelle 4-37: Operationalisierung des Endpunkts patientenberichtete Symptomatik

Studie	Operationalisierung
JADE DARE	<p>Der Patient-Oriented Eczema Measure (POEM) ist ein patientenberichtetes und validiertes Messinstrument zur Selbstbeurteilung der Häufigkeit des Auftretens von 7 Symptomen (Juckreiz, Schlafstörungen, Blutungen der Haut, nässende Haut, rissige Haut, schuppige Haut, trockene/ raue Haut) bei atopischer Dermatitis innerhalb der vergangenen Woche auf einer 5-Punkte-Skala (0 = an keinem Tag, 1 = an 1-2 Tagen, 2 = an 3-4 Tagen, 3 = an 5-6 Tagen, 4 = jeden Tag). Der POEM-Gesamtscore kann Werte von 0-28 annehmen, wobei ein höherer Wert einem häufigeren Auftreten der Symptome entspricht. Ein Gesamtscore im POEM von 0-2 entspricht einem normalen bis fast normalem Hautbild. Die Einzelitems (Fragen) können Werte von 0-4 annehmen.</p> <p><u>Hauptanalysen:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten mit POEM-Gesamtscore = 0-2 zu Woche 26 • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 5 Punkte im POEM-Gesamtscore gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 1 Punkt in den POEM-Einzelitems Blutungen der Haut, nässende Haut, rissige Haut, schuppige Haut und trockene/ raue Haut gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen • Mittlere Veränderung des POEM-Gesamtscores gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 12, 16 und 26 <p>Die Responderanalysen erfolgten im modifizierten <i>Intention-to-treat-Analysis-Set</i> (mITT), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten und das entsprechende <i>Response</i>-Kriterium zu <i>Baseline</i> erreicht haben. Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response</i>-Imputation ergänzt.</p> <p>Die MMRM-Analyse erfolgte im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben.</p> <p>Die folgenden <i>Estimands</i> wurden betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 1 (Composite Estimand)</i> für Responderanalysen • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-Haenszel Tests. Die Stratifizierung beinhaltete den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares</i> (LS)-Mean-Differenzen und des Hedges'g mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.</p>
	<p>FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; MMRM: <i>Mixed Model Repeated Measures</i>; OR: <i>Odds Ratio</i>; POEM: <i>Patient-Oriented Eczema Measure</i>; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus</p>

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 4-38: Bewertung des Verzerrungspotenzials für POEM 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-treat; POEM: Patient-Oriented Eczema Measure; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

Tabelle 4-39: Beobachtungsdauer für POEM in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	171,9 (41,36)	176,2 (31,02)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-212,0)
Quelle: Table 14.2.10.11 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; POEM: <i>Patient-Oriented Eczema Measures</i> ; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-40: Rücklaufquote für POEM in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 12	327/362 (90,3)	341/365 (93,4)
Woche 16	323/362 (89,2)	337/365 (92,3)
Woche 26	303/362 (83,7)	322/365 (88,2)
Quelle: Table 14.2.10.5 (57) FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS; POEM: <i>Patient-Oriented Eczema Measures</i>		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für den POEM zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

Tabelle 4-41: Ergebnisse für patientenberichtete Symptomatik (POEM) aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit POEM-Gesamtscore = 0-2 zu Woche 26				
106/358 (29,6)	69/363 (19,0)	1,56 [1,19; 2,03] p=0,0011	1,79 [1,27; 2,54] p=0,0010	10,6 [4,4; 16,8] p=0,0009
Anteil der Patienten mit einer Verbesserung um ≥ 5 Punkte gegenüber <i>Baseline</i> im POEM-Gesamtscore zu Woche 26				
271/356 (76,1)	289/363 (79,6)	0,96 [0,88; 1,03] p=0,2519	0,81 [0,57; 1,16] p=0,2516	-3,5 [-9,6; 2,5] p=0,2518
Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für Blutungen der Haut zu Woche 26 (ergänzend)				
247/325 (76,0)	270/327 (82,6)	0,92 [0,85; 1,00] p=0,0384	0,67 [0,46; 0,98] p=0,0383	-6,6 [-12,8; -0,4] p=0,0379
Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für nässende Haut zu Woche 26 (ergänzend)				
205/272 (75,4)	232/283 (82,0)	0,92 [0,84; 1,00] p=0,0579	0,67 [0,45; 1,01] p=0,0571	-6,6 [-13,4; 0,2] p=0,0567
Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für rissige Haut zu Woche 26 (ergänzend)				
234/330 (70,9)	258/349 (73,9)	0,96 [0,87; 1,05] p=0,3651	0,86 [0,61; 1,20] p=0,3644	-3,1 [-9,8; 3,6] p=0,3648
Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für schuppige Haut zu Woche 26 (ergänzend)				
245/340 (72,1)	250/353 (70,8)	1,02 [0,92; 1,12] p=0,7378	1,06 [0,76; 1,47] p=0,7378	1,1 [-5,6; 7,9] p=0,7379
Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für trockene/ raue Haut zu Woche 26 (ergänzend)				
247/358 (69,0)	248/363 (68,3)	1,01 [0,91; 1,11] p=0,8534	1,03 [0,75; 1,41] p=0,8533	0,6 [-6,1; 7,4] p=0,8535

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Quelle: Table 14.2.10.8.1, Table 14.2.10.9.1, Table 14.2.10.10.1 (57)				
Werte nach Therapieabbruch und nach <i>Rescue</i> -Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i> , siehe Abschnitt 4.2.5.2).				
a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i> -Kriterium erreicht haben (mITT).				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).				
IGA: <i>Investigator's Global Assessment</i> ; FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; mg: Milligramm; mITT: modifiziertes <i>Intention-to-treat-Analysis-Set</i> ; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i> ; POEM: <i>Patient-Oriented Eczema Measures</i> ; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus				

Zu Woche 26 zeigte sich für den Anteil der Patienten mit POEM-Gesamtscore = 0-2 ein statistisch signifikanter Unterschied zugunsten von Abrocitinib gegenüber Dupilumab (RR [95 %-KI]: 1,56 [1,19; 2,03]; p-Wert = 0,0011). Unter Abrocitinib wiesen 29,6 % der Patienten einen POEM-Gesamtscore von 0-2 und somit nahezu eine Symptombefreiung auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 19,0 %.

Zu Woche 26 wiesen im Abrocitinib-Arm 76,1 % der Patienten eine Verbesserung um ≥ 5 Punkte im POEM-Gesamtscore gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 79,6 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 0,96 [0,88; 1,03]; p-Wert = 0,2519).

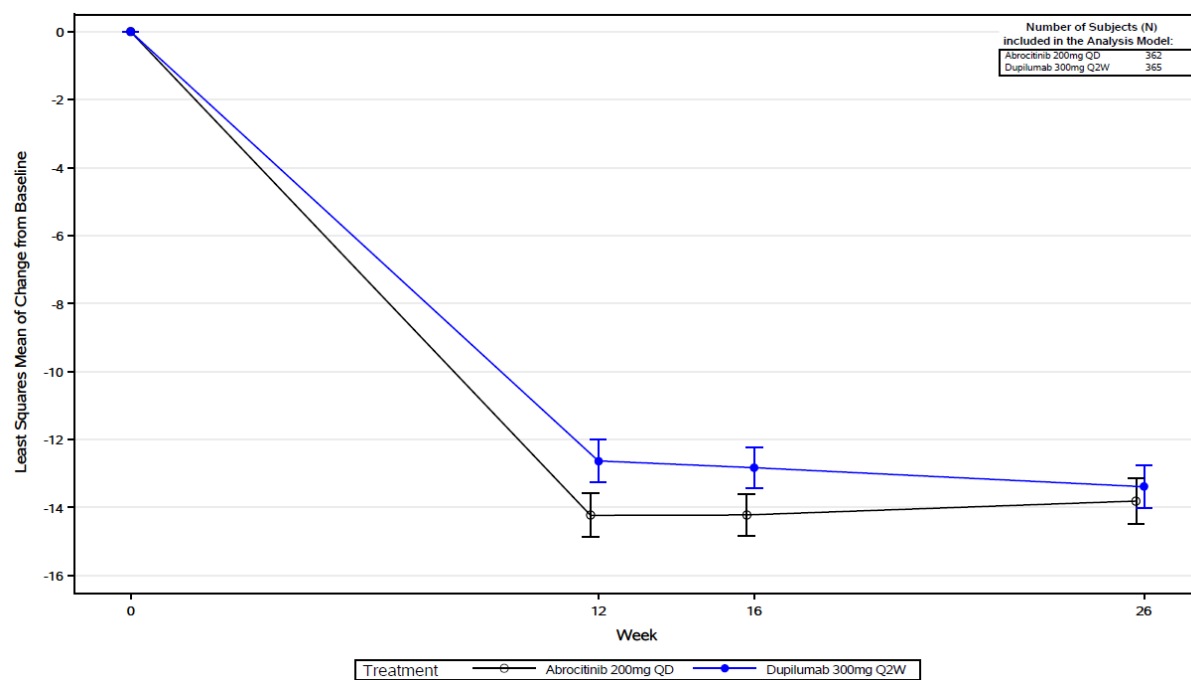
Zu Woche 26 zeigte sich ein statistisch signifikanter Unterschied in der Verbesserung um ≥ 1 Punkt gegenüber dem Ausgangswert für das ergänzend dargestellte POEM-Einzelitem Blutungen der Haut (RR [95 %-KI]: 0,92 [0,85; 1,00]; p-Wert = 0,0384) zuungunsten von Abrocitinib gegenüber Dupilumab. Für die weiteren POEM-Einzelitems nässende Haut, rissige Haut, schuppige Haut und trockene/ raue Haut lagen keine statistisch signifikanten Unterschiede vor.

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

Tabelle 4-42: Ergebnisse für mittlere Veränderung im POEM-Gesamtscore; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu <i>Baseline</i> (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu <i>Baseline</i> (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung im POEM-Gesamtscore								
<i>Baseline</i>	362	20,4 (5,8)	NA	365	20,9 (5,3)	NA	NA	NA
Woche 12	327	6,3 (6,1)	-14,23 (0,33)	341	8,0 (6,2)	-12,63 (0,32)	-1,60 [-2,49; -0,70] p=0,0005	-0,27 [-0,42; -0,12]
Woche 16	323	6,2 (5,8)	-14,22 (0,31)	337	7,8 (5,7)	-12,83 (0,30)	-1,39 [-2,24; -0,54] p=0,0013	-0,25 [-0,40; -0,10]
Woche 26	303	6,5 (6,4)	-13,81 (0,33)	322	7,2 (5,8)	-13,39 (0,33)	-0,42 [-1,34; 0,50] p=0,3684	-0,07 [-0,23; 0,09]
<p>Quelle: Table 14.2.10.1, Table 14.2.10.5 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie wurden als fehlende Werte betrachtet (<i>Hypothetical Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Deskriptive Analyse basierend auf den Patienten im FAS mit <i>Baseline</i>-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).</p> <p>b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem <i>Restricted Maximum Likelihood</i> Ansatz.</p> <p>FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mg: Milligramm; MMRM: <i>Mixed Model Repeated Measures</i>; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit <i>Baseline</i>-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; POEM: <i>Patient-Oriented Eczema Measures</i>; SD: Standardabweichung (<i>Standard Deviation</i>); SE: Standardfehler (<i>Standard Error</i>); vs.: Versus</p>								

In der Analyse der mittleren Veränderung des POEM-Gesamtscores gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 12 und 16 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind nicht als klinisch relevant zu bewerten, da das 95 %-KI der SMD (Hedges'g) innerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	12	16	26
N* Abrocitinib 200 mg	362	327	323	303
N* Dupilumab 300 mg	365	341	337	322

Abbildung 4-6: Verlaufskurve für die mittlere Veränderung des POEM-Gesamtscores von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel
mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren Veränderungen im POEM zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-6).

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 die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.1.1.5 Juckreiz – RCT

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

Tabelle 4-43: Operationalisierung des Endpunkts Juckreiz

Studie	Operationalisierung
JADE DARE	<p>Die <i>Peak Pruritus Numerical Rating Scale</i> (NRS) ist eine patientenberichtete Skala, in der die maximale Stärke des Juckreizes innerhalb der letzten 24 Stunden bewertet wird (0: kein Juckreiz bis 10: schlimmster vorstellbarer Juckreiz).</p> <p><u>Hauptanalyse:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 4 Punkte gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Mittlere Veränderung gegenüber dem <i>Baseline</i>-Wert anhand eines <i>Mixed Model Repeated Measures</i> (MMRM) zu Woche 2, 4, 8, 12, 16, 20 und 26 <p>Auf der <i>Scoring Atopic Dermatitis</i> (SCORAD) Visuellen Analogskala (VAS) für Juckreiz beurteilen die Patienten den Juckreiz in den letzten 3 Tage bzw. Nächten auf einer Skala von 0-10 (0: kein Juckreiz bis 10: schlimmster vorstellbarer Juckreiz).</p> <p><u>Hauptanalyse:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 2 Punkte gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Mittlere Veränderung gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 2, 4, 8, 12, 16, 20 und 26 <p>Im ersten Item (Frage) des <i>Patient-Oriented Eczema Measures</i> (POEM) („An wie vielen Tagen in der letzten Woche juckte Ihre Haut aufgrund Ihres Ekzems?“) berichtet der Patient die Häufigkeit des Auftretens von Juckreiz auf einer 5-Punkte-Skala (0 = an keinem Tag, 1 = an 1-2 Tagen, 2 = an 3-4 Tagen, 3 = an 5-6 Tagen, 4 = jeden Tag). Das Item kann einen Wert von 0-4 annehmen.</p> <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 1 Punkt gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen (ergänzend) <p>Die Responderanalysen erfolgten im modifizierten <i>Intention-to-treat-Analysis-Set</i> (mITT), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten und das entsprechende <i>Response</i>-Kriterium zu <i>Baseline</i> erreicht haben. Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response</i>-Imputation ergänzt.</p> <p>Die MMRM-Analysen erfolgten im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben.</p> <p>Die folgenden <i>Estimands</i> wurden betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 1 (Composite Estimand)</i> für Responderanalysen • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-HaenszelTests. Die Stratifizierung beinhaltete den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p>

Studie	Operationalisierung
	Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i> -Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares (LS)-Mean</i> -Differenzen und des Hedges'g mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.
IGA: <i>Investigator's Global Assessment</i> ; FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i> ; MMRM: <i>Mixed Model Repeated Measures</i> ; NRS: <i>Numerical Rating Scale</i> ; OR: <i>Odds Ratio</i> ; POEM: <i>Patient-Oriented Eczema Measure</i> ; RD: Risikodifferenz; RR: Relatives Risiko; SCORAD: <i>Scoring Atopic Dermatitis</i> ; VAS: Visuelle Analogskala; vs.: Versus	

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 4-44: Bewertung des Verzerrungspotenzials für Juckreiz 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

4.3.1.3.1.1.5.1 Peak Pruritus NRS

Tabelle 4-45: Beobachtungsdauer für *Peak Pruritus* NRS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	176,9 (35,03)	180,5 (24,69)
Median (min-max)	183,0 (7,0-282,0)	183,0 (14,0-221,0)
Quelle: Table 14.2.4.5.8 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; NRS: <i>Numerical Rating Scale</i> ; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-46: Rücklaufquote für *Peak Pruritus* NRS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 2	354/362 (97,8)	355/365 (97,3)
Woche 4	349/362 (96,4)	351/365 (96,2)
Woche 8	345/362 (95,3)	351/365 (96,2)
Woche 12	333/362 (92,0)	347/365 (95,1)
Woche 16	331/362 (91,4)	342/365 (93,7)
Woche 20	326/362 (90,1)	338/365 (92,6)
Woche 26	313/362 (86,5)	328/365 (89,9)
Quelle: Table 14.2.4.1.1 (57) FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS; NRS: <i>Numerical Rating Scale</i>		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für die *Peak Pruritus* NRS zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 86 %.

Tabelle 4-47: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber *Baseline* in der *Peak Pruritus* NRS aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber <i>Baseline</i> in der <i>Peak Pruritus</i> NRS zu Woche 26				
241/357 (67,5)	229/364 (62,9)	1,07 [0,96; 1,19] p=0,1978	1,22 [0,90; 1,67] p=0,1977	4,6 [-2,4; 11,5] p=0,1981
<p>Quelle: Table 14.2.4.5.6.1 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i>-Kriterium erreicht haben (mITT).</p> <p>b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; mg: Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; NRS: <i>Numerical Rating Scale</i>; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus</p>				

Zu Woche 26 wiesen im Abrocitinib-Arm 67,5 % der Patienten eine Verbesserung des *Peak Pruritus* NRS-Scores um ≥ 4 Punkte gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 62,9 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 1,07 [0,96; 1,19]; p-Wert = 0,1978).

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

Tabelle 4-48: Ergebnisse für mittlere Veränderung der *Peak Pruritus* NRS; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung der <i>Peak Pruritus</i> NRS								
Baseline	362	7,4 (1,6)	NA	365	7,4 (1,6)	NA	NA	NA
Woche 2	354	3,8 (2,3)	-3,60 (0,11)	355	5,0 (2,2)	-2,37 (0,11)	-1,23 [-1,54; -0,92] p<0,0001	-0,58 [-0,73; -0,43]
Woche 4	349	3,2 (2,3)	-4,24 (0,12)	351	4,2 (2,2)	-3,23 (0,12)	-1,01 [-1,33; -0,68] p<0,0001	-0,46 [-0,61; -0,31]
Woche 8	345	2,8 (2,3)	-4,63 (0,12)	351	3,6 (2,4)	-3,81 (0,12)	-0,81 [-1,15; -0,48] p<0,0001	-0,36 [-0,51; -0,21]
Woche 12	333	2,6 (2,3)	-4,74 (0,12)	347	3,1 (2,2)	-4,24 (0,12)	-0,50 [-0,82; -0,17] p=0,0029	-0,23 [-0,38; -0,08]
Woche 16	331	2,6 (2,3)	-4,76 (0,12)	342	3,0 (2,2)	-4,33 (0,12)	-0,43 [-0,76; -0,10] p=0,0109	-0,20 [-0,35; -0,05]
Woche 20	326	2,5 (2,2)	-4,83 (0,12)	338	2,8 (2,1)	-4,50 (0,12)	-0,32 [-0,65; 0,00] p=0,0497	-0,15 [-0,30; 0,00]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Woche 26	313	2,3 (2,2)	-4,92 (0,12)	328	2,7 (2,1)	-4,65 (0,12)	-0,27 [-0,60; 0,06] p=0,1068	-0,13 [-0,28; 0,03]

Quelle: Table 14.2.4.1.1, Table 14.2.4.4.1 (57)

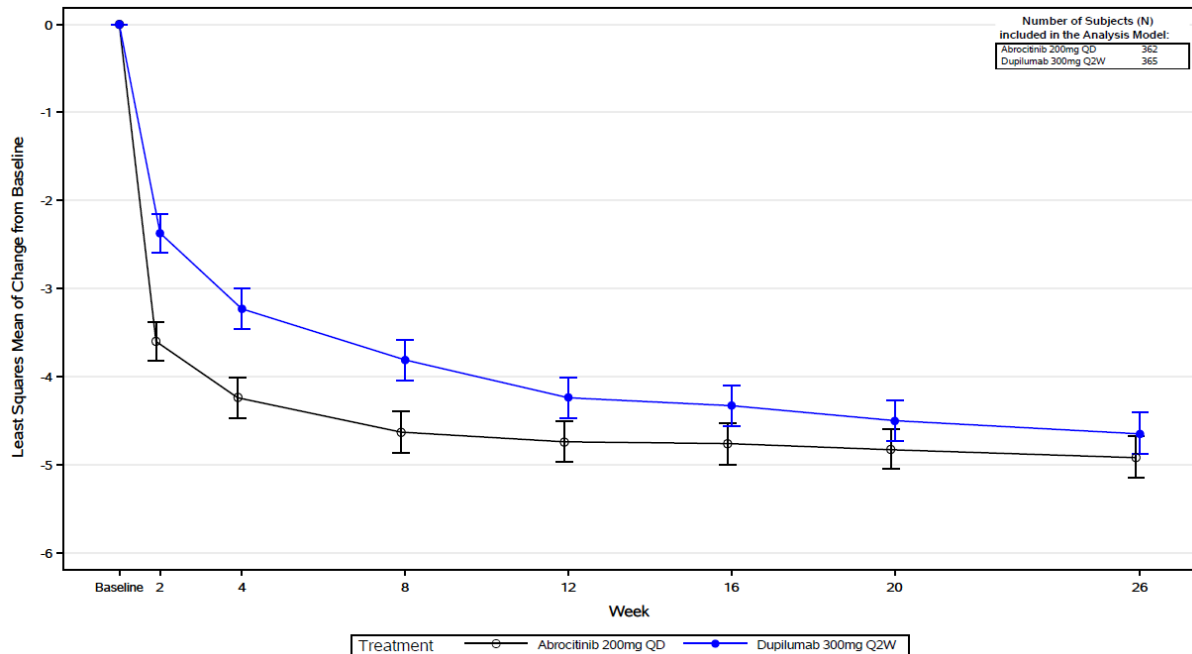
Werte nach Therapieabbruch und nach *Rescue*-Therapie wurden als fehlende Werte betrachtet (*Hypothetical Estimand*, siehe Abschnitt 4.2.5.2).

a: Deskriptive Analyse basierend auf den Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).

b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie *Baseline*-Wert + Krankheitsschwere zu *Baseline* als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem *Restricted Maximum Likelihood* Ansatz.

FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (*Least Squares*); mg: Milligramm; MMRM: *Mixed Model Repeated Measures*; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; NRS: *Numerical Rating Scale*; SD: Standardabweichung (*Standard Deviation*); SE: Standardfehler (*Standard Error*); vs.: Versus

In der Analyse der mittleren Veränderung der *Peak Pruritus* NRS gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 2, 4, 8, 12, 16 und 20 ein statistisch signifikanter Unterschied der *LS-Mean-Differenz* zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind zu Woche 2, 4 und 8 als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges'g) außerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	2	4	8	12	16	20	26
N* Abrocitinib 200 mg	362	354	349	345	333	331	326	313
N* Dupilumab 300 mg	365	355	351	351	347	342	338	328

Abbildung 4-7: Verlaufskurve für die mittlere Veränderung der *Peak Pruritus* NRS von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel
mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren Veränderungen in der *Peak Pruritus* NRS zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-7).

4.3.1.3.1.1.5.2 SCORAD VAS für Juckreiz

Tabelle 4-49: Beobachtungsdauer für SCORAD VAS für Juckreiz in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	175,1 (37,81)	179,2 (26,50)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-221,0)
Quelle: Table 14.2.13.9.12 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SCORAD: <i>Scoring Atopic Dermatitis</i> ; SD: Standardabweichung (<i>Standard Deviation</i>); VAS: Visuelle Analogskala		

Tabelle 4-50: Rücklaufquote für SCORAD VAS für Juckreiz in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
Baseline	362/362 (100,0)	365/365 (100,0)
Woche 2	346/362 (95,6)	345/365 (94,5)
Woche 4	337/362 (93,1)	351/365 (96,2)
Woche 8	333/362 (92,0)	345/365 (94,5)
Woche 12	326/362 (90,1)	341/365 (93,4)
Woche 16	324/362 (89,5)	336/365 (92,1)
Woche 20	319/362 (88,1)	333/365 (91,2)
Woche 26	300/362 (82,9)	323/365 (88,5)
Quelle: Table 14.2.13.9.4.3 (57) FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS; SCORAD: <i>Scoring Atopic Dermatitis</i> ; VAS: Visuelle Analogskala		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für die SCORAD VAS für Juckreiz zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 82 %.

Tabelle 4-51: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 2 Punkte gegenüber *Baseline* in der SCORAD VAS für Juckreiz aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit einer Verbesserung um ≥ 2 Punkte gegenüber <i>Baseline</i> in der SCORAD VAS für Juckreiz zu Woche 26				
273/362 (75,4)	294/365 (80,5)	0,94 [0,87; 1,01] p=0,0926	0,74 [0,52; 1,05] p=0,0922	-5,2 [-11,2; 0,8] p=0,0919
<p>Quelle: Table 14.2.13.9.8.1 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i>-Kriterium erreicht haben (mITT).</p> <p>b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; mg: Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; SCORAD: <i>Scoring Atopic Dermatitis</i>; VAS: Visuelle Analogskala; vs.: Versus</p>				

Zu Woche 26 wiesen im Abrocitinib-Arm 75,4 % der Patienten eine Verbesserung in der SCORAD VAS für Juckreiz um ≥ 2 Punkte gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 80,5 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 0,94 [0,87; 1,01]; p-Wert = 0,0926).

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

Tabelle 4-52: Ergebnisse für mittlere Veränderung in der SCORAD VAS für Juckreiz; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung in der SCORAD VAS für Juckreiz								
Baseline	362	7,7 (1,6)	NA	365	7,7 (1,6)	NA	NA	NA
Woche 2	346	3,5 (2,3)	-4,17 (0,12)	345	4,9 (2,4)	-2,79 (0,12)	-1,37 [-1,71; -1,03] p<0,0001	-0,60 [-0,75; -0,45]
Woche 4	337	2,6 (2,1)	-5,11 (0,12)	351	3,8 (2,3)	-3,88 (0,12)	-1,23 [-1,56; -0,90] p<0,0001	-0,56 [-0,71; -0,41]
Woche 8	333	2,3 (2,2)	-5,37 (0,13)	345	3,2 (2,4)	-4,50 (0,12)	-0,87 [-1,22; -0,52] p<0,0001	-0,38 [-0,53; -0,23]
Woche 12	326	2,1 (2,2)	-5,52 (0,12)	341	2,7 (2,3)	-4,97 (0,12)	-0,55 [-0,88; -0,21] p=0,0014	-0,25 [-0,40; -0,10]
Woche 16	324	2,1 (2,3)	-5,51 (0,12)	336	2,5 (2,1)	-5,22 (0,12)	-0,29 [-0,63; 0,05] p=0,0938	-0,13 [-0,28; 0,02]
Woche 20	319	1,9 (2,1)	-5,71 (0,12)	333	2,2 (2,1)	-5,42 (0,12)	-0,29 [-0,62; 0,04] p=0,0882	-0,13 [-0,29; 0,02]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Woche 26	300	2,0 (2,2)	-5,58 (0,13)	323	2,2 (2,2)	-5,49 (0,12)	-0,09 [-0,43; 0,25] p=0,6106	-0,04 [-0,20; 0,12]

Quelle: Table 14.2.13.9.4.3, Table 14.2.13.9.7.1 (57)

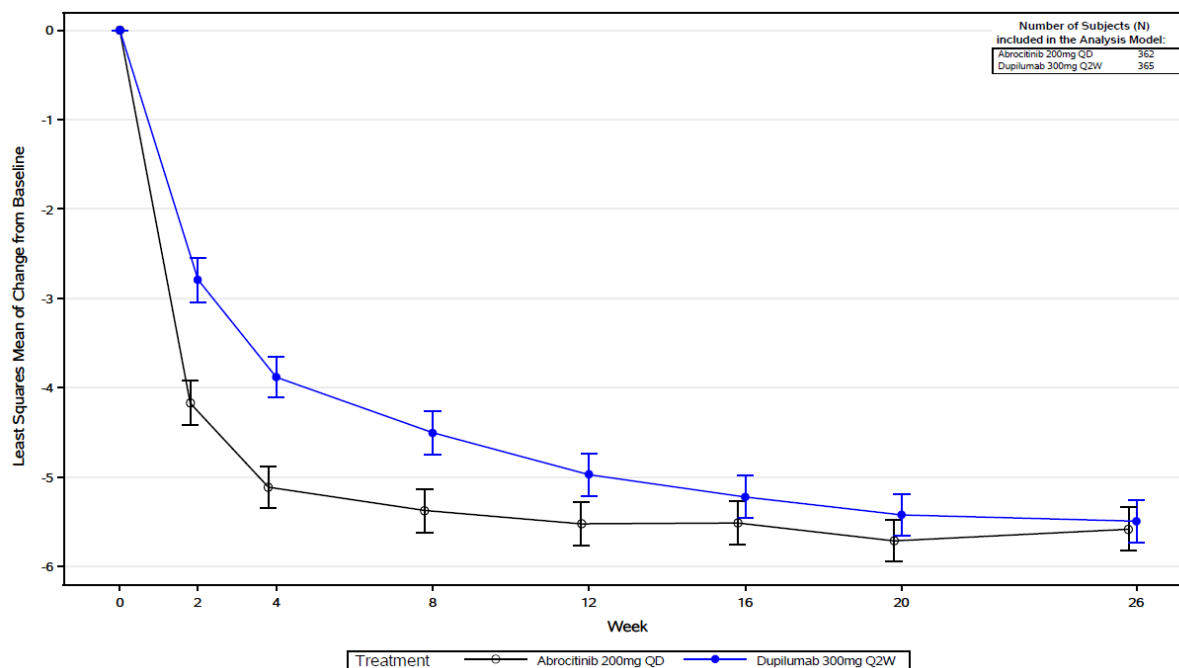
Werte nach Therapieabbruch und nach *Rescue*-Therapie wurden als fehlende Werte betrachtet (*Hypothetical Estimand*, siehe Abschnitt 4.2.5.2).

a: Deskriptive Analyse basierend auf den Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).

b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie *Baseline*-Wert + Krankheitsschwere zu *Baseline* als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem *Restricted Maximum Likelihood* Ansatz.

FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (*Least Squares*); mg: Milligramm; MMRM: *Mixed Model Repeated Measures*; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SCORAD: *Scoring Atopic Dermatitis*; SD: Standardabweichung (*Standard Deviation*); SE: Standardfehler (*Standard Error*); VAS: Visuelle Analogskala

In der Analyse der mittleren Veränderung des SCORAD VAS für Juckreiz gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 2, 4, 8 und 12 ein statistisch signifikanter Unterschied der *LS-Mean-Differenz* zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind zu Woche 2, 4 und 8 als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges'g) außerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	2	4	8	12	16	20	26
N* Abrocitinib 200 mg	362	346	337	333	326	324	319	300
N* Dupilumab 300 mg	365	345	351	345	341	336	333	323

Abbildung 4-8: Verlaufskurve für die mittlere Veränderung der SCORAD VAS für Juckreiz von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren Veränderungen der SCORAD VAS für Juckreiz zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-8).

4.3.1.3.1.1.5.3 POEM-Einzelitem für Juckreiz

Tabelle 4-53: Beobachtungsdauer für POEM-Einzelitem für Juckreiz in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	171,9 (41,36)	176,2 (31,02)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-212,0)
Quelle: Table 14.2.10.11 (57) max: Maximal; min: Minimal; mg: Milligramm; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; POEM: <i>Patient-Oriented Eczema Measures</i> ; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-54: Rücklaufquote für POEM-Einzelitem für Juckreiz in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	361/362 (99,7)	365/365 (100,0)
Woche 12	328/362 (90,6)	340/365 (93,2)
Woche 16	322/362 (89,0)	337/365 (92,3)
Woche 26	303/362 (83,7)	322/365 (88,2)
Quelle: Table 14.2.10.5 (57) FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS; POEM: <i>Patient-Oriented Eczema Measures</i>		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für das POEM-Einzelitem für Juckreiz zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

Tabelle 4-55: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber *Baseline* im POEM-Einzelitem für Juckreiz aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für Juckreiz zu Woche 26 (ergänzend)				
241/358 (67,3)	249/364 (68,4)	0,98 [0,89; 1,09] p=0,7524	0,95 [0,70; 1,30] p=0,7524	-1,1 [-7,9; 5,7] p=0,7527
Quelle: Table 14.2.10.10.1 (57)				
Werte nach Therapieabbruch und nach <i>Rescue</i> -Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i> , siehe Abschnitt 4.2.5.2).				
a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i> -Kriterium erreicht haben (mITT).				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).				
IGA: <i>Investigator's Global Assessment</i> ; FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; mg: Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i> ; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i> ; POEM: <i>Patient-Oriented Eczema Measures</i> ; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus				

Zu Woche 26 wiesen im Abrocitinib-Arm 67,3 % der Patienten eine Verbesserung des POEM-Einzelitems für Juckreiz um ≥ 1 Punkt gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 68,4 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI: 0,98 [0,89; 1,09]; p-Wert = 0,7524).

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 die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.1.1.6 Schlafstörungen – RCT

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

Tabelle 4-56: Operationalisierung des Endpunkts Schlafstörungen

Studie	Operationalisierung
JADE DARE	<p>Die <i>Medical Outcomes Study (MOS) Sleep Scale</i> ist ein vom Patienten selbst dokumentiertes Instrument zur Beurteilung der Schlafqualität in den letzten 4 Wochen. Sie umfasst insgesamt 12 Items in den 6 Domänen Schlafstörung, Schlafadäquanz, Schlafquantität, Somnolenz, Schnarchen, Kurzatmigkeit oder Kopfschmerzen. Die Items können zu 2 Indices zusammengefasst werden. Der <i>MOS Sleep Scale Score Index-1</i> umfasst 6 Items (Schlafprobleme ohne Schlafquantität) und der <i>MOS Sleep Scale Score Index-2</i> umfasst 9 Items (Gesamtübersicht der Schlafprobleme). Der Gesamtscore der <i>MOS Sleep Scale Score Indices</i> kann jeweils Werte von 0-100 annehmen.</p> <p><u>Hauptanalysen:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 15 Punkte im <i>MOS Sleep Scale Score Index-1</i> gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 15 Punkte im <i>MOS Sleep Scale Score Index-2</i> gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen <p><u>Sensitivitätsanalysen:</u></p> <ul style="list-style-type: none"> • Mittlere Veränderung des <i>MOS Sleep Scale Score Index-1</i> gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 12, 16 und 26 • Mittlere Veränderung des <i>MOS Sleep Scale Score Index-2</i> gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 12, 16 und 26 <p>Auf der <i>Scoring Atopic Dermatitis (SCORAD)</i> Visuellen Analogskala (VAS) für Schlafstörungen beurteilen die Patienten die Schlafstörungen der letzten 3 Tage bzw. Nächte auf einer Skala von 0-10 (0: keine Schlafstörungen; 10: schlimmste vorstellbare Schlafstörungen).</p> <p><u>Hauptanalyse:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 2 Punkte in der <i>SCORAD VAS</i> für Schlafstörung gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Mittlere Veränderung der <i>SCORAD VAS</i> für Schlafstörungen gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 2, 4, 8, 12, 16, 20 und 26 <p>Im zweiten Item (Frage) des <i>Patient-Oriented Eczema Measures (POEM)</i> („In wie vielen Nächten in der letzten Woche war Ihre Nachtruhe aufgrund Ihres Ekzems gestört?“) berichtet der Patient die Häufigkeit des Auftretens von Schlafstörungen auf einer 5-Punkte-Skala (0 = in keiner Nacht, 1 = in 1-2 Nächten, 2 = in 3-4 Nächten, 3 = in 5-6 Nächten, 4 = jede Nacht). Das Item kann einen Wert von 0-4 annehmen.</p> <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 1 Punkt im <i>POEM-Einzelitem</i> Schlafstörungen gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen (ergänzend) <p>Die Responderanalysen erfolgten im modifizierten <i>Intention-to-treat-Analysis-Set (mITT)</i>, welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten und das entsprechende <i>Response</i>-Kriterium zu <i>Baseline</i> erreicht haben. Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt.</p> <p>Die MMRM-Analysen erfolgten im <i>Full-Analysis-Set (FAS)</i>, welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben.</p>

Studie	Operationalisierung
	<p>Die folgenden <i>Estimands</i> wurden betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 1 (Composite Estimand)</i> für Responderanalysen • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-Haenszel-Tests. Die Stratifizierung beinhaltete den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares (LS)-Mean-Differenzen</i> und des Hedges'g mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.</p>
	<p>IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; MMRM: <i>Mixed Model Repeated Measures</i>; MOS: <i>Medical Outcomes Study</i>; OR: <i>Odds Ratio</i>; POEM: <i>Patient-Oriented Eczema Measures</i>; RD: Risikodifferenz; RR: Relatives Risiko; SCORAD: <i>Scoring Atopic Dermatitis</i>; VAS: Visuelle Analogskala; vs.: Versus</p>

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

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

Tabelle 4-57: Bewertung des Verzerrungspotenzials für Schlafstörungen 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung

begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

4.3.1.3.1.1.6.1 MOS Sleep Scale

Tabelle 4-58: Beobachtungsdauer für MOS Sleep Scale in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	170,9 (43,25)	176,2 (31,02)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-212,0)
Quelle: Table 14.2.11.9 (57) max: Maximal; min: Minimal; mg: Milligramm; MOS: <i>Medical Outcomes Study</i> ; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-59: Rücklaufquote für MOS Sleep Scale Score Index-1 und MOS Sleep Scale Score Index-2 in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
MOS Sleep Scale Score Index-1		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 12	327/362 (90,3)	341/365 (93,4)
Woche 16	321/362 (88,7)	337/365 (92,3)
Woche 26	301/362 (83,1)	322/365 (88,2)
MOS Sleep Scale Score Index-2		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 12	327/362 (90,3)	341/365 (93,4)
Woche 16	321/362 (88,7)	337/365 (92,3)
Woche 26	301/362 (83,1)	322/365 (88,2)
Quelle: Table 14.2.11.3 (57) FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; MOS: <i>Medical Outcomes Study</i> ; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für den MOS *Sleep Scale Score* Index-1 und Index-2 zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

Tabelle 4-60: Ergebnisse für MOS *Sleep Scale* aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit einer Verbesserung um ≥ 15 Punkte gegenüber <i>Baseline</i> im MOS <i>Sleep Scale Score</i> Index-1 zu Woche 26				
131/362 (36,2)	117/363 (32,2)	1,12 [0,92; 1,37] p=0,2642	1,19 [0,88; 1,63] p=0,2639	3,9 [-2,9; 10,7] p=0,2637
Anteil der Patienten mit einer Verbesserung um ≥ 15 Punkte gegenüber <i>Baseline</i> im MOS <i>Sleep Scale Score</i> Index-2 zu Woche 26				
139/362 (38,4)	140/364 (38,5)	1,00 [0,83; 1,20] p=0,9724	0,99 [0,74; 1,34] p=0,9724	-0,1 [-7,1; 6,9] p=0,9724
Quelle: Table 14.2.11.7.1, Table 4.2.11.8.1 (57)				
Werte nach Therapieabbruch und nach <i>Rescue</i> -Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i> , siehe Abschnitt 4.2.5.2).				
a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i> -Kriterium erreicht haben (mITT).				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).				
IGA: <i>Investigator's Global Assessment</i> ; FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; mg: Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i> ; MOS: <i>Medical Outcomes Study</i> ; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i> ; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus				

Zu Woche 26 wiesen im Abrocitinib-Arm 36,2 % der Patienten eine Verbesserung des MOS *Sleep Scale Scores* Index-1 um ≥ 15 Punkte gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 32,2 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 1,12 [0,92; 1,37]; p-Wert = 0,2642).

Zu Woche 26 wiesen im Abrocitinib-Arm 38,4 % der Patienten eine Verbesserung des MOS *Sleep Scale Scores* Index-2 um ≥ 15 Punkte gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 38,5 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 1,00 [0,83; 1,20]; p-Wert = 0,9724).

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

Tabelle 4-61: Ergebnisse für MOS *Sleep Scale*; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung im MOS <i>Sleep Scale</i> Score Index-1								
Baseline	362	45,1 (14,6)	NA	365	45,2 (14,7)	NA	NA	NA
Woche 12	327	32,9 (12,4)	-12,07 (0,66)	341	34,1 (13,3)	-10,91 (0,65)	-1,16 [-2,97; 0,65] p=0,2100	-0,10 [-0,25; 0,05]
Woche 16	321	31,6 (12,0)	-13,25 (0,64)	337	34,2 (12,8)	-11,00 (0,62)	-2,26 [-4,00; -0,51] p=0,0114	-0,20 [-0,35; -0,04]
Woche 26	301	31,9 (12,4)	-12,93 (0,63)	322	33,0 (12,0)	-12,11 (0,61)	-0,82 [-2,53; 0,90] p=0,3489	-0,08 [-0,23; 0,08]
Mittlere Veränderung im MOS <i>Sleep Scale</i> Score Index-2								
Baseline	362	49,0 (15,3)	NA	365	49,1 (15,5)	NA	NA	NA
Woche 12	327	34,5 (12,7)	-14,41 (0,67)	341	36,6 (14,1)	-12,22 (0,66)	-2,19 [-4,04; -0,34] p=0,0205	-0,18 [-0,33; -0,03]
Woche 16	321	33,1 (12,2)	-15,71 (0,66)	337	36,2 (13,6)	-12,76 (0,64)	-2,95 [-4,75; -1,14] p=0,0014	-0,25 [-0,40; -0,10]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Woche 26	301	33,4 (13,1)	-15,38 (0,65)	322	34,9 (12,6)	-13,98 (0,64)	-1,40 [-3,19; 0,39] p=0,1259	-0,12 [-0,28; 0,03]

Quelle: Table 14.2.11.1, Table 14.2.11.3 (57)

Werte nach Therapieabbruch und nach *Rescue*-Therapie wurden als fehlende Werte betrachtet (*Hypothetical Estimand*, siehe Abschnitt 4.2.5.2).

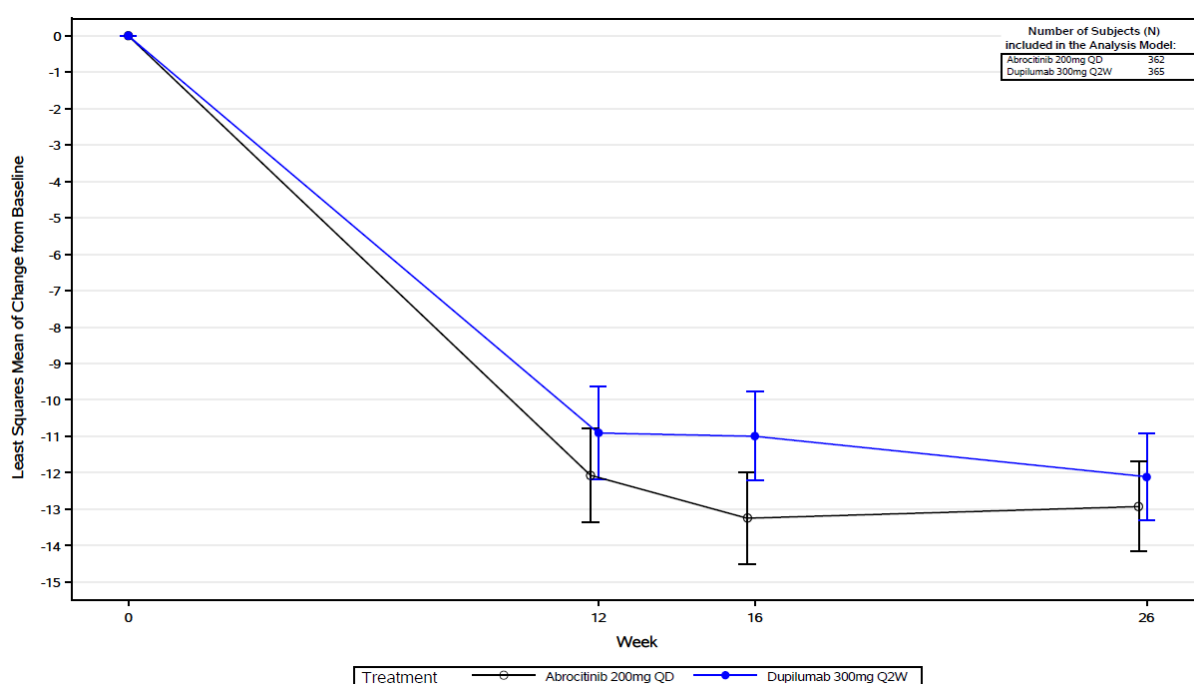
a: Deskriptive Analyse basierend auf den Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).

b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie *Baseline*-Wert + Krankheitsschwere zu *Baseline* als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem *Restricted Maximum Likelihood* Ansatz.

FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (*Least Squares*); mg: Milligramm; MMRM: *Mixed Model Repeated Measures*; MOS: *Medical Outcomes Study*; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SD: Standardabweichung (*Standard Deviation*); SE: Standardfehler (*Standard Error*); vs.: Versus

In der Analyse der mittleren Veränderung im MOS *Sleep Scale Score* Index-1 gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 16 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind nicht als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges'g) innerhalb des Irrelevanzbereichs von 0,2 liegt.

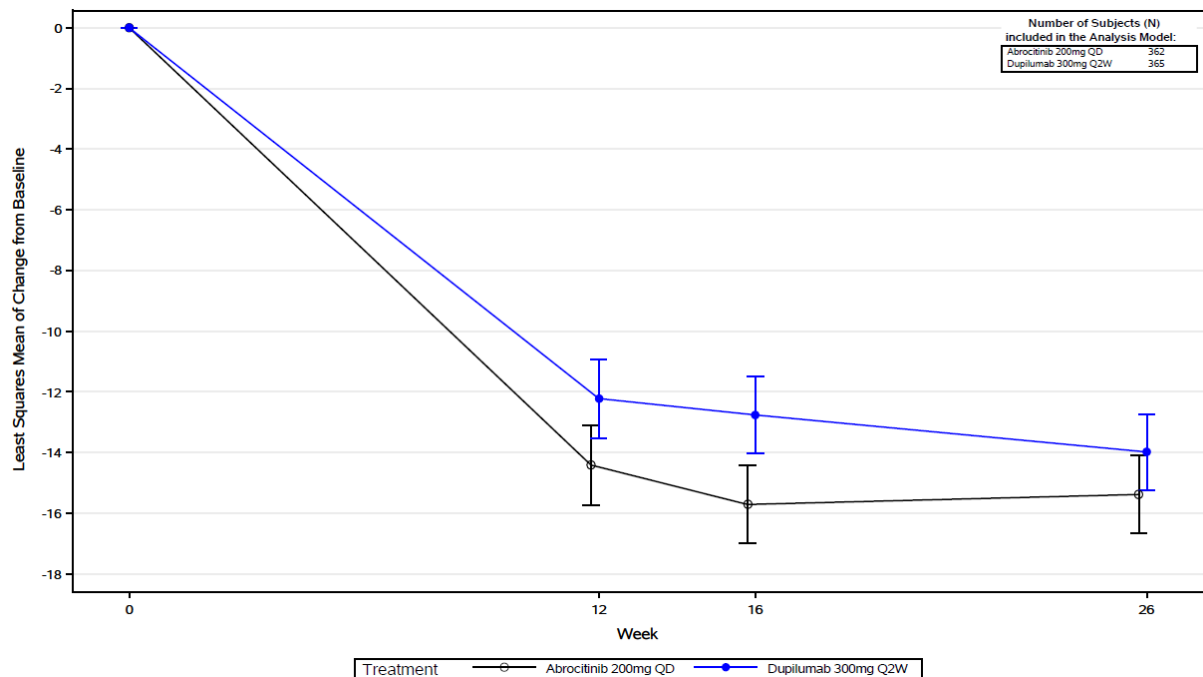
In der Analyse der mittleren Veränderung im MOS *Sleep Scale Score* Index-2 gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 12 und 16 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind nicht als klinisch relevant zu bewerten, da das gesamte 95 %-KI der standardisierten Mittelwertsdifferenz (Hedges'g) innerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	12	16	26
N* Abrocitinib 200 mg	362	327	321	301
N* Dupilumab 300 mg	365	341	337	322

Abbildung 4-9: Verlaufskurve für die mittlere Veränderung des MOS *Sleep Scale* Index-1 von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel
mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die VerlaufsDarstellungen der mittleren Veränderungen im MOS *Sleep Scale Score* Index-1 zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-9).



Woche	0	12	16	26
N* Abrocitinib 200 mg	362	327	321	301
N* Dupilumab 300 mg	365	341	337	322

Abbildung 4-10: Verlaufskurve für die mittlere Veränderung des MOS *Sleep Scale* Index-2 von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren Veränderungen im MOS *Sleep Scale* Score Index-2 zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-10).

4.3.1.3.1.1.6.2 SCORAD VAS für Schlafstörungen

Tabelle 4-62: Beobachtungsdauer für SCORAD VAS für Schlafstörungen in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	175,1 (37,81)	179,2 (26,50)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-221,0)
Quelle: Table 14.2.13.9.11 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SCORAD: <i>Scoring Atopic Dermatitis</i> ; SD: Standardabweichung (<i>Standard Deviation</i>); VAS: Visuelle Analogskala		

Tabelle 4-63: Rücklaufquote für SCORAD VAS für Schlafstörungen in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 2	346/362 (95,6)	345/365 (94,5)
Woche 4	337/362 (93,1)	351/365 (96,2)
Woche 8	333/362 (92,0)	345/365 (94,5)
Woche 12	327/362 (90,3)	341/365 (93,4)
Woche 16	324/362 (89,5)	336/365 (92,1)
Woche 20	319/362 (88,1)	333/365 (91,2)
Woche 26	300/362 (82,9)	323/365 (88,5)
Quelle: Table 14.2.13.9.4.2 (57) mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS; SCORAD: <i>Scoring Atopic Dermatitis</i> ; VAS: Visuelle Analogskala		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für die SCORAD VAS für Schlafstörungen zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 82 %.

Tabelle 4-64: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 2 Punkte gegenüber *Baseline* in der SCORAD VAS für Schlafstörungen aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit einer Verbesserung um ≥ 2 Punkte gegenüber <i>Baseline</i> in der SCORAD VAS für Schlafstörungen zu Woche 26				
253/329 (76,9)	265/333 (79,6)	0,97 [0,89; 1,05] p=0,3960	0,85 [0,59; 1,23] p=0,3957	-2,7 [-9,0; 3,5] p=0,3960
<p>Quelle: Table 14.2.13.9.3.1.1 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i>-Kriterium erreicht haben (mITT).</p> <p>b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; mg: Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; SCORAD: <i>Scoring Atopic Dermatitis</i>; VAS: Visuelle Analogskala; vs.: Versus</p>				

Zu Woche 26 wiesen im Abrocitinib-Arm 76,9 % der Patienten eine Verbesserung in der SCORAD VAS für Schlafstörungen um ≥ 2 Punkte gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 79,6 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 0,97 [0,89; 1,05]; p-Wert = 0,3960).

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

Tabelle 4-65: Ergebnisse für mittlere Veränderung in der SCORAD VAS für Schlafstörungen; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung in der SCORAD VAS für Schlafstörungen								
Baseline	362	6,1 (2,7)	NA	365	6,0 (2,7)	NA	NA	NA
Woche 2	346	2,4 (2,3)	-3,68 (0,13)	345	3,5 (2,7)	-2,56 (0,13)	-1,12 [-1,47; -0,77] p<0,0001	-0,48 [-0,63; -0,32]
Woche 4	337	1,6 (1,9)	-4,43 (0,11)	351	2,5 (2,4)	-3,49 (0,11)	-0,93 [-1,25; -0,62] p<0,0001	-0,44 [-0,60; -0,29]
Woche 8	333	1,4 (2,0)	-4,65 (0,11)	345	1,9 (2,2)	-4,14 (0,11)	-0,51 [-0,83; -0,20] p=0,0014	-0,25 [-0,40; -0,10]
Woche 12	327	1,3 (2,0)	-4,70 (0,11)	341	1,6 (2,2)	-4,36 (0,11)	-0,35 [-0,66; -0,03] p=0,0295	-0,17 [-0,32; -0,02]
Woche 16	324	1,2 (1,9)	-4,75 (0,11)	336	1,4 (2,0)	-4,57 (0,11)	-0,18 [-0,48; 0,12] p=0,2389	-0,09 [-0,24; 0,06]
Woche 20	319	1,1 (1,8)	-4,87 (0,11)	333	1,3 (2,1)	-4,61 (0,11)	-0,26 [-0,56; 0,03] p=0,0825	-0,14 [-0,29; 0,02]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Woche 26	300	1,1 (1,9)	-4,83 (0,11)	323	1,3 (1,9)	-4,65 (0,11)	-0,18 [-0,48; 0,11] p=0,2268	-0,10 [-0,25; 0,06]

Quelle: Table 14.2.13.3, Table 14.2.13.9.4.2 (57)

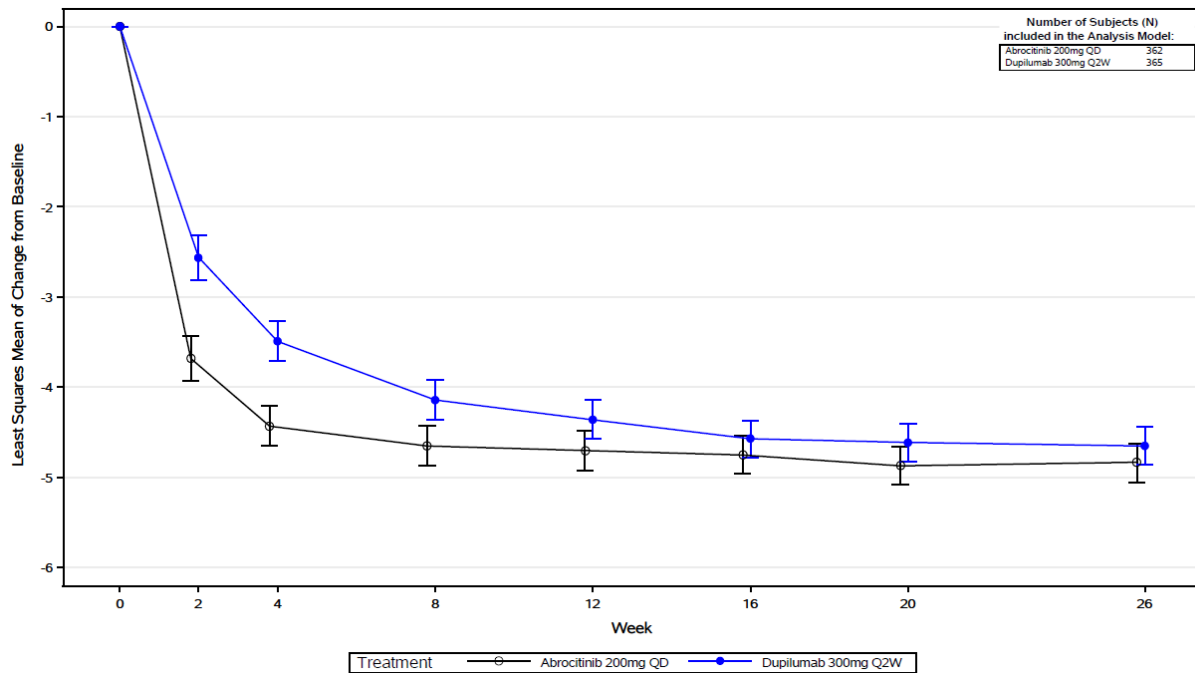
Werte nach Therapieabbruch und nach *Rescue*-Therapie wurden als fehlende Werte betrachtet (*Hypothetical Estimand*, siehe Abschnitt 4.2.5.2).

a: Deskriptive Analyse basierend auf den Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).

b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie *Baseline*-Wert + Krankheitsschwere zu *Baseline* als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem *Restricted Maximum Likelihood* Ansatz.

FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (*Least Squares*); mg: Milligramm; MMRM: *Mixed Model Repeated Measures*; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SCORAD: *Scoring Atopic Dermatitis*; SD: Standardabweichung (*Standard Deviation*); SE: Standardfehler (*Standard Error*); VAS: Visuelle Analogskala; vs.: Versus

In der Analyse der mittleren Veränderung des SCORAD VAS für Schlafstörungen gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 2, 4, 8 und 12 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Die beobachteten Behandlungsunterschiede sind zu Woche 2 und 4 als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges'g) außerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	2	4	8	12	16	20	26
N* Abrocitinib 200 mg	362	346	337	333	327	324	319	300
N* Dupilumab 300 mg	365	345	351	345	341	336	333	323

Abbildung 4-11: Verlaufskurve für die mittlere Veränderung der SCORAD VAS für Schlafstörungen von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren Veränderungen der SCORAD VAS für Schlafstörungen zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-11).

4.3.1.3.1.1.6.3 POEM-Einzelitem für Schlafstörungen

Tabelle 4-66: Beobachtungsdauer für POEM-Einzelitem für Schlafstörungen in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	171,9 (41,36)	176,2 (31,02)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-212,0)
Quelle: Table 14.2.10.11 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; POEM: <i>Patient-Oriented Eczema Measures</i> ; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-67: Rücklaufquote für POEM-Einzelitem für Schlafstörung in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	362/362 (100,0)	364/365 (99,7)
Woche 12	328/362 (90,6)	341/365 (93,4)
Woche 16	323/362 (89,2)	337/365 (92,3)
Woche 26	303/362 (83,7)	322/365 (88,2)
Quelle: Table 14.2.10.5 (57) FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS; POEM: <i>Patient-Oriented Eczema Measures</i>		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für das POEM-Einzelitem für Schlafstörungen zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

Tabelle 4-68: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber *Baseline* im POEM-Einzelitem für Schlafstörungen aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit einer Verbesserung um ≥ 1 Punkt gegenüber <i>Baseline</i> im POEM-Einzelitem für Schlafstörungen zu Woche 26 (ergänzend)				
257/344 (74,7)	270/346 (78,0)	0,96 [0,88; 1,04] p=0,3060	0,83 [0,58; 1,18] p=0,3059	-3,3 [-9,6; 3,0] p=0,3062
Quelle: Table 14.2.10.10.1 (57)				
Werte nach Therapieabbruch und nach <i>Rescue</i> -Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i> , siehe Abschnitt 4.2.5.2).				
a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i> -Kriterium erreicht haben (mITT).				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).				
IGA: <i>Investigator's Global Assessment</i> ; FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; mg: Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i> ; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i> ; POEM: <i>Patient-Oriented Eczema Measures</i> ; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus				

Zu Woche 26 wiesen im Abrocitinib-Arm 74,7 % der Patienten eine Verbesserung des POEM-Einzelitem für Schlafstörungen um ≥ 1 Punkt gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 78,0 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 0,96 [0,88; 1,04]; p-Wert = 0,3060).

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 die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.1.1.7 Hautschmerz (*Skin Pain NRS*) – RCT

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

Tabelle 4-69: Operationalisierung des Endpunkts Hautschmerz

Studie	Operationalisierung
JADE DARE	<p>Die <i>Skin Pain Numerical Rating Scale</i> (NRS) ist eine patientenberichtete Skala, in der die maximale Stärke des Schmerzes innerhalb der letzten 24 Stunden bewertet wird (0: kein Schmerz bis 10: schlimmster vorstellbarer Schmerz).</p> <p><u>Hauptanalyse:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 4 Punkte gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen <p>Die Hauptanalyse erfolgte im modifizierten <i>Intention-to-treat-Analysis-Set</i> (mITT), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten und das entsprechende <i>Response</i>-Kriterium zu <i>Baseline</i> erreicht haben. Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response</i>-Imputation ersetzt.</p> <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Mittlere Veränderung gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 2, 12, 16, 20 und 26 <p>Die Sensitivitätsanalyse erfolgte im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben.</p> <p>Die folgenden <i>Estimands</i> wurden betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 1 (Composite Estimand)</i> für Responderanalysen • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-Haenszel-Tests. Die Stratifizierung beinhaltete den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares</i> (LS)-Mean-Differenzen und des Hedges'g mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.</p>
<p>IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; MMRM: <i>Mixed Model Repeated Measures</i>; NRS: <i>Numerical Rating Scale</i>; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus</p>	

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 4-70: Bewertung des Verzerrungspotenzials für Hautschmerz 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

Tabelle 4-71: Beobachtungsdauer für *Skin Pain* NRS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	171,5 (41,88)	176,6 (29,76)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-212,0)
Quelle: Table 14.2.12.8 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; NRS: <i>Numerical Rating Scale</i> ; N: Anzahl der in der Analyse eingeschlossenen Patienten; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-72: Rücklaufquote für *Skin Pain* NRS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 2	348/362 (96,1)	351/365 (96,2)
Woche 12	327/362 (90,3)	341/365 (93,4)
Woche 16	321/362 (88,7)	337/365 (92,3)
Woche 20	317/362 (87,6)	332/365 (91,0)
Woche 26	301/362 (83,1)	323/365 (88,5)
Quelle: Table 14.2.12.5 (57) FAS: <i>Full-Analysis-Set</i> ;mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS; NRS: <i>Numerical Rating Scale</i>		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für die *Skin Pain* NRS zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

Tabelle 4-73: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber *Baseline* in der *Skin Pain* NRS aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber <i>Baseline</i> in der <i>Skin Pain</i> NRS zu Woche 26				
205/316 (64,9)	202/325 (62,2)	1,04 [0,93; 1,17] p=0,4745	1,12 [0,82; 1,55] p=0,4746	2,7 [-4,7; 10,2] p=0,4751
<p>Quelle: Table 14.2.12.4.1.1 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i>-Kriterium erreicht haben (mITT).</p> <p>b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; mg Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; NRS: <i>Numerical Rating Scale</i>; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus</p>				

Zu Woche 26 wiesen im Abrocitinib-Arm 64,9 % der Patienten eine Verbesserung des *Skin Pain* NRS Scores um ≥ 4 Punkte gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 62,2 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 1,04 [0,93; 1,17]; p-Wert = 0,4745).

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

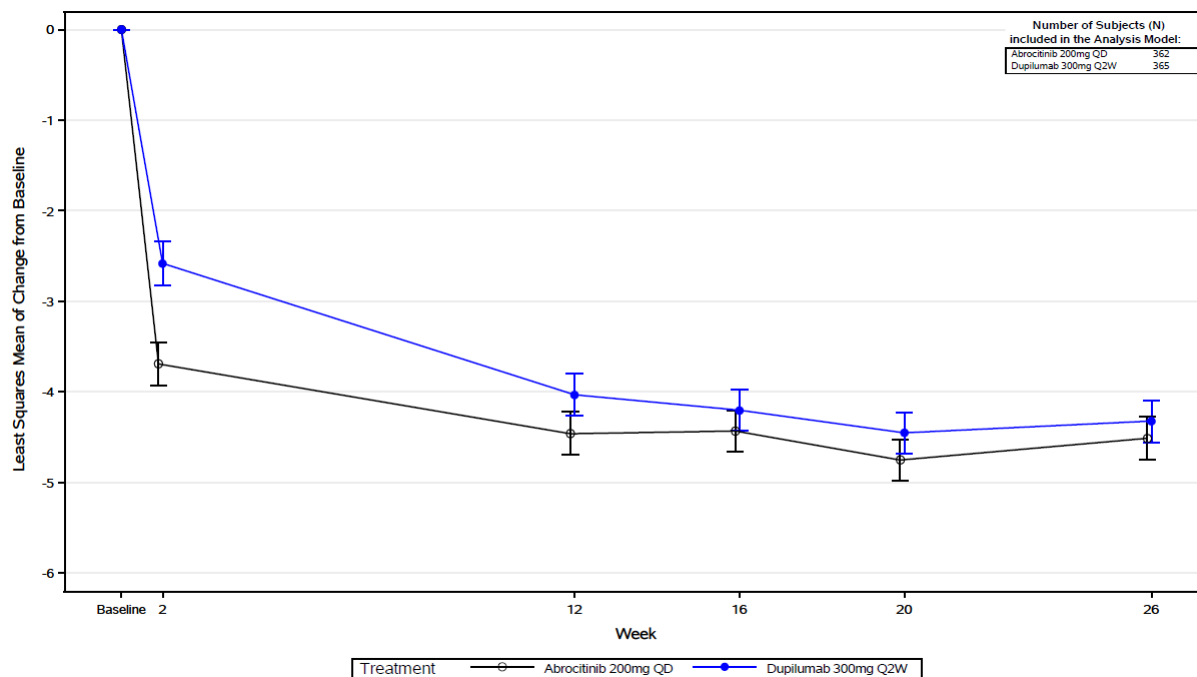
Tabelle 4-74: Ergebnisse für mittlere Veränderung in der *Skin Pain* NRS; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung in der <i>Skin Pain</i> NRS								
Baseline	362	6,5 (2,4)	NA	365	6,3 (2,3)	NA	NA	NA
Woche 2	348	2,7 (2,4)	-3,69 (0,12)	351	3,8 (2,5)	-2,58 (0,12)	-1,11 [-1,45; -0,77] p<0,0001	-0,48 [-0,63; -0,33]
Woche 12	327	1,9 (2,2)	-4,46 (0,12)	341	2,3 (2,3)	-4,03 (0,12)	-0,43 [-0,76; -0,10] p=0,0116	-0,20 [-0,35; -0,04]
Woche 16	321	1,9 (2,3)	-4,43 (0,12)	337	2,1 (2,1)	-4,20 (0,12)	-0,23 [-0,56; 0,09] p=0,1613	-0,11 [-0,26; 0,04]
Woche 20	317	1,6 (2,1)	-4,75 (0,12)	332	1,8 (2,1)	-4,45 (0,11)	-0,30 [-0,62; 0,02] p=0,0648	-0,15 [-0,30; 0,01]
Woche 26	301	1,8 (2,1)	-4,51 (0,12)	323	2,0 (2,2)	-4,32 (0,12)	-0,19 [-0,53; 0,14] p=0,2657	-0,09 [-0,25; 0,07]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
<p>Quelle: Table 14.2.12.1, Table 14.2.12.5 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie wurden als fehlende Werte betrachtet (<i>Hypothetical Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Deskriptive Analyse basierend auf den Patienten im FAS mit <i>Baseline</i>-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).</p> <p>b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem <i>Restricted Maximum Likelihood</i> Ansatz.</p> <p>FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mg: Milligramm; MMRM: <i>Mixed Model Repeated Measures</i>; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit <i>Baseline</i>-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; NRS: <i>Numerical Rating Scale</i>; SD: Standardabweichung (<i>Standard Deviation</i>); SE: Standardfehler (<i>Standard Error</i>); vs.: Versus</p>								

In der Analyse der mittleren Veränderung der *Skin Pain* NRS gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 2 und 12 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Der beobachtete Behandlungsunterschied ist zu Woche 2 als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges'g) außerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	2	12	16	20	26
N* Abrocitinib 200 mg	362	348	327	321	317	301
N* Dupilumab 300 mg	365	351	341	337	332	323

Abbildung 4-12: Verlaufskurve für die mittlere Veränderung der *Skin Pain* NRS von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel
mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren Veränderungen in der *Skin Pain* NRS zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-12).

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 die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.1.1.8 Ängste und Depressionen (HADS) – RCT

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

Tabelle 4-75: Operationalisierung des Endpunkts Ängste und Depressionen

Studie	Operationalisierung
JADE DARE	<p>Die <i>Hospital Anxiety and Depression Scale</i> (HADS) ist ein validiertes und patientenberichtetes Instrument zur Beurteilung von Angstzuständen und Depression während der vergangenen Woche und berücksichtigt die psychosozialen Belastungen durch die atopischen Dermatitis. Die HADS setzt sich aus 2 <i>Subscores</i> (Angst und Depression) mit jeweils 7 Items zusammen, die auf einer 4-Punkte-Skala (0-3) bewertet werden. Die HADS-<i>Subscores</i> können jeweils Werte von 0-21 annehmen, wobei ein höherer Wert einer größeren psychosozialen Belastung entspricht. Ab 8 Punkten liegt eine klinisch relevante Symptomatik im jeweiligen <i>Subscore</i> vor.</p> <p><u>Hauptanalysen:</u></p> <ul style="list-style-type: none"> • Mittlere Veränderung im <i>Subscore</i> Angst gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 12, 16 und 26 • Mittlere Veränderung im <i>Subscore</i> Depression gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 12, 16 und 26 <p>Die Analysen erfolgten im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben.</p> <p>Der folgende <i>Estimand</i> wurde betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-Haenszel-Tests. Die Stratifizierung beinhaltete den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares</i> (LS)-<i>Mean-Differenzen</i> und des Hedges'g mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.</p>
<p>HADS: <i>Hospital Anxiety and Depression Scale</i>; IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); MMRM: <i>Mixed Model Repeated Measures</i>; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus</p>	

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

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

Tabelle 4-76: Bewertung des Verzerrungspotenzials für Ängste und Depressionen 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

Tabelle 4-77: Beobachtungsdauer für HADS in RCT JADE DARE mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	171,9 (41,36)	176,2 (31,02)
Median (min-max)	183,0 (1,0-268,0)	183,0 (1,0-212,0)
Quelle: Table 14.2.9.12.1 (57) HADS: <i>Hospital Anxiety and Depression Scale</i> ; max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-78: Rücklaufquote für HADS Subscore für Angst und HADS Subscore für Depression in RCT JADE DARE mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
HADS-Subscore für Angst		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 12	329/362 (90,9)	341/365 (93,4)
Woche 16	321/362 (88,7)	337/365 (92,3)
Woche 26	303/362 (83,7)	322/365 (88,2)
HADS-Subscore für Depression		
<i>Baseline</i>	362/362 (100,0)	365/365 (100,0)
Woche 12	329/362 (90,9)	341/365 (93,4)
Woche 16	321/362 (88,7)	337/365 (92,3)
Woche 26	303/362 (83,7)	322/365 (88,2)
Quelle: Table 14.2.9.5 (57) FAS: <i>Full-Analysis-Set</i> ; HADS: <i>Hospital Anxiety and Depression Scale</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für den HADS-Subscore für Angst und für den HADS-Subscore für Depression zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

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

Tabelle 4-79: Ergebnisse für Ängste und Depressionen (HADS); MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung im HADS-Subscore für Angst								
Baseline	362	5,1 (3,7)	NA	365	5,2 (3,6)	NA	NA	NA
Woche 12	329	4,3 (3,5)	-0,81 (0,14)	341	4,3 (3,4)	-0,84 (0,14)	0,03 [-0,36; 0,41] p=0,8949	0,01 [-0,14; 0,16]
Woche 16	321	4,0 (3,5)	-1,05 (0,14)	337	4,0 (3,3)	-1,16 (0,14)	0,11 [-0,28; 0,50] p=0,5814	0,04 [-0,11; 0,20]
Woche 26	303	4,0 (3,5)	-1,06 (0,16)	322	3,9 (3,3)	-1,21 (0,15)	0,15 [-0,29; 0,58] p=0,4991	0,05 [-0,10; 0,21]
Mittlere Veränderung im HADS-Subscore für Depression								
Baseline	362	3,3 (3,2)	NA	365	3,3 (3,0)	NA	NA	NA
Woche 12	329	2,5 (2,9)	-0,75 (0,12)	341	2,5 (2,7)	-0,74 (0,12)	-0,01 [-0,35; 0,33] p=0,9601	0,00 [-0,16; 0,15]
Woche 16	321	2,5 (3,0)	-0,76 (0,13)	337	2,4 (2,7)	-0,90 (0,12)	0,14 [-0,21; 0,49] p=0,4398	0,06 [-0,09; 0,21]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 365)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Woche 26	303	2,5 (3,0)	-0,78 (0,14)	322	2,2 (2,7)	-1,01 (0,13)	0,24 [-0,14; 0,61] p=0,2132	0,10 [-0,06; 0,26]

Quelle: Table 14.2.9.1.1.1, Table 14.2.9.1.2.1, Table 14.2.9.5 (57)

Werte nach Therapieabbruch und nach *Rescue*-Therapie wurden als fehlende Werte betrachtet (*Hypothetical Estimand*, siehe Abschnitt 4.2.5.2).

a: Deskriptive Analyse basierend auf den Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).

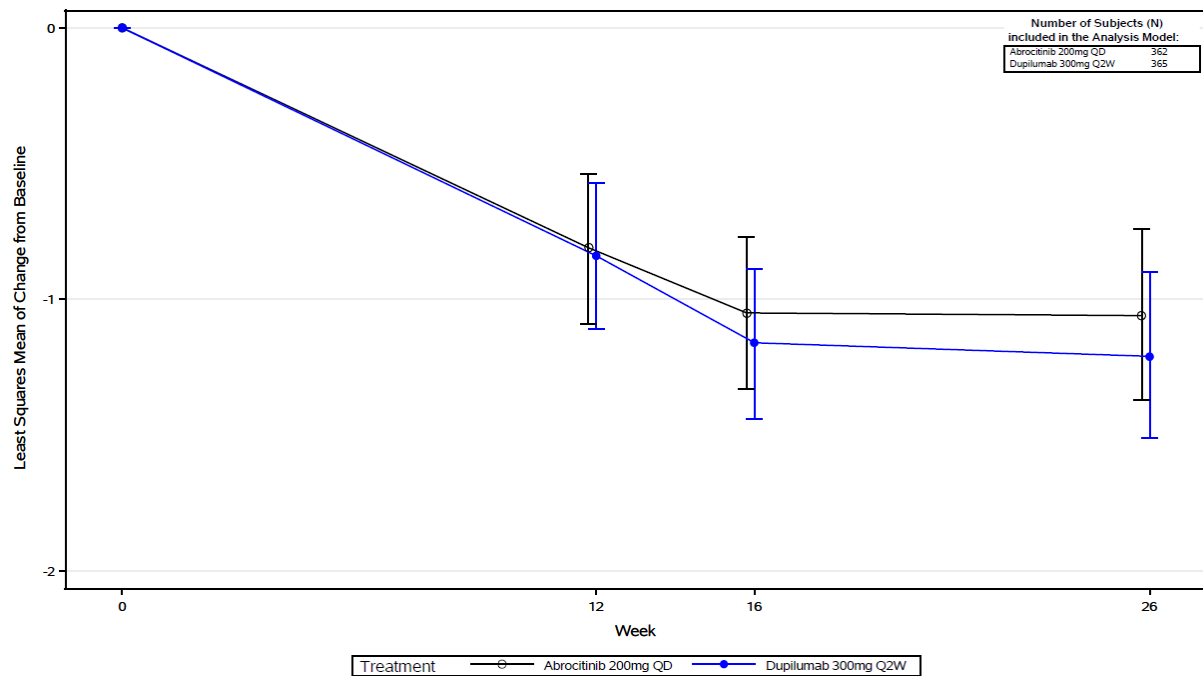
b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie *Baseline*-Wert + Krankheitsschwere zu *Baseline* als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem *Restricted Maximum Likelihood* Ansatz.

FAS: *Full-Analysis-Set*; HADS: *Hospital Anxiety and Depression Scale*; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (*Least Squares*);

mg: Milligramm; MMRM: *Mixed Model Repeated Measures*; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SD: Standardabweichung (*Standard Deviation*);

SE: Standardfehler (*Standard Error*); vs.: Versus

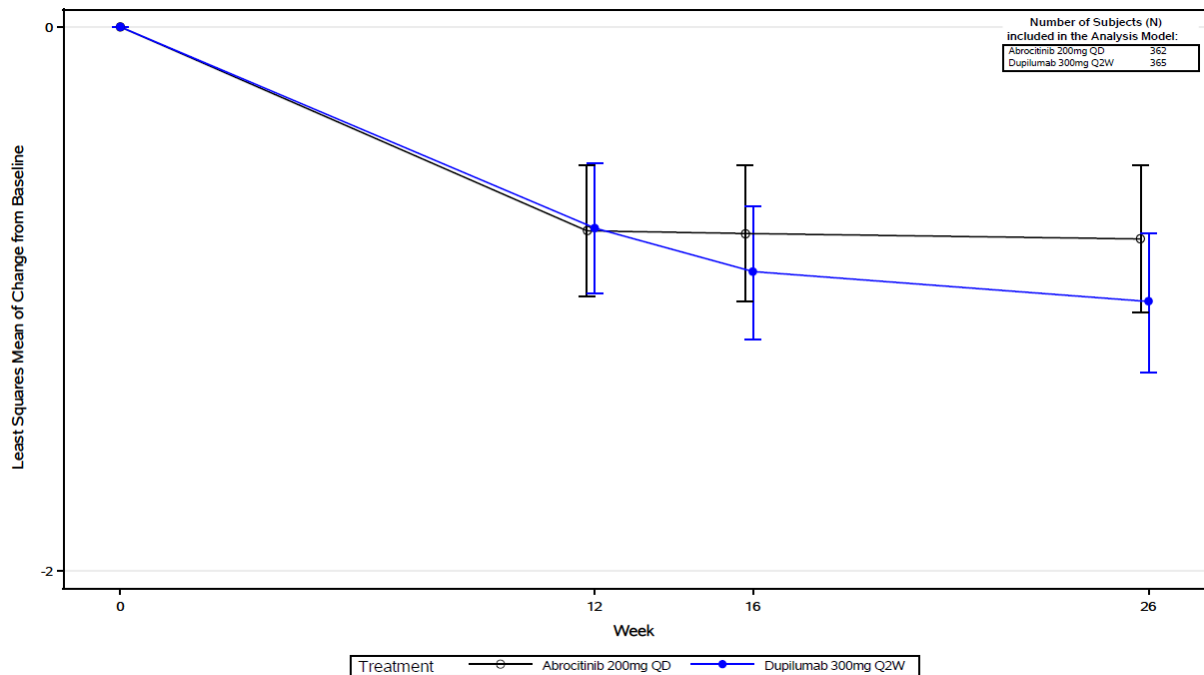
In der Analyse der mittleren Veränderung des HADS-Subscores für Angst und des HADS-Subscores für Depression gegenüber *Baseline* mittels MMRM zeigte sich zu keinem der Bewertungszeitpunkte ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen.



Woche	0	12	16	26
N* Abrocitinib 200 mg	362	329	321	303
N* Dupilumab 300 mg	365	341	337	322

Abbildung 4-13: Verlaufskurve für die mittlere Veränderung des HADS-Subscores für Angst von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die VerlaufsDarstellungen der mittleren Veränderungen des HADS-Subscores für Angst zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf vergleichbares Ansprechen zwischen den Behandlungsarmen (siehe Abbildung 4-13).



Woche	0	12	16	26
N* Abrocitinib 200 mg	362	329	321	303
N* Dupilumab 300 mg	365	341	337	322

Abbildung 4-14: Verlaufskurve für die mittlere Veränderung des HADS-Subscores für Depression von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die VerlaufsDarstellungen der mittleren Veränderungen des HADS-Subscores für Depression zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf vergleichbares Ansprechen zwischen den Behandlungsarmen (siehe Abbildung 4-14).

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 die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.1.1.9 Gesundheitszustand (EQ-5D VAS) – RCT

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

Tabelle 4-80: Operationalisierung des Endpunkts Gesundheitszustand

Studie	Operationalisierung
JADE DARE	<p>Die <i>European Quality of Life Group 5 Dimensions Visuelle Analogskala (EQ-5D VAS)</i> ist ein etabliertes und patientenberichtetes Instrument zur Messung des Gesundheitszustands. Die Patienten bewerten ihren Gesundheitszustand auf einer Skala von 0 (schlechtester vorstellbarer Gesundheitszustand) bis 100 (bester vorstellbarer Gesundheitszustand).</p> <p><u>Hauptanalyse:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 15 Punkte gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen <p>Die Hauptanalyse erfolgte im modifizierten <i>Intention-to-treat-Analysis-Set</i> (mITT), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten und das entsprechende <i>Response</i>-Kriterium zu <i>Baseline</i> erreicht haben. Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response</i>-Imputation ersetzt.</p> <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Mittlere Veränderung gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 12, 16 und 26 <p>Die Sensitivitätsanalyse erfolgte im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben. Die folgenden <i>Estimands</i> wurden betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 1 (Composite Estimand)</i> für Responderanalysen • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-Haenszel-Tests. Die Stratifizierung beinhaltete den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares (LS)-Mean-Differenzen</i> und des Hedges'g mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.</p>
<p>IGA: <i>Investigator's Global Assessment</i>; EQ-5D VAS: <i>European Quality of Life Group 5 Dimensions Visuelle Analogskala</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; MMRM: <i>Mixed Model Repeated Measures</i>; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus</p>	

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

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

Tabelle 4-81: Bewertung des Verzerrungspotenzials für Gesundheitszustand 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig

ITT: *Intention-to-treat*; RCT: Randomisierte kontrollierte Studie;

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

Tabelle 4-82: Beobachtungsdauer für EQ-5D VAS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 362
Beobachtungsdauer (Tage)		
MW (SD)	174,9 (39,68)	178,7 (29,44)
Median (min-max)	183,0 (1,0-282,0)	183,0 (1,0-221,0)

Quelle: Table 14.2.8.10 (57)
EQ-5D VAS: *European Quality of Life Group 5 Dimensions* Visuelle Analogskala; max: Maximal; min: Minimal; mg: Milligramm; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SD: Standardabweichung (*Standard Deviation*)

Tabelle 4-83: Rücklaufquote für EQ-5D VAS in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	362/362 (100,0)	364/365 (99,7)
Woche 12	329/362 (90,9)	341/365 (93,4)
Woche 16	323/362 (89,2)	337/365 (92,3)
Woche 26	303/362 (83,7)	324/365 (88,8)
Quelle: Table 14.2.8.3 (57) EQ-5D VAS: <i>European Quality of Life Group 5 Dimensions</i> Visuelle Analogskala; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für die EQ-5D VAS zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

Tabelle 4-84: Ergebnisse für Anteil der Patienten mit einer Verbesserung um ≥ 15 Punkte gegenüber *Baseline* in der EQ-5D VAS aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
		RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit einer Verbesserung um ≥ 15 Punkte gegenüber <i>Baseline</i> in der EQ-5D VAS zu Woche 26				
13/357 (3,6)	6/361 (1,7)	2,20 [0,85; 5,71] p=0,1040	2,27 [0,85; 6,07] p=0,1030	2,0 [-0,4; 4,4] p=0,0950
<p>Quelle: Table 14.2.8.8.1 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response-Imputation</i> ergänzt (<i>Composite Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i>-Kriterium erreicht haben (mITT).</p> <p>b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>EQ-5D VAS: <i>European Quality of Life Group 5 Dimensions</i> Visuelle Analogskala; IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; mg: Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; vs. Versus</p>				

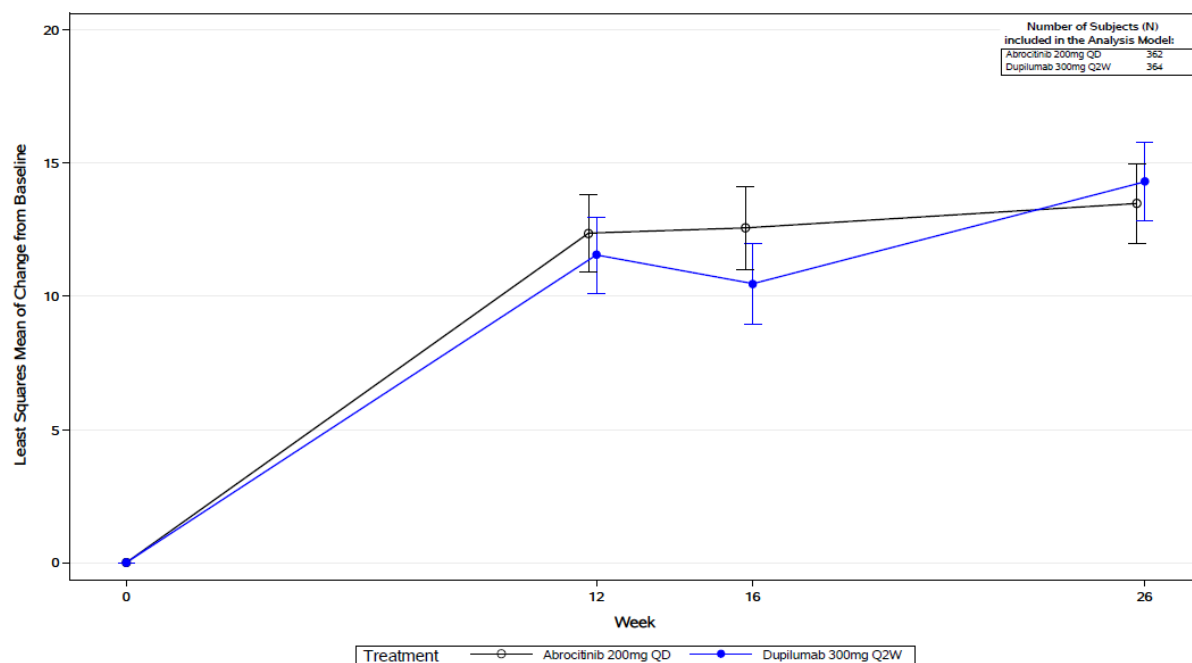
Zu Woche 26 wiesen im Abrocitinib-Arm 3,6 % der Patienten eine Verbesserung in der EQ-5D VAS um ≥ 15 Punkte gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 1,7 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 2,20 [0,85; 5,71]; p-Wert = 0,1040).

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

Tabelle 4-85: Ergebnisse für mittlere Veränderung der EQ-5D VAS; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 362)			Dupilumab 300 mg (N = 364)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung der EQ-5D VAS								
Baseline	362	68,4 (19,5)	NA	364	67,7 (18,3)	NA	NA	NA
Woche 12	329	80,5 (15,3)	12,37 (0,74)	340	79,5 (13,4)	11,55 (0,73)	0,82 [-1,22; 2,86] p=0,4310	0,06 [-0,09; 0,21]
Woche 16	323	80,8 (14,9)	12,57 (0,79)	336	78,5 (15,9)	10,47 (0,78)	2,09 [-0,08; 4,27] p=0,0592	0,15 [-0,01; 0,30]
Woche 26	303	81,8 (15,3)	13,48 (0,76)	323	82,5 (13,5)	14,30 (0,75)	-0,82 [-2,91; 1,28] p=0,4450	-0,06 [-0,22; 0,10]
<p>Quelle: Table 14.2.8.3, Table 14.2.8.4 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie wurden als fehlende Werte betrachtet (<i>Hypothetical Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Deskriptive Analyse basierend auf den Patienten im FAS mit <i>Baseline</i>-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).</p> <p>b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem <i>Restricted Maximum Likelihood</i> Ansatz.</p> <p>EQ-5D VAS: <i>European Quality of Life Group 5 Dimensions</i> Visuelle Analogskala; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; mg: Milligramm; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); MMRM: <i>Mixed Model Repeated Measures</i>; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit <i>Baseline</i>-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SD: Standardabweichung (<i>Standard Deviation</i>); SE: Standardfehler (<i>Standard Error</i>); vs.: Versus</p>								

In der Analyse der mittleren Veränderung der EQ-5D VAS gegenüber *Baseline* mittels MMRM zeigte sich zu keinem der Bewertungszeitpunkte ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen.



Woche	0	12	16	26
N* Abrocitinib 200 mg	362	329	323	303
N* Dupilumab 300 mg	364	340	336	323

Abbildung 4-15: Verlaufskurve für die mittlere Veränderung der EQ-5D VAS von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

EQ-5D VAS: European Quality of Life Group 5 Dimensions Visuelle Analogskala; mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die Verlaufsdarstellungen der mittleren Veränderungen in der EQ-5D VAS zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf vergleichbares Ansprechen zwischen den Behandlungsarmen (siehe Abbildung 4-15).

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 die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.1.2 Gesundheitsbezogene Lebensqualität

4.3.1.3.1.2.1 Krankheitsspezifische Lebensqualität (DLQI) – RCT

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

Tabelle 4-86: Operationalisierung des Endpunkts krankheitsspezifische Lebensqualität

Studie	Operationalisierung
JADE DARE	<p>Der <i>Dermatology Life Quality Index</i> (DLQI) ist ein validierter und patientenberichteter Fragebogen zur Erhebung der krankheitsspezifischen Lebensqualität bei dermatologischen Erkrankungen. Der DLQI besteht aus 10 Items, die den folgenden 6 Subskalen zugeordnet werden: Symptome und Gefühle, tägliche Aktivitäten, Freizeit, Arbeit/ Schule, persönliche Beziehungen und Auswirkungen der Behandlung auf den Alltag. Die Items werden auf einer 4-Punkte-Skala beantwortet (0: gar nicht, 1: etwas, 2: stark, 3: sehr stark). Der DLQI-Gesamtscore kann Werte von 0-30 annehmen, wobei ein höherer Wert einer größeren Beeinträchtigung der Lebensqualität entspricht. Ein Gesamtscore im DLQI von 0 oder 1 entspricht keiner Beeinträchtigung der Lebensqualität.</p> <p><u>Hauptanalysen:</u></p> <ul style="list-style-type: none"> • Anteil der Patienten mit DLQI-Gesamtscore = 0 oder 1 zu Woche 26 • Anteil der Patienten, die eine Verbesserung (Reduktion) um ≥ 5 Punkte gegenüber dem <i>Baseline</i>-Wert zu Woche 26 erreichen <p>Die Hauptanalysen erfolgten im modifizierten <i>Intention-to-treat-Analysis-Set</i> (mITT), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten und das entsprechende <i>Response</i>-Kriterium zu <i>Baseline</i> erreicht haben. Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie sowie fehlende Werte wurden mittels <i>Non-Response</i>-Imputation ersetzt.</p> <p><u>Sensitivitätsanalyse:</u></p> <ul style="list-style-type: none"> • Mittlere Veränderung gegenüber dem <i>Baseline</i>-Wert anhand eines MMRM zu Woche 2, 12, 16, 20 und 26 <p>Die Sensitivitätsanalyse erfolgte im <i>Full-Analysis-Set</i> (FAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben.</p> <p>Die folgenden <i>Estimands</i> wurden betrachtet:</p> <ul style="list-style-type: none"> • <i>Estimand 1 (Composite Estimand)</i> für Responderanalysen • <i>Estimand 2 (Hypothetical Estimand)</i> für die mittlere Änderung <p>Die Methodik der <i>Estimands</i> ist in Abschnitt 4.2.5.2 beschrieben.</p> <p>Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand des stratifizierten Cochran-Mantel-Haenszel-Tests. Die Stratifizierung beinhaltet den Faktor Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).</p> <p>Das MMRM umfasste die Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Im Rahmen der MMRM-Analysen erfolgte der Gruppenvergleich anhand der <i>Least Squares</i> (LS)-<i>Mean</i>-Differenzen und des Hedges'g mit den korrespondierenden 95 %-KI zu jedem Beobachtungszeitpunkt.</p>
	<p>DLQI: <i>Dermatology Life Quality Index</i>; IGA: <i>Investigator's Global Assessment</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i>; MMRM: <i>Mixed Model Repeated Measures</i>; OR: <i>Odds Ratio</i>; RD: Risikodifferenz; RR: Relatives Risiko; vs: Versus</p>

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 4-87: Bewertung des Verzerrungspotenzials für krankheitsspezifische Lebensqualität 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
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

Tabelle 4-88: Beobachtungsdauer für DLQI in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Beobachtungsdauer (Tage)		
MW (SD)	175,4 (38,20)	179,0 (28,42)
Median (min-max)	183,0 (1,0-282,0)	183,0 (1,0-221,0)
Quelle: Table 14.2.6.12 (57) DLQI: <i>Dermatology Life Quality Index</i> ; max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der in der Analyse eingeschlossenen Patienten; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-89: Rücklaufquote für DLQI in RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Zeitpunkt	Abrocitinib 200 mg n/N (%)	Dupilumab 300 mg n/N (%)
JADE DARE		
<i>Baseline</i>	361/362 (99,7)	363/365 (99,5)
Woche 2	347/362 (95,9)	351/365 (96,2)
Woche 12	329/362 (90,9)	341/365 (93,4)
Woche 16	325/362 (89,8)	338/365 (92,6)
Woche 20	320/362 (88,4)	333/365 (91,2)
Woche 26	303/362 (83,7)	325/365 (89,0)
Quelle: Table 14.2.6.9 (57) DLQI: <i>Dermatology Life Quality Index</i> ; FAS: <i>Full-Analysis-Set</i> ; mg: Milligramm; n: Anzahl der Patienten mit vorliegendem Wert zum beobachteten Zeitpunkt; N: Anzahl der Patienten im FAS		

Bezogen auf die FAS-Population lag die Rücklaufquote in der Studie JADE DARE für den DLQI zu allen Bewertungszeitpunkten in beiden Behandlungsarmen über 83 %.

Tabelle 4-90: Ergebnisse für DLQI aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Anteil der Patienten mit DLQI-Gesamtscore von 0 oder 1 zu Woche 26				
137/358 (38,3)	114/361 (31,6)	1,21 [0,99; 1,48] p=0,0603	1,34 [0,99; 1,83] p=0,0596	6,7 [-0,2; 13,7] p=0,0596
Anteil der Patienten mit einer Verbesserung um ≥ 5 Punkte gegenüber <i>Baseline</i> im DLQI zu Woche 26				
243/336 (72,3)	270/345 (78,3)	0,92 [0,85; 1,01] p=0,0707	0,72 [0,51; 1,03] p=0,0701	-6,0 [-12,4; 0,5] p=0,0699
Quelle: Table 14.2.6.5.1, Table 14.2.6.10.1 (57)				
Werte nach Therapieabbruch und nach <i>Rescue</i> -Therapie sowie fehlende Werte wurden mittels <i>Non-Response</i> -Imputation ergänzt (<i>Composite Estimand</i> , siehe Abschnitt 4.2.5.2).				
a: Anzahl der Patienten im FAS, die zu <i>Baseline</i> das entsprechende <i>Response</i> -Kriterium erreicht haben (mITT).				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4).				
DLQI: <i>Dermatology Life Quality Index</i> ; IGA: <i>Investigator's Global Assessment</i> ; FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; mg: Milligramm; mITT: Modifiziertes <i>Intention-to-treat-Analysis-Set</i> ; n: Anzahl der Patienten mit Ereignis; N: Anzahl der in der Analyse eingeschlossenen Patienten; OR: <i>Odds Ratio</i> ; RD: Risikodifferenz; RR: Relatives Risiko; vs.: Versus				

Zu Woche 26 erreichten im Abrocitinib-Arm 38,3 % der Patienten einen DLQI-Gesamtscore von 0 oder 1 und somit nahezu keine Beeinträchtigung der Lebensqualität durch die AD; unter Dupilumab lag der entsprechende Anteil der Patienten bei 31,6 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 1,21 [0,99; 1,48]; p-Wert = 0,0603).

Zu Woche 26 wiesen im Abrocitinib-Arm 72,3 % der Patienten eine Verbesserung des DLQI-Gesamtscores um ≥ 5 Punkte gegenüber dem Ausgangswert auf; unter Dupilumab lag der entsprechende Anteil der Patienten bei 78,3 %. Der Unterschied zwischen den Behandlungsarmen war nicht statistisch signifikant (RR [95 %-KI]: 0,92 [0,85; 1,01]; p-Wert = 0,0707).

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

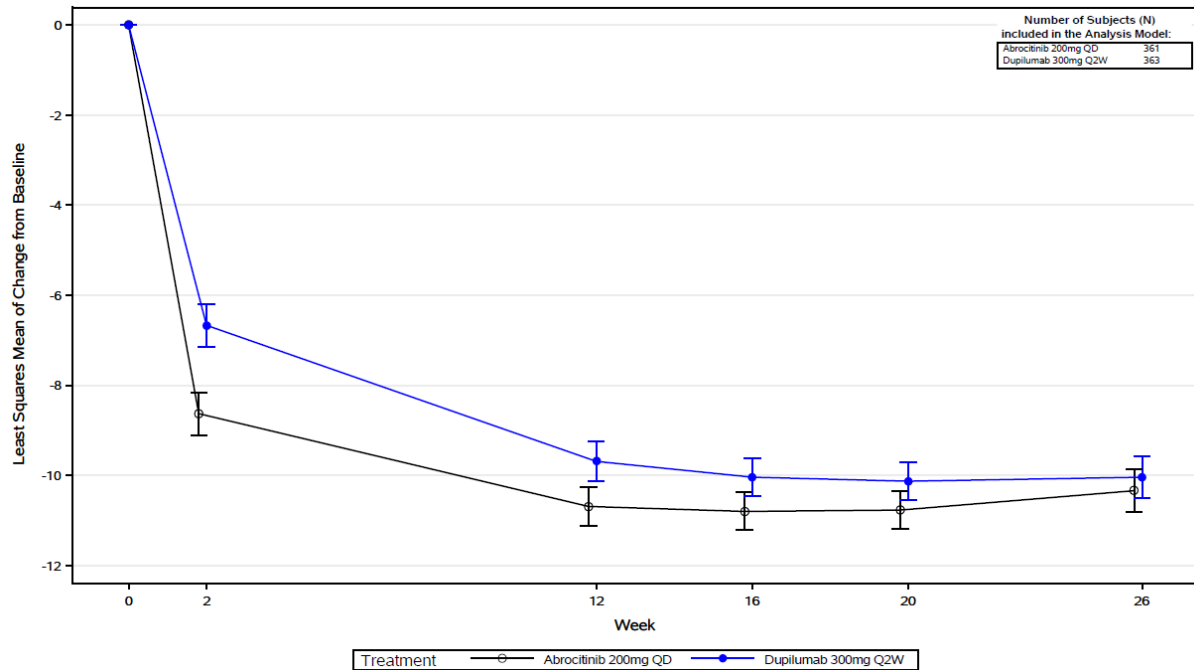
Tabelle 4-91: Ergebnisse für mittlere Veränderung des DLQI; MMRM-Analyse aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 361)			Dupilumab 300 mg (N = 363)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
Mittlere Veränderung des DLQI								
Baseline	361	14,0 (6,8)	NA	363	14,2 (6,3)	NA	NA	NA
Woche 2	346	5,4 (5,1)	-8,63 (0,24)	349	7,4 (5,1)	-6,67 (0,24)	-1,96 [-2,64; -1,29] p<0,0001	-0,43 [-0,58; -0,28]
Woche 12	328	3,5 (4,0)	-10,69 (0,22)	339	4,3 (4,7)	-9,69 (0,22)	-1,00 [-1,62; -0,38] p=0,0016	-0,25 [-0,40; -0,09]
Woche 16	324	3,2 (3,9)	-10,80 (0,22)	336	4,0 (4,3)	-10,04 (0,21)	-0,76 [-1,35; -0,16] p=0,0126	-0,19 [-0,35; -0,04]
Woche 20	319	3,3 (4,0)	-10,77 (0,22)	331	3,8 (4,1)	-10,13 (0,21)	-0,64 [-1,23; -0,04] p=0,0376	-0,16 [-0,32; -0,01]
Woche 26	302	3,6 (4,7)	-10,34 (0,24)	323	4,0 (4,5)	-10,04 (0,24)	-0,30 [-0,97; 0,37] p=0,3814	-0,07 [-0,23; 0,09]

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

Endpunkt Beobachtungs- zeitpunkt	Abrocitinib 200 mg (N = 361)			Dupilumab 300 mg (N = 363)			Abrocitinib 200 mg vs. Dupilumab 300 mg	
	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	N*	MW (SD) ^a	LS-Mean Änderung zu Baseline (SE) ^b	LS-Mean-Differenz [95 %-KI] p-Wert ^b	Hedges'g [95 %-KI]
<p>Quelle: Table 14.2.6.1.1, Table 14.2.6.9 (57)</p> <p>Werte nach Therapieabbruch und nach <i>Rescue</i>-Therapie wurden als fehlende Werte betrachtet (<i>Hypothetical Estimand</i>, siehe Abschnitt 4.2.5.2).</p> <p>a: Deskriptive Analyse basierend auf den Patienten im FAS mit <i>Baseline</i>-Wert und Wert zum jeweiligen Beobachtungszeitpunkt (N*).</p> <p>b: Analyse basieren auf einem MMRM mit den Faktoren Behandlung + Visite + Visite*Behandlung als feste Effekte sowie <i>Baseline</i>-Wert + Krankheitsschwere zu <i>Baseline</i> als Kovariaten. Es wurde eine unstrukturierte Kovarianz-Matrix angenommen. Die Analyse basiert auf dem <i>Restricted Maximum Likelihood</i> Ansatz.</p> <p>DLQI: <i>Dermatology Life Quality Index</i>; FAS: <i>Full-Analysis-Set</i>; KI: Konfidenzintervall; LS: Methode der kleinsten Quadrate (<i>Least Squares</i>); mg: Milligramm; MMRM: <i>Mixed Model Repeated Measures</i>; MW: Mittelwert; N: Anzahl der eingeschlossenen Patienten im MMRM; N*: Anzahl der Patienten im FAS mit <i>Baseline</i>-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; NA: Nicht anwendbar; SD: Standardabweichung (<i>Standard Deviation</i>); SE: Standardfehler (<i>Standard Error</i>); vs.: Versus</p>								

In der Analyse der mittleren Veränderung des DLQI-Gesamtscores gegenüber *Baseline* mittels MMRM zeigte sich zu Woche 2, 12, 16 und 20 ein statistisch signifikanter Unterschied der LS-Mean-Differenz zugunsten von Abrocitinib. Der beobachtete Behandlungsunterschied ist zu Woche 2 als klinisch relevant zu bewerten, da das gesamte 95 %-KI der SMD (Hedges'g) außerhalb des Irrelevanzbereichs von 0,2 liegt.



Woche	0	2	12	16	20	26
N* Abrocitinib 200 mg	361	346	328	324	319	302
N* Dupilumab 300 mg	363	349	339	336	331	323

Abbildung 4-16: Verlaufskurve für die mittlere Veränderung des DLQI von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel
mg: Milligramm; N*: Anzahl der Patienten im FAS mit *Baseline*-Wert und Wert zum jeweiligen Beobachtungszeitpunkt; QD: Täglich; Q2W: Alle 2 Wochen

Die VerlaufsDarstellungen der mittleren Veränderungen im DLQI zu den Beobachtungszeitpunkten zeigten ein über den Studienverlauf hinweg konsistent höheres Ansprechen unter Abrocitinib gegenüber Dupilumab (siehe Abbildung 4-16).

4.3.1.3.1.3 Verträglichkeit – RCT

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

Tabelle 4-92: Operationalisierung der Endpunkte zu Verträglichkeit

Studie	Operationalisierung
JADE DARE	<p>Die Erhebung der unerwünschten Ereignisse (UE) erfolgte ab Studienbeginn bis hin zu 28 Tagen nach der letzten Dosis der Studienmedikation (<i>Safety Follow-up</i>). Die UE wurden standardmäßig nach verbindlichen internationalen Standards (Gute klinische Praxis, <i>Good Clinical Practice</i>, [GCP]) erhoben sowie gemäß Version 24.0 des <i>Medical Dictionary for Regulatory Activities</i> (MedDRA) kodiert.</p> <p>Zur Evaluierung der für die Bewertung relevanten Aspekte werden die folgenden Endpunkte dargestellt:</p> <ul style="list-style-type: none"> • UE jeglichen Schweregrades <ul style="list-style-type: none"> ○ Gesamtrate (ergänzend) ○ Häufige UE nach SOC und PT, die bei mindestens 10 % der Patienten oder mindestens 10 Patienten und 1 % der Patienten in mindestens einem Studienarm aufgetreten sind. • Schwere UE <ul style="list-style-type: none"> ○ Gesamtrate ○ Häufige schwere UE nach SOC und PT, die bei mindestens 5 % der Patienten oder mindestens 10 Patienten und 1 % der Patienten in mindestens einem Studienarm aufgetreten sind. • Schwerwiegende UE (SUE) <ul style="list-style-type: none"> ○ Gesamtrate ○ Häufige SUE nach SOC und PT, die bei mindestens 5 % der Patienten oder mindestens 10 Patienten und 1 % der Patienten in mindestens einem Studienarm aufgetreten sind. • UE, die zum Therapieabbruch führten <ul style="list-style-type: none"> ○ Gesamtrate ○ Darstellung nach SOC und PT (deskriptiv) • Spezifische UE, jeweils als Gesamtrate und differenziert nach Schweregrad (schwer, schwerwiegend): <ul style="list-style-type: none"> ○ Superinfektionen ○ Herpes zoster ○ Konjunktivitis ○ Akne ○ Follikulitis <p>Spezifische UE wurden anhand von <i>customized MedDRA-Queries</i> (Herpes zoster) oder im Rahmen des detaillierten CRF (Konjunktivitis, Akne und Follikulitis) erfasst. Superinfektionen waren definiert als Infektionen, die > 2 Wochen mit antibiotischen, antiviralen oder fungiziden Wirkstoffen behandelt wurden.</p> <p>Ergebnisse zu UE jeglichen Schweregrades, schwere UE und SUE ohne Progressionsereignisse sind in Anhang 4-G dargestellt.</p> <p>Die Analysen erfolgten im <i>Safety-Analysis-Set</i> (SAS), welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienintervention erhalten haben. Im SAS wurden alle Patienten gemäß ihrer erhaltenen Studienintervention analysiert.</p>

Studie	Operationalisierung
	Der Gruppenvergleich mit der Schätzung des relativen Risikos (RR), Odds Ratios (OR) und der Risikodifferenz (RD) mit den korrespondierenden 95 %-Konfidenzintervallen (KI) und p-Werten erfolgte anhand der unstratifizierten Approximation der Normalverteilung nach der Wald-Methode.
CRF: <i>Case Report Form</i> ; GCP: Gute klinische Praxis (<i>Good Clinical Practice</i>); KI: Konfidenzintervall; MedDRA: <i>Medical Dictionary for Regulatory Activities</i> ; OR: <i>Odds Ratio</i> ; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RD: Risikodifferenz; RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; SOC: Systemorganklasse (<i>System Organ Class</i>); SUE: Schwerwiegendes UE; UE: Unerwünschtes Ereignis	

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

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

Tabelle 4-93: Bewertung des Verzerrungspotenzials für Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
JADE DARE	niedrig	ja	ja	ja	ja	niedrig
ITT: <i>Intention-to-treat</i> ; RCT: Randomisierte kontrollierte Studie						

Begründen Sie für jede Studie die abschließende Einschätzung.

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Für die Analysen zur Verträglichkeit wurde das SAS herangezogen, welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienmedikation erhalten haben. Des Weiteren wurden keine das Verzerrungspotenzial beeinflussenden Aspekte oder eine ergebnisgesteuerte Berichterstattung identifiziert. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft.

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

Tabelle 4-94: Beobachtungsdauer für Verträglichkeit in RCT JADE DARE mit dem zu bewertenden Arzneimittel

Studie JADE DARE	Abrocitinib 200 mg N = 362	Dupilumab 300 mg N = 365
Behandlungsdauer orale Gabe (Tage)		
MW (SD)	172,5 (39,3)	177,0 (27,0)
Median (min-max)	183,0 (3,0-268,0)	183,0 (2,0-212,0)
Behandlungsdauer Injektion (Tage)		
MW (SD)	171,7 (38,2)	176,4 (26,3)
Median (min-max)	182,0 (14,0-269,0)	182,0 (14,0-203,0)
Quelle: Table 14.4.1.2 (57) max: Maximal; mg: Milligramm; min: Minimal; MW: Mittelwert; N: Anzahl der Patienten im SAS; SAS: <i>Safety-Analysis-Set</i> ; SD: Standardabweichung (<i>Standard Deviation</i>)		

Tabelle 4-95: Ergebnisse für Verträglichkeit aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
UE jeglichen Schweregrades (ergänzend)				
268/362 (74,0)	239/365 (65,5)	1,13 [1,03; 1,24] p=0,0124	1,50 [1,09; 2,07] p=0,0123	8,6 [1,9; 15,2] p=0,0117
Schwere UE				
11/362 (3,0)	8/365 (2,2)	1,39 [0,56; 3,41] p=0,4763	1,40 [0,56; 3,52] p=0,4761	0,8 [-1,5; 3,2] p=0,4743
SUE				
6/362 (1,7)	6/365 (1,6)	1,01 [0,33; 3,10] p=0,9885	1,01 [0,32; 3,16] p=0,9885	0,0 [-1,8; 1,9] p=0,9885
UE, die zum Therapieabbruch führten				
9/362 (2,5)	9/365 (2,5)	1,01 [0,40; 2,51] p=0,9859	1,01 [0,40; 2,57] p=0,9859	0,0 [-2,2; 2,3] p=0,9859
Quelle: Table 14.3.1.5.1, Table 14.3.1.5.3, Table 14.3.1.5.2, Table 14.3.1.5.15 (57)				
a: Anzahl der Patienten im SAS.				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet anhand der unstratifizierten Approximation der Normalverteilung nach der Wald-Methode.				
KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; OR: <i>Odds Ratio</i> ; RD: Risikodifferenz; RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; SUE: Schwerwiegendes UE; UE: Unerwünschtes Ereignis; vs.: Versus				

Bis Woche 26 trat bei 74,0 % der Patienten im Abrocitinib-Arm und 65,5 % der Patienten im Dupilumab-Arm mindestens ein UE auf.

Für das Auftreten von schweren UE (RR [95 %-KI]: 1,39 [0,33; 3,10]; p-Wert = 0,4763), SUE (RR [95 %-KI]: 1,01 [0,33; 3,10]; p-Wert = 0,9885) und UE, die zum Therapieabbruch führten (RR [95 %-KI]: 1,01 [0,40; 2,51]; p-Wert = 0,9859), zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen.

Die Ergebnisse zu UE, schweren UE und SUE ohne Progressionsereignisse sind vergleichbar zu den Analysen, die die Progressionsereignisse einschlossen, und in Anhang 4-G dargestellt.

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

Tabelle 4-96: Häufige UE nach SOC und PT aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

SOC ^a PT ^a	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
	n/N ^b (%)	n/N ^b (%)	RR [95 %-KI] p-Wert ^c	OR [95 %-KI] p-Wert ^c	RD [95 %-KI] p-Wert ^c
Infektionen und parasitäre Erkrankungen					
Jegliche PT	68/362 (18,8)	76/365 (20,8)	0,90 [0,67; 1,21] p=0,4911	0,88 [0,61; 1,27] p=0,4909	-2,0 [-7,8; 3,8] p=0,4905
COVID-19	15/362 (4,1)	12/365 (3,3)	1,26 [0,60; 2,65] p=0,5427	1,27 [0,59; 2,76] p=0,5426	0,9 [-1,9; 3,6] p=0,5418
Konjunktivitis	8/362 (2,2)	35/365 (9,6)	0,23 [0,11; 0,49] p=0,0001	0,21 [0,10; 0,47] p=0,0001	-7,4 [-10,8; -4,0] p<0,0001
Folikulitis	12/362 (3,3)	3/365 (0,8)	4,03 [1,15; 14,17] p=0,0296	4,14 [1,16; 14,79] p=0,0289	2,5 [0,4; 4,6] p=0,0179
Herpes simplex	12/362 (3,3)	5/365 (1,4)	2,42 [0,86; 6,80] p=0,0936	2,47 [0,86; 7,08] p=0,0928	2,0 [-0,3; 4,1] p=0,0826
Nasopharyngitis	14/362 (3,9)	12/365 (3,3)	1,18 [0,55; 2,51] p=0,6742	1,18 [0,54; 2,59] p=0,6742	0,6 [-2,1; 3,3] p=0,6739
Oraler Herpes	9/362 (2,5)	15/365 (4,1)	0,60 [0,27; 1,36] p=0,2260	0,59 [0,26; 1,38] p=0,2253	-1,6 [-4,2; 1,0] p=0,2197
Infektion der oberen Atemwege	10/362 (2,8)	9/365 (2,5)	1,12 [0,46; 2,72] p=0,8022	1,12 [0,45; 2,80] p=0,8021	0,3 [-2,0; 2,6] p=0,8021

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

SOC ^a PT ^a	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
	n/N ^b (%)	n/N ^b (%)	RR [95 %-KI] p-Wert ^c	OR [95 %-KI] p-Wert ^c	RD [95 %-KI] p-Wert ^c
Erkrankungen des Gastrointestinaltrakts					
Jegliche PT	70/362 (19,3)	11/365 (3,0)	6,42 [3,46; 11,91] p<0,0001	7,71 [4,01; 14,84] p<0,0001	16,3 [11,9; 20,8] p<0,0001
Übelkeit	70/362 (19,3)	8/365 (2,2)	8,82 [4,31; 18,07] p<0,0001	10,70 [5,07; 22,59] p<0,0001	17,2 [12,8; 21,5] p<0,0001
Erbrechen	11/362 (3,0)	6/365 (1,6)	1,85 [0,69; 4,95] p=0,2211	1,88 [0,69; 5,13] p=0,2204	1,4 [-0,8; 3,6] p=0,2134
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort					
Jegliche PT	10/362 (2,8)	5/365 (1,4)	2,02 [0,70; 5,84] p=0,1962	2,05 [0,69; 6,04] p=0,1955	1,4 [-0,7; 3,5] p=0,1867
Erschöpfung	10/362 (2,8)	5/365 (1,4)	2,02 [0,70; 5,84] p=0,1962	2,05 [0,69; 6,04] p=0,1955	1,4 [-0,7; 3,5] p=0,1867
Untersuchungen					
Jegliche PT	38/362 (10,5)	26/365 (7,1)	1,47 [0,91; 2,37] p=0,1112	1,53 [0,91; 2,58] p=0,1105	3,4 [-0,7; 7,5] p=0,1081
Kreatinphosphokinase im Blut erhöht	14/362 (3,9)	13/365 (3,6)	1,09 [0,52; 2,28] p=0,8275	1,09 [0,50; 2,35] p=0,8275	0,3 [-2,4; 3,1] p=0,8275

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

SOC ^a PT ^a	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
	n/N ^b (%)	n/N ^b (%)	RR [95 %-KI] p-Wert ^c	OR [95 %-KI] p-Wert ^c	RD [95 %-KI] p-Wert ^c
Abnahme der natürlichen Killerzellen	10/362 (2,8)	0	20,19 [1,18; 344,43] p=0,0378	20,74 [1,21; 356,39] p=0,0367	2,8 [1,1; 4,5] p=0,0013
SARS-CoV-2 positiv	15/362 (4,1)	13/365 (3,6)	1,16 [0,56; 2,41] p=0,6838	1,17 [0,55; 2,50] p=0,6837	0,6 [-2,2; 3,4] p=0,6835
Erkrankungen des Nervensystems					
Jegliche PT	55/362 (15,2)	27/365 (7,4)	2,05 [1,33; 3,18] p=0,0012	2,24 [1,38; 3,65] p=0,0011	7,8 [3,2; 12,4] p=0,0008
Schwindelgefühl	10/362 (2,8)	4/365 (1,1)	2,52 [0,80; 7,96] p=0,1152	2,56 [0,80; 8,25] p=0,1144	1,7 [-0,3; 3,7] p=0,1021
Kopfschmerzen	47/362 (13,0)	24/365 (6,6)	1,97 [1,23; 3,16] p=0,0045	2,12 [1,27; 3,55] p=0,0042	6,4 [2,1; 10,7] p=0,0035
Erkrankungen der Haut und des Unterhautgewebes					
Jegliche PT	61/362 (16,9)	23/365 (6,3)	2,67 [1,69; 4,22] p<0,0001	3,01 [1,82; 4,99] p<0,0001	10,6 [6,0; 15,1] p<0,0001
Akne	46/362 (12,7)	10/365 (2,7)	4,64 [2,38; 9,05] p<0,0001	5,17 [2,57; 10,41] p<0,0001	10,0 [6,2; 13,8] p<0,0001
Dermatitis atopisch	17/362 (4,7)	14/365 (3,8)	1,22 [0,61; 2,45] p=0,5666	1,24 [0,60; 2,55] p=0,5666	0,9 [-2,1; 3,8] p=0,5659

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

SOC ^a PT ^a	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
	n/N ^b (%)	n/N ^b (%)	RR [95 %-KI] p-Wert ^c	OR [95 %-KI] p-Wert ^c	RD [95 %-KI] p-Wert ^c
<p>Quelle: Table 14.3.1.5.4 (57)</p> <p>a: MedDRA-Version 24.0.</p> <p>b: Anzahl der Patienten im SAS.</p> <p>c: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet anhand der unstratifizierten Approximation der Normalverteilung nach der Wald-Methode.</p> <p>CoV: Coronavirus; KI: Konfidenzintervall; MedDRA: <i>Medical Dictionary for Regulatory Activities</i>; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; OR: <i>Odds Ratio</i>; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RD: Risikodifferenz; RR: Relatives Risiko; SARS: <i>Severe Acute Respiratory Syndrome</i>; SAS: <i>Safety-Analysis-Set</i>; SOC: Systemorganklasse (<i>System Organ Class</i>); vs.: Versus</p>					

Für die SOC Erkrankungen des Gastrointestinaltrakts zeigte sich in der Kategorie UE jeglichen Schweregrades zu Woche 26 ein statistisch signifikanter Unterschied zuungunsten von Abrocitinib (RR [95 %-KI]: 6,42 [3,46; 11,91]; p-Wert < 0,0001). Die dazugehörige PT Übelkeit trat bei 19,3 % der Patienten im Abrocitinib-Arm und bei 2,2 % der Patienten im Dupilumab-Arm auf. Die PT Übelkeit machte somit den überwiegenden Anteil der Patienten mit UE für die SOC Erkrankungen des Gastrointestinaltrakts (19,3 % im Abrocitinib-Arm versus 3,0 % im Dupilumab-Arm) aus. Der Unterschied in der PT Übelkeit jeglichen Schweregrades war statistisch signifikant zuungunsten von Abrocitinib (RR [95 %-KI]: 8,82 [4,31; 18,07]; p-Wert < 0,0001). Alle aufgetretenen Ereignisse waren nicht schwer (Anteil der Patienten mit mindestens einem nicht schweren UE aus der SOC Erkrankungen des Gastrointestinaltrakts: 19,3 % im Abrocitinib-Arm versus 3,0 % im Dupilumab-Arm; nicht schwere UE mit der PT Übelkeit: 19,3 % im Abrocitinib-Arm versus 2,2 % im Dupilumab-Arm; siehe Anhang 4-G).

Für die SOC Erkrankungen des Nervensystems zeigte sich in der Kategorie UE jeglichen Schweregrades zu Woche 26 ein statistisch signifikanter Unterschied zuungunsten von Abrocitinib (RR [95 %-KI]: 2,05 [1,33; 3,18]; p-Wert = 0,0012). Die dazugehörige PT Kopfschmerzen trat bei 13,0 % der Patienten im Abrocitinib-Arm und bei 6,6 % der Patienten im Dupilumab-Arm auf. Die PT Kopfschmerzen machte somit den überwiegenden Anteil der Patienten mit UE für die SOC Erkrankungen des Nervensystems (15,2 % im Abrocitinib-Arm versus 7,4 % im Dupilumab-Arm) aus. Der Unterschied in der PT Kopfschmerzen jeglichen Schweregrades war statistisch signifikant zuungunsten von Abrocitinib (RR [95 %-KI]: 1,97 [1,23; 3,16]; p-Wert = 0,0045). Die aufgetretenen Ereignisse waren größtenteils nicht schwer (Anteil der Patienten mit mindestens einem nicht schweren UE aus der SOC Erkrankungen des Nervensystems: 14,9 % im Abrocitinib-Arm versus 7,4 % im Dupilumab-Arm; nicht schwere UE mit der PT Kopfschmerzen: 12,7 % im Abrocitinib-Arm versus 6,6 % im Dupilumab-Arm; siehe Anhang 4-G).

Für die SOC Erkrankungen der Haut und des Unterhautgewebes zeigte sich in der Kategorie UE jeglichen Schweregrades zu Woche 26 ein statistisch signifikanter Unterschied zuungunsten von Abrocitinib (RR [95 %-KI]: 2,67 [1,69; 4,22]; p-Wert < 0,0001). Die dazugehörige PT Akne trat bei 12,7 % der Patienten im Abrocitinib-Arm und bei 2,7 % der Patienten im Dupilumab-Arm auf. Die PT Akne machte somit den überwiegenden Anteil der Patienten mit UE für die SOC Erkrankungen der Haut und des Unterhautgewebes (16,9 % im Abrocitinib-Arm vs. 6,3 % im Dupilumab-Arm) aus. Der Unterschied in der PT Akne jeglichen Schweregrades war statistisch signifikant zuungunsten von Abrocitinib (RR [95 %-KI]: 4,64 [2,38; 9,05]; p-Wert < 0,0001). Die aufgetretenen Ereignisse waren größtenteils nicht schwer (Anteil der Patienten mit mindestens einem nicht schweren UE aus der SOC Erkrankungen der Haut und des Unterhautgewebes: 16,6 % im Abrocitinib-Arm versus 6,3 % im Dupilumab-Arm; nicht schwere UE mit der PT Akne: 12,7 % im Abrocitinib-Arm versus 2,7 % im Dupilumab-Arm; siehe Anhang 4-G).

Für die PT Konjunktivitis aus der SOC Infektionen und parasitäre Erkrankungen zeigte sich in der Kategorie UE jeglichen Schweregrades zu Woche 26 ein statistisch signifikanter

Unterschied zugunsten von Abrocitinib (RR [95 %-KI]: 0,23 [0,11; 0,49]; p-Wert = 0,0001). Unter Abrocitinib lag der Anteil der Patienten mit aufgetretener Konjunktivitis bei 2,2 %; unter Dupilumab lag der entsprechende Anteil der Patienten bei 9,6 %. Alle aufgetretenen Ereignisse waren nicht schwer (Anteil der Patienten mit mindestens einem nicht schweren UE mit der PT Konjunktivitis: 2,2 % im Abrocitinib-Arm versus 9,6 % im Dupilumab-Arm; siehe Anhang 4-G).

Für die PT Follikulitis aus der SOC Infektionen und parasitäre Erkrankungen zeigte sich in der Kategorie UE jeglichen Schweregrades zu Woche 26 ein statistisch signifikanter Unterschied zuungunsten von Abrocitinib (RR [95 %-KI]: 4,03 [1,15; 14,17]; p-Wert = 0,0296). Unter Abrocitinib lag der Anteil der Patienten mit aufgetretener Follikulitis bei 3,3 %; unter Dupilumab lag der entsprechende Anteil der Patienten bei 0,8 %. Alle aufgetretenen Ereignisse waren nicht schwer (Anteil der Patienten mit mindestens einem nicht schweren UE mit der PT Follikulitis: 3,3 % im Abrocitinib-Arm versus 0,8 % im Dupilumab-Arm; siehe Anhang 4-G).

Für die PT Abnahme der natürlichen Killerzellen aus der SOC Untersuchungen zeigte sich in der Kategorie UE jeglichen Schweregrades zu Woche 26 ein statistisch signifikanter Unterschied zuungunsten von Abrocitinib (RR [95 %-KI]: 20,19 [1,18; 344,43]; p-Wert = 0,0378). Unter Abrocitinib lag der Anteil der Patienten mit aufgetretener Abnahme der natürlichen Killerzellen bei 2,8 %; unter Dupilumab trat bei keinem der Patienten eine Abnahme der natürlichen Killerzellen auf. Alle aufgetretenen Ereignisse waren nicht schwer (Anteil der Patienten mit mindestens einem nicht schweren UE mit der PT Abnahme der natürlichen Killerzellen: 2,8 % im Abrocitinib-Arm; siehe Anhang 4-G).

Zu Woche 26 traten keine schweren UE und SUE nach SOC und PT in mindestens 5 % der Patienten in mindestens einem Behandlungsarm auf.

Tabelle 4-97: Ergebnisse für UE, die zum Therapieabbruch geführt haben, nach SOC und PT aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

SOC ^a PT ^a	Abrocitinib 200 mg n/N ^b (%)	Dupilumab 300 mg n/N ^b (%)
Erkrankungen des Blutes und des Lymphsystems		
Jegliche PT	3 (0,8)	0
Anämie	1 (0,3)	0
Leukopenie	1 (0,3)	0
Normozytäre Anämie	1 (0,3)	0
Thrombozytopenie	1 (0,3)	0
Augenerkrankungen		
Jegliche PT	0	1 (0,3)
Allergische Konjunktivitis	0	1 (0,3)
Trockene Augen	0	1 (0,3)
Erkrankungen des Gastrointestinaltrakts		
Jegliche PT	2 (0,6)	1 (0,3)
Durchfall	0	1 (0,3)
Erbrechen	2 (0,6)	0
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort		
Jegliche PT	1 (0,3)	0
Erschöpfung	1 (0,3)	0
Untersuchungen		
Jegliche PT	0	2 (0,5)
Alanin-Aminotransferase erhöht	0	1 (0,3)
Hepatische Enzyme erhöht	0	1 (0,3)
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen		
Jegliche PT	0	1 (0,3)
Rhabdomyolyse	0	1 (0,3)
Erkrankungen des Nervensystems		
Jegliche PT	1 (0,3)	0
Kopfschmerzen	1 (0,3)	0
Schwangerschaft, Wochenbett und perinatale Erkrankungen		
Jegliche PT	0	1 (0,3)
Schwangerschaft	0	1 (0,3)

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

SOC ^a PT ^a	Abrocitinib 200 mg n/N ^b (%)	Dupilumab 300 mg n/N ^b (%)
Psychiatrische Erkrankungen		
Jegliche PT	1 (0,3)	2 (0,5)
Depressive Stimmung	0	1 (0,3)
Depression	1 (0,3)	0
Suizidgedanken	0	1 (0,3)
Erkrankungen der Nieren und Harnwege		
Jegliche PT	0	1 (0,3)
Nephrotisches Syndrom	0	1 (0,3)
Erkrankungen der Haut und des Unterhautgewebes		
Jegliche PT	1 (0,3)	0
Dermatitis atopisch	1 (0,3)	0
Quelle: Table 14.3.1.5.9 (57) a: MedDRA-Version 24.0. b: Anzahl der Patienten im SAS. MedDRA: <i>Medical Dictionary for Regulatory Activities</i> ; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); SAS: <i>Safety-Analysis-Set</i> ; SOC: Systemorganklasse (<i>System Organ Class</i>)		

Tabelle 4-98: Ergebnisse für spezifische UE aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg		
n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b	OR [95 %-KI] p-Wert ^b	RD [95 %-KI] p-Wert ^b
Superinfektionen				
31/362 (8,6)	22/365 (6,0)	1,42 [0,84; 2,41] p=0,1912	1,46 [0,83; 2,57] p=0,1906	2,54 [-1,24; 6,31] p=0,1882
Herpes zoster				
9/362 (2,5)	2/365 (0,5)	4,54 [0,99; 20,85] p=0,0520	4,63 [0,99; 21,57] p=0,0511	1,94 [0,16; 3,71] p=0,0322
Konjunktivitis				
10/362 (2,8)	39/365 (10,7)	0,26 [0,13; 0,51] p<0,0001	0,24 [0,12; 0,48] p<0,0001	-7,92 [-11,51; -4,33] p<0,0001
Akne				
48/362 (13,3)	11/365 (3,0)	4,40 [2,32; 8,33] p<0,0001	4,92 [2,51; 9,64] p<0,0001	10,25 [6,34; 14,16] p<0,0001
Follikulitis				
11/362 (3,0)	3/365 (0,8)	3,70 [1,04; 13,14] p=0,0433	3,78 [1,05; 13,67] p=0,0425	2,22 [0,22; 4,21] p=0,0295
Quelle: Table 14.3.1.5.8, Table 14.3.1.5.10.1, Table 14.3.1.5.10.4, Table 14.3.1.6.10.2, Table 14.3.1.6.10.3 (57)				
a: Anzahl der Patienten im SAS.				
b: RR, OR und RD inklusive 95 %-KI und p-Wert berechnet anhand der unstratifzierten Approximation der Normalverteilung nach der Wald-Methode.				
KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; OR: <i>Odds Ratio</i> ; RD: Risikodifferenz; RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus				

Zu Woche 26 zeigte sich für Konjunktivitis in der Kategorie spezifische UE jeglichen Schweregrades ein statistisch signifikanter Unterschied zugunsten von Abrocitinib (RR [95 %-KI]: 0,26 [0,13; 0,51]; p-Wert < 0,0001). Für die spezifischen UE Akne (RR [95 %-KI]: 4,40 [2,32; 8,33]; p-Wert < 0,0001) und Follikulitis (RR [95 %-KI]: 3,70 [1,04; 13,14]; p-Wert = 0,0433) zeigte sich ein statistisch signifikanter Unterschied zuungunsten von Abrocitinib.

Bis Woche 26 traten keine schweren oder schwerwiegenden spezifischen UE auf.

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

Es wurde keine Meta-Analyse durchgeführt, da die Voraussetzungen dafür nicht erfüllt sind.

4.3.1.3.2 Subgruppenanalysen – RCT

Für die Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.¹⁶

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

- Subgruppenanalysen sind nur für die Merkmale (z.B. Alter) durchzuführen, bei denen die resultierenden Subgruppen jeweils mindestens 10 Patienten umfassen.
- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.
- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ($p < 0,05$) dargestellt werden.
- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.
- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.

¹⁶ unbesetzt

- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmodifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.
- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.
- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

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

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

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

Tabelle 4-99: Matrix der durchgeführten Subgruppenanalysen

Kategorie Endpunkt Studie	Alter	Geschlecht	Abstammung	Region	Krankheitsschwere zu <i>Baseline</i>	Vorbehandlung mit Ciclosporin	Gewicht	Krankheitsdauer	EASI zu <i>Baseline</i>	Betroffene BSA (%) zu <i>Baseline</i>	Vorherige AD- Medikation	Peak Pruritus NRS zu <i>Baseline</i>
	Morbidität											
EASI												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
SCORAD												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
POEM-Gesamtscore												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
Peak Pruritus NRS												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
SCORAD VAS für Juckreiz												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
MOS Sleep Scale												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												

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

Kategorie Endpunkt Studie	Alter	Geschlecht	Abstammung	Region	Krankheitsschwere zu <i>Baseline</i>	Vorbehandlung mit Ciclosporin	Gewicht	Krankheitsdauer	EASI zu <i>Baseline</i>	Betroffene BSA (%) zu <i>Baseline</i>	Vorherige AD- Medikation	Peak Pruritus NRS zu <i>Baseline</i>
SCORAD VAS für Schlafstörungen												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
Skin Pain NRS												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
EQ-5D VAS												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
Gesundheitsbezogene Lebensqualität												
DLQI												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
Verträglichkeit												
JADE	•	•	•	•	•	•	•	•	•	•	•	○
DARE												
<p>•: <i>A priori</i> geplante Subgruppenanalyse. ○: <i>Posthoc</i> durchgeführte Subgruppenanalyse.</p> <p>DLQI: <i>Dermatology Life Quality Index</i>; EASI: <i>Eczema Area and Severity Index</i>; EQ-5D VAS: <i>European Quality of Life Group 5 Dimensions</i>; MOS: <i>Medical Outcomes Study</i>; NRS: <i>Numerical Rating Scale</i>; POEM: <i>Patient-Oriented Eczema Measure</i>; SCORAD: <i>Scoring Atopic Dermatitis</i>; VAS: <i>Visuelle Analogskala</i></p>												

Die Subgruppenanalysen werden für die Studie JADE DARE für die Endpunkte EASI, SCORAD, POEM, *Peak Pruritus* NRS, SCORAD VAS für Juckreiz, MOS *Sleep Scale*, SCORAD VAS für Schlafstörungen, *Skin Pain* NRS, EQ-5D VAS, DLQI und Verträglichkeit dargestellt.

Folgende Subgruppenanalysen waren in der Studie JADE DARE präspezifiziert und sind gemäß Dossiervorlage gefordert:

- Alter in Jahren (Kategorie 1: < 40 versus ≥ 40 ; Kategorie 2: < 65 versus ≥ 65)
- Geschlecht (männlich versus weiblich)
- Abstammung (kaukasisch versus schwarz oder afroamerikanisch versus asiatisch versus andere)
- Region der Studieneinschreibung (USA/ Kanada/ Australien versus Europa versus Asien)
- Gewicht in kg (< 70 versus 70-100 versus > 100 kg)
- Krankheitsdauer in Jahren (< 26 versus ≥ 26)
- Krankheitsschwere zu *Baseline* (IGA = 3 versus IGA = 4)
- EASI-Score zu *Baseline* (16-25 versus > 25)
- Betroffene BSA (%) zu *Baseline* (10-30 % versus > 30 -50 % versus > 50 %)
- Vorbehandlung mit Ciclosporin (ja versus nein)
- Vorbehandlung mit AD-Medikation (nur topische Mittel versus systemische Mittel)

Die Subgruppenanalyse nach *Peak Pruritus* NRS zu *Baseline* (4-6 versus ≥ 7) wurde für die Studie JADE DARE *posthoc* durchgeführt und war nicht präspezifiziert.

Stellen Sie anschließend in Tabelle 4-100 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante ($p < 0,05$) Interaktionsterme.

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

Tabelle 4-100: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für JADE DARE

Endpunkt Studie	Alter (Kategorie 1)	Alter (Kategorie 2)	Geschlecht	Abstammung	Region	Krankheitsschwere zu Baseline	Vorbehandlung mit Ciclosporin	Gewicht	Krankheitsdauer	EASI zu Baseline	Betroffene BSA (%) zu Baseline	Vorherige AD-Medikation	Peak Pruritus NRS zu Baseline
EASI-75													
JADE DARE	p=0,0951	p=0,3406	p=0,0280	p=0,7979	p=0,5379	p=0,4236	p=0,9774	p=0,1945	p=0,6670	p=0,8047	p=0,7288	p=0,6934	p=0,7079
EASI-90													
JADE DARE	p=0,6895	p=0,7423	p=0,2428	p=0,6013	p=0,5107	p=0,4499	p=0,4840	p=0,9402	p=0,7685	p=0,9290	p=0,5795	p=0,9876	p=0,5129
SCORAD-75													
JADE DARE	p=0,2069	p=0,7717	p=0,1104	p=0,5541	p=0,3197	p=0,0893	p=0,8468	p=0,2141	p=0,9023	p=0,8380	p=0,5201	p=0,9673	p=0,3469
SCORAD-90													
JADE DARE	p=0,5685	p=0,3700	p=0,1356	p=0,5375	p=0,1251	p=0,4942	p=0,5056	p=0,3544	p=0,6921	p=0,6057	p=0,4215	p=0,5561	p=0,9542
Anteil der Patienten mit POEM-Gesamtscore = 0-2													
JADE DARE	p=0,1194	p=0,2006	p=0,5289	p=0,8964	p=0,6651	p=0,6804	p=0,4725	p=0,1801	p=0,7500	p=0,0746	p=0,3059	p=0,1437	p=0,3124
Anteil der Patienten mit einer Verbesserung um ≥ 5 Punkte gegenüber Baseline im POEM-Gesamtscore													
JADE DARE	p<0,0009	p=0,8870	p=0,3931	p=0,2692	p=0,7478	p=0,7829	p=0,7696	p=0,5173	p=0,1430	p=0,4393	p=0,4428	p=0,2312	p=0,1616

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

Endpunkt Studie	Alter (Kategorie 1)	Alter (Kategorie 2)	Geschlecht	Abstammung	Region	Krankheitsschwere zu Baseline	Vorbehandlung mit Ciclosporin	Gewicht	Krankheitsdauer	EASI zu Baseline	Betroffene BSA (%) zu Baseline	Vorherige AD-Medikation	Peak Pruritus NRS zu Baseline
Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber Baseline in der Peak Pruritus NRS													
JADE	p=0,1958	p=0,9232	p=0,2229	p=0,2597	p=0,1556	p=0,8374	p=0,9615	p=0,5631	p=0,1407	p=0,9914	p=0,8503	p=0,3973	p=0,4363
DARE													
Anteil der Patienten mit einer Verbesserung um ≥ 2 Punkte gegenüber Baseline in der SCORAD VAS für Juckreiz													
JADE	p=0,0903	p=0,9781	p=0,3857	p=0,5656	p=0,3963	p=0,3454	p=0,2309	p=0,6038	p=0,9401	p=0,3352	p=0,2423	p=0,1165	p=0,7145
DARE													
Anteil der Patienten mit einer Verbesserung um ≥ 15 Punkte gegenüber Baseline im MOS Sleep Scale Score Index-1													
JADE	p=0,3940	p=0,3637	p=0,3725	p=0,4172	p=0,3894	p=0,1335	p=0,5276	p=0,1814	p=0,5742	p=0,3034	p=0,2529	p=0,3797	p=0,9030
DARE													
Anteil der Patienten mit einer Verbesserung um ≥ 15 Punkte gegenüber Baseline im MOS Sleep Scale Score Index-2													
JADE	p=0,5610	p=0,5482	p=0,3603	p=0,3260	p=0,6552	p=0,6161	p=0,3725	p=0,1229	p=0,8907	p=0,0687	p=0,2762	p=0,6937	p=0,7081
DARE													
Anteil der Patienten mit einer Verbesserung um ≥ 2 Punkte gegenüber Baseline in der SCORAD VAS für Schlafstörungen													
JADE	p=0,0459	p=0,9325	p=0,0775	p=0,4097	p=0,8875	p=0,5640	p=0,9960	p=0,3275	p=0,1886	p=0,6615	p=0,7363	p=0,5060	p=0,7786
DARE													
Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber Baseline in der Skin Pain NRS													
JADE	p=0,0077	p=0,6366	p=0,5105	p=0,8161	p=0,1887	p=0,7312	p=0,8239	p=0,4450	p=0,4446	p=0,0129	p=0,2829	p=0,7597	p=0,8456
DARE													
Anteil der Patienten mit einer Verbesserung um ≥ 15 Punkte gegenüber Baseline in der EQ-5D VAS													
JADE	p=0,5001	p=0,4773	p=0,5382	p=0,3886	p=0,2624	p=0,6435	p=0,7114	p=0,7993	p=0,1313	p=0,1478	p=0,8397	p=0,4218	p=0,3376
DARE													

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

Endpunkt Studie	Alter (Kategorie 1)	Alter (Kategorie 2)	Geschlecht	Abstammung	Region	Krankheitsschwere zu Baseline	Vorbehandlung mit Ciclosporin	Gewicht	Krankheitsdauer	EASI zu Baseline	Betroffene BSA (%) zu Baseline	Vorherige AD-Medikation	Peak Pruritus NRS zu Baseline
Anteil der Patienten mit DLQI-Gesamtscore von 0 oder 1													
JADE	p=0,0807	p=0,3704	p=0,5958	p=0,1397	p=0,8316	p=0,5303	p=0,7545	p=0,5957	p=0,7757	p=0,1012	p=0,7131	p=0,8143	p=0,2106
DARE													
Anteil der Patienten mit einer Verbesserung um ≥ 5 Punkte gegenüber Baseline im DLQI													
JADE	p=0,0733	p=0,6291	p=0,2304	p=0,7485	p=0,7941	p=0,9183	p=0,8908	p=0,2055	p=0,4041	p=0,0668	p=0,4202	p=0,4812	p=0,7479
DARE													
UE jeglichen Schweregrades													
JADE	p=0,5042	p=0,0270	p=0,6631	p=0,8236	p=0,9036	p=0,8701	p=0,2343	p=0,9647	p=0,5818	p=0,6971	p=0,4238	p=0,6515	p=0,9473
DARE													
Schwere UE													
JADE	p=0,4442	p=0,9746	p=0,2392	p=0,9140	p=0,8432	p=0,4201	p=0,9626	p=0,2866	p=0,6625	p=0,8295	p=0,0913	p=0,2278	p=0,6107
DARE													
SUE													
JADE	p=0,6669	p=0,9338	p=0,9268	p=0,9908	p=0,9862	p=0,9783	p=0,8605	p=0,8662	p=0,5428	p=0,6546	p=0,4082	p=0,0783	p=0,7687
DARE													
UE, die zum Therapieabbruch führten													
JADE	p=0,1535	p=0,9371	p=0,4137	p=0,9256	p=0,5066	p=0,3041	p=0,8585	p=0,8244	p=0,3635	p=0,3380	p=0,6067	p=0,6323	p=0,9485
DARE													
Häufige UE nach SOC und PT - Erkrankungen des Gastrointestinaltrakts													
JADE	p=0,0148	p=0,1717	p<0,0005	p=0,9617	p=0,0039	p=0,4477	p=0,0922	p=0,0173	p=0,2029	p=0,8434	p=0,4972	p=0,0544	p=0,8019
DARE													

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

Endpunkt Studie	Alter (Kategorie 1)	Alter (Kategorie 2)	Geschlecht	Abstammung	Region	Krankheitsschwere zu Baseline	Vorbehandlung mit Ciclosporin	Gewicht	Krankheitsdauer	EASI zu Baseline	Betroffene BSA (%) zu Baseline	Vorherige AD-Medikation	Peak Pruritus NRS zu Baseline
Häufige UE nach SOC und PT - Erkrankungen des Gastrointestinaltrakts: Übelkeit													
JADE	p=0,0132	p=0,1452	p<0,0006	p=0,9505	p=0,0019	p=0,7016	p=0,1189	p=0,0092	p=0,1439	p=0,7611	p=0,3398	p=0,0663	p=0,7776
DARE													
Häufige UE nach SOC und PT - Erkrankungen des Gastrointestinaltrakts: Erbrechen													
JADE	p=0,2587	p=0,9175	p=0,0936	p=0,9776	p=0,8619	p=0,9906	p=0,7718	p=0,1135	p=0,0339	p=0,7877	p=0,9452	p=0,2463	p=0,4769
DARE													
Häufige UE nach SOC und PT - Infektionen und parasitäre Erkrankungen: Konjunktivitis													
JADE	p=0,3803	NA	p=0,4937	p=0,0592	p=0,2628	p=0,2401	p=0,0867	p=0,6025	p=0,0308	p=0,3185	p=0,3154	p=0,7727	p=0,3673
DARE													
Häufige UE nach SOC und PT - Infektionen und parasitäre Erkrankungen: Follikulitis													
JADE	p=0,7526	p=0,8040	p=0,0729	p=0,6077	p=0,7124	p=0,1476	p=0,7582	p=0,1856	p=0,5480	p=0,2102	p=0,7224	p=0,8891	p=0,0134
DARE													
Häufige UE nach SOC und PT - Erkrankungen des Nervensystems													
JADE	p=0,6772	p=0,8339	p=0,7148	p=0,9278	p=0,9048	p=0,2787	p=0,2271	p=0,9870	p=0,5644	p=0,0056	p=0,0490	p=0,5041	p=0,3173
DARE													
Häufige UE nach SOC und PT - Erkrankungen des Nervensystems: Kopfschmerzen													
JADE	p=0,4922	p=0,7091	p=0,5872	p=0,8658	p=0,9443	p=0,6865	p=0,7697	p=0,8210	p=0,9000	p=0,0103	p=0,0555	p=0,9856	p=0,1768
DARE													
Häufige UE nach SOC und PT - Untersuchungen: Kreatinphosphokinase im Blut erhöht													
JADE	p=0,0468	p=0,9716	p=0,4755	p=0,8989	p=0,9243	p=0,5547	p=0,3250	p=0,5176	p=0,6188	p=0,5514	p=0,2399	p=0,3129	p=0,1228
DARE													

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

Endpunkt Studie	Alter (Kategorie 1)	Alter (Kategorie 2)	Geschlecht	Abstammung	Region	Krankheitsschwere zu Baseline	Vorbehandlung mit Ciclosporin	Gewicht	Krankheitsdauer	EASI zu Baseline	Betroffene BSA (%) zu Baseline	Vorherige AD-Medikation	Peak Pruritus NRS zu Baseline
Häufige UE nach SOC und PT - Erkrankungen der Haut und des Unterhautgewebes: Akne													
JADE	p=0,7370	NA	p=0,7950	p=0,1873	p=0,1633	p=0,1346	p=0,8429	p=0,6330	p=0,0458	p=0,6243	p=0,5795	p=0,8281	p=0,9248
DARE													
Häufige UE nach SOC und PT - Erkrankungen der Haut und des Unterhautgewebes: Dermatitis atopisch													
JADE	p=0,0481	NA	p=0,5105	p=0,9095	p=0,5650	p=0,3289	p=0,1643	p=0,3776	p=0,7317	p=0,6937	p=0,6344	p=0,3327	p=0,0212
DARE													
Spezifische UE - Superinfektionen													
JADE	p=0,8103	p=0,7502	p=0,4216	p=0,8119	p=0,9965	p=0,6238	p=0,6622	p=0,1061	p=0,3944	p=0,2103	p=0,4403	p=0,2329	p=0,1899
DARE													
Spezifische UE – Herpes Zoster													
JADE	p=0,4405	p=0,6698	p=0,8457	p=0,7564	p=0,9636	p=0,2721	p=0,5160	p=0,4008	p=0,4379	p=0,1675	p=0,4239	p=0,9671	p=0,0571
DARE													
Spezifische UE – Akne													
JADE	p=0,7300	p=0,0601	p=0,5116	p=0,1881	p=0,1721	p=0,1911	p=0,7933	p=0,8508	p=0,1090	p=0,7282	p=0,4888	p=0,7302	p=0,8001
DARE													
Spezifische UE - Follikulitis													
JADE	p=0,9812	p=0,8331	p=0,1122	p=0,6766	p=0,5893	p=0,2092	p=0,8368	p=0,1597	p=0,6790	p=0,3039	p=0,8770	p=0,6669	p=0,0234
DARE													
Spezifische UE - Konjunktivitis													
JADE	p=0,8350	p=0,3731	p=0,3281	p=0,0408	p=0,1573	p=0,0225	p=0,3211	p=0,2836	p=0,0177	p=0,1658	p=0,5188	p=0,5396	p=0,2035
DARE													

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

Endpunkt	Alter (Kategorie 1)	Alter (Kategorie 2)	Geschlecht	Abstammung	Region	Krankheitsschwere zu Baseline	Vorbehandlung mit Ciclosporin	Gewicht	Krankheitsdauer	EASI zu Baseline	Betroffene BSA (%) zu Baseline	Vorherige AD-Medikation	Peak Pruritus NRS zu Baseline
Studie													
<p>Quelle: (57)</p> <p>a: Das Subgruppenmerkmal Alter (Kategorie 1) betrachtet die Altersgruppen < 40 Jahre vs. ≥ 40 Jahre.</p> <p>b: Das Subgruppenmerkmal Alter (Kategorie 2) betrachtet die Altersgruppen < 65 Jahre vs. ≥ 65 Jahre.</p> <p>AD: Atopische Dermatitis; BSA: Körperoberfläche (<i>Body Surface Area</i>); DLQI: <i>Dermatology Life Quality Index</i>; EASI: <i>Eczema Area and Severity Index</i>; EQ-5D: <i>European Quality of Life Group 5 Dimensions</i>; MOS: <i>Medical Outcomes Study</i>; NA: Nicht anwendbar; NRS: <i>Numerical Rating Scale</i>; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); POEM: <i>Patient-Oriented Eczema Measure</i>; SCORAD: <i>Scoring Atopic Dermatitis</i>; SOC: Systemorganklasse (<i>System Organ Class</i>); SUE: Schwerwiegendes unerwünschtes UE; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala; vs.: Versus</p>													

Stellen Sie schließlich alle Subgruppenergebnisse dar.

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

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Tabelle 4-101: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: EASI-75

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Geschlecht, p=0,0280			
Männlich	137/193 (71,0)	134/204 (65,7)	1,08 [0,95; 1,24] p=0,2568
Weiblich	117/169 (69,2)	127/161 (78,9)	0,88 [0,77; 1,00] p=0,0464

Quelle: Table 14.2.2.6.1.2 (57)

a: Anzahl der Patienten im FAS in der jeweiligen Subgruppenkategorie, die zu *Baseline* das entsprechende *Response*-Kriterium erreicht haben.

b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.

EASI: *Eczema Area and Severity Index*; FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; RR: Relatives Risiko

Es zeigte sich für den Endpunkt „EASI-75“ ein Beleg für Effektmodifikation durch das Merkmal „Geschlecht“. In der Subgruppe „weiblich“ ist der Effekt gleichgerichtet zur Hauptanalyse und zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „männlich“ ist der Effekt nicht gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Tabelle 4-102: Detaillierte Ergebnisse der Subgruppenanalyse für den Anteil der Patienten mit einer Verbesserung um ≥ 5 Punkte gegenüber *Baseline* im POEM-Gesamtscore mit statistisch signifikantem Interaktionsterm ($p < 0,05$)

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Alter (Kategorie 1), p=0,0009			
< 40 Jahre	158/225 (70,2)	200/246 (81,3)	0,86 [0,78; 0,96] p=0,0058
≥ 40 Jahre	113/131 (86,3)	89/117 (76,1)	1,13 [1,00; 1,28] p=0,0442
Quelle: Table 14.2.10.9.5 (57)			
a: Anzahl der Patienten im FAS in der jeweiligen Subgruppenkategorie, die zu <i>Baseline</i> das entsprechende <i>Response</i> -Kriterium erreicht haben.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
FAS: <i>Full-Analysis-Set</i> ; KI: Konfidenzintervall; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; POEM: <i>Patient-Oriented Eczema Measures</i> ; RR: Relatives Risiko; vs.: Versus			

Es zeigte sich für den Endpunkt POEM-Gesamtscore (Verbesserung um ≥ 5 Punkte gegenüber *Baseline*) ein Beleg für Effektmodifikation durch das Merkmal „Alter (Kategorie 1)“. In der Subgruppe „< 40 Jahre“ ist der Effekt gleichgerichtet zur Hauptanalyse und zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „ ≥ 40 Jahre“ ist der Effekt nicht gleichgerichtet zur Hauptanalyse und zuungunsten von Abrocitinib statistisch signifikant.

Tabelle 4-103: Detaillierte Ergebnisse der Subgruppenanalyse für Anteil der Patienten mit einer Verbesserung um ≥ 4 Punkte gegenüber *Baseline* in der *Skin Pain* NRS mit statistisch signifikantem Interaktionsterm ($p < 0,05$)

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Alter (Kategorie 1), p=0,0077			
< 40 Jahre	118/204 (57,8)	133/212 (62,7)	0,92 [0,79; 1,08] p=0,3091
≥ 40 Jahre	87/112 (77,7)	69/113 (61,1)	1,27 [1,07; 1,52] p=0,0079
EASI zu <i>Baseline</i>, p=0,0129			
16-25 Punkte	90/156 (57,7)	106/164 (64,6)	0,89 [0,75; 1,06] p=0,2050
> 25 Punkte	111/153 (72,5)	95/158 (60,1)	1,21 [1,03; 1,42] p=0,0215

Quelle: Table 14.2.12.4.2 (57)

a: Anzahl der Patienten im FAS in der jeweiligen Subgruppenkategorie, die zu *Baseline* das entsprechende *Response*-Kriterium erreicht haben.

b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.

EASI: *Eczema Area and Severity Index*; FAS: *Full-Analysis-Set*; KI: Konfidenzintervall; mg: Milligramm;

N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis;

NRS: *Numerical Rating Scale*; RR: Relatives Risiko; vs.: Versus

Es zeigte sich für den Endpunkt *Skin Pain* NRS (Verbesserung um ≥ 4 Punkte gegenüber *Baseline*) ein Beleg für Effektmodifikation durch das Merkmal „Alter (Kategorie 1)“. In der Subgruppe „ ≥ 40 Jahre“ ist der Effekt gleichgerichtet zur Hauptanalyse und zugunsten von Abrocitinib statistisch signifikant. In der Subgruppe „< 40 Jahre“ ist der Effekt nicht gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Es zeigte sich für den Endpunkt *Skin Pain* NRS (Verbesserung um ≥ 4 Punkte gegenüber *Baseline*) ein Beleg für Effektmodifikation durch das Merkmal „EASI zu *Baseline*“. In der Subgruppe „> 25 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und zugunsten von Abrocitinib statistisch signifikant. In der Subgruppe „16-25 Punkte“ ist der Effekt nicht gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Tabelle 4-104: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach SOC Erkrankungen des Gastrointestinaltrakts

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Alter (Kategorie 1), p=0,0148			
< 40 Jahre	53/230 (23,0)	7/247 (2,8)	8,13 [3,77; 17,52] p<0,0001
≥ 40 Jahre	17/132 (12,9)	4/118 (3,4)	3,80 [1,32; 10,97] p=0,0136
Geschlecht, p=0,0005			
Männlich	21/193 (10,9)	4/204 (2,0)	5,55 [1,94; 15,87] p=0,0014
Weiblich	49/169 (29,0)	7/161 (4,3)	6,67 [3,11; 14,29] p<0,0001
Region, p=0,0005			
Vereinigte Staaten/ Kanada/ Australien	39/177 (22,0)	6/195 (3,1)	7,16 [3,11; 16,51] p<0,0001
Europa	30/150 (20,0)	5/132 (3,8)	5,28 [2,11; 13,21] p=0,0004
Asien	1/17 (5,9)	0/19 (0,0)	2,29 [0,08; 64,21] p=0,6252
Latein-amerika	0/18 (0,0)	0/19 (0,0)	NA
Gewicht, p=0,0173			
< 70 kg	35/132 (26,5)	3/136 (2,2)	12,02 [3,79; 38,13] p<0,0001
70 bis 100 kg	31/196 (15,8)	4/184 (2,2)	7,28 [2,62; 20,21] p=0,0001
> 100 kg	4/34 (11,8)	4/45 (8,9)	1,32 [0,36; 4,92] p=0,6755

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
<p>Quelle: Table 14.3.1.6.4 (57)</p> <p>a: Anzahl der Patienten im SAS.</p> <p>b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.</p> <p>kg: Kilogramm; KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i>; SOC: Systemorganklasse (<i>System Organ Class</i>); vs.: Versus</p>			

Es zeigte sich für den Endpunkt UE nach SOC Erkrankungen des Gastrointestinaltrakts ein Beleg für Effektmodifikation durch das Merkmal „Alter (Kategorie 1)“. In den Subgruppen „< 40 Jahre“ und „≥ 40 Jahre“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant.

Es zeigte sich für den Endpunkt UE nach SOC Erkrankungen des Gastrointestinaltrakts ein Beleg für Effektmodifikation durch das Merkmal „Geschlecht“. In den Subgruppen „weiblich“ und „männlich“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant.

Es zeigte sich für den Endpunkt UE nach SOC Erkrankungen des Gastrointestinaltrakts ein Beleg für Effektmodifikation durch das Merkmal „Region“. In den Subgruppen „Vereinigte Staaten/ Kanada/ Australien“ und „Europa“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „Asien“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant. In der Subgruppe „Lateinamerika“ sind in keiner der beiden Behandlungsgruppen entsprechende Ereignisse aufgetreten.

Es zeigte sich für den Endpunkt UE nach SOC Erkrankungen des Gastrointestinaltrakts ein Beleg für Effektmodifikation durch das Merkmal „Gewicht“. In den Subgruppen „< 70 kg“ und „70 bis 100 kg“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „> 100 kg“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Tabelle 4-105: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach SOC Erkrankungen des Nervensystems

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
EASI zu <i>Baseline</i>, p=0,0056			
16-25 Punkte	21/181 (11,6)	19/183 (10,4)	1,12 [0,62; 2,01] p=0,7101
> 25 Punkte	33/172 (19,2)	8/174 (4,6)	4,17 [1,98; 8,77] p=0,0002
Betroffene BSA (%) zu <i>Baseline</i>, p=0,0490			
10-30 %	13/122 (10,7)	13/133 (9,8)	1,09 [0,53; 2,26] p=0,8163
> 30-50 %	22/132 (16,7)	10/121 (8,3)	2,02 [1,00; 4,08] p=0,0514
> 50 %	20/108 (18,5)	4/111 (3,6)	5,14 [1,82; 14,54] p=0,0020
Quelle: Table 14.3.1.6.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
BSA: Körperoberfläche (<i>Body Surface Area</i>); EASI: <i>Eczema Area and Severity Index</i> ;			
KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; SOC: Systemorganklasse (<i>System Organ Class</i>); vs.: Versus			

Es zeigte sich für den Endpunkt UE nach SOC Erkrankungen des Nervensystems ein Beleg für Effektmodifikation durch das Merkmal „EASI zu *Baseline*“. In der Subgruppe „>25 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „16-25 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen nicht statistisch signifikant.

Es zeigte sich für den Endpunkt UE nach SOC Erkrankungen des Nervensystems ein Beleg für Effektmodifikation durch das Merkmal „Betroffene BSA (%) zu *Baseline*“. In der Subgruppe > 50 %“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant. In den Subgruppen

„10-30 %“ und „> 30-50 %“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Tabelle 4-106: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Konjunktivitis

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Krankheitsdauer, p=0,0308			
< 26 Jahre	3/222 (1,4)	12/220 (5,5)	0,25 [0,07; 0,87] p=0,0289
≥ 26 Jahre	5/140 (3,6)	23/145 (15,9)	0,23 [0,09; 0,58] p=0,0019
Quelle: Table 14.3.1.6.10.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus			

Es zeigte sich für den Endpunkt UE nach PT Konjunktivitis ein Beleg für Effektmodifikation durch das Merkmal „Krankheitsdauer“. In den Subgruppen „<26 Jahre“ und „≥ 26 Jahre“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zugunsten von Abrocitinib statistisch signifikant.

Tabelle 4-107: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Follikulitis

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Peak Pruritus NRS zu Baseline, p=0,0134			
4-6 Punkte	8/83 (9,6)	0/105 (0,0)	20,34 [1,19; 349,00] p=0,0378
≥ 7 Punkte	4/274 (1,5)	3/259 (1,2)	1,26 [0,28; 5,58] p=0,7604
Quelle: Table 14.3.1.6.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; NRS: <i>Numerical Rating Scale</i> ; ; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus			

Es zeigte sich für den Endpunkt UE nach PT Follikulitis ein Beleg für Effektmodifikation durch das Merkmal „Peak Pruritus NRS zu Baseline“. In der Subgruppe „4-6 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „≥ 7 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Tabelle 4-108: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Übelkeit

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Alter (Kategorie 1), p=0,0132			
< 40 Jahre	53/230 (23,0)	5/247 (2,0)	11,38 [4,63; 27,98] p<0,0001
≥ 40 Jahre	17/132 (12,9)	3/118 (2,5)	5,07 [1,52; 16,85] p=0,0082
Geschlecht, p=0,0006			
Männlich	21/193 (10,9)	2/204 (1,0)	11,10 [2,64; 46,70] p=0,0010
Weiblich	49/169 (29,0)	6/161 (3,7)	7,78 [3,43; 17,66] p<0,0001
Region, p=0,0019			
Vereinigte Staaten/ Kanada/ Australien	39/177 (22,0)	4/195 (2,1)	10,74 [3,29; 29,45] p<0,0001
Europa	30/150 (20,0)	4/132 (3,0)	6,60 [2,39; 18,24] p=0,0003
Asien	1/17 (5,9)	0/19 (0,0)	2,29 [0,08; 64,21] p=0,6252
Latein-amerika	0/18 (0,0)	0/19 (0,0)	NA
Gewicht, p=0,0092			
< 70 kg	35/132 (26,5)	1/136 (0,7)	36,06 [5,01; 259,43] p=0,0004
70 bis 100 kg	31/196 (15,8)	3/184 (1,6)	9,70 [3,02; 31,19] p=0,0001
> 100 kg	4/34 (11,8)	4/45 (8,9)	1,32 [0,36; 4,92] p=0,6755

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Quelle: Table 14.3.1.6.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
kg: Kilogramm; KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; NA: Nicht anwendbar; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus			

Es zeigte sich für den Endpunkt UE nach PT Übelkeit ein Beleg für Effektmodifikation durch das Merkmal „Alter (Kategorie 1)“. In den Subgruppen „< 40 Jahre“ und „≥ 40 Jahre“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant.

Es zeigte sich für den Endpunkt UE nach PT Übelkeit ein Beleg für Effektmodifikation durch das Merkmal „Geschlecht“. In den Subgruppen „weiblich“ und „männlich“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant.

Es zeigte sich für den Endpunkt UE nach PT Übelkeit ein Beleg für Effektmodifikation durch das Merkmal „Region“. In den Subgruppen „Vereinigte Staaten/ Kanada/ Australien“ und „Europa“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „Asien“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant. In der Subgruppe „Lateinamerika“ sind in keiner der beiden Behandlungsgruppen entsprechende Ereignisse aufgetreten.

Es zeigte sich für den Endpunkt UE nach PT Übelkeit ein Beleg für Effektmodifikation durch das Merkmal „Gewicht“. In den Subgruppen „< 70 kg“ und „70 bis 100 kg“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „> 100 kg“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Tabelle 4-109: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Erbrechen

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Krankheitsdauer, p=0,0339			
< 26 Jahre	10/222 (4,5)	3/220 (1,4)	3,30 [0,92; 11,84] p=0,0666
≥ 26 Jahre	1/140 (0,7)	3/145 (2,1)	0,35 [0,04; 3,28] p=0,3545
Quelle: Table 14.3.1.6.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko;			
SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus			

Es zeigte sich für den Endpunkt UE nach PT Erbrechen ein Beleg für Effektmodifikation durch das Merkmal „Krankheitsdauer“. Für keine der beiden Subgruppen „< 26 Jahre“ und „≥ 26 Jahre“ zeigte sich ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen.

Tabelle 4-110: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Kopfschmerzen

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
EASI zu <i>Baseline</i>, p = 0,0103			
16-25 Punkte	18/181 (9,9)	17/183 (9,3)	1,07 [0,57; 2,01] p=0,8322
> 25 Punkte	28/172 (16,3)	7/174 (4,0)	4,05 [1,82; 9,01] p=0,0006
Quelle: Table 14.3.1.6.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
EASI: <i>Eczema Area and Severity Index</i> ; KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus			

Es zeigte sich für den Endpunkt UE nach PT Kopfschmerzen ein Beleg für Effektmodifikation durch das Merkmal „EASI zu *Baseline*“. In der Subgruppe „> 25 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „16-25 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Tabelle 4-111: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Häufige UE nach PT Akne

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Krankheitsdauer, p=0,0458			
< 26 Jahre	22/222 (9,9)	7/220 (3,2)	3,11 [1,36; 7,14] p=0,0073
≥ 26 Jahre	24/140 (17,1)	3/145 (2,1)	8,29 [2,55; 26,90] p=0,0004
Quelle: Table 14.3.1.6.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus			

Es zeigte sich für den Endpunkt UE nach PT Akne ein Beleg für Effektmodifikation durch das Merkmal „Krankheitsdauer“. In den Subgruppen „< 26 Jahre“ und „≥ 26 Jahre“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen ist zuungunsten von Abrocitinib statistisch signifikant.

Tabelle 4-112: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Spezifische UE Follikulitis

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Peak Pruritus NRS zu Baseline, p=0,0234			
4-6 Punkte	7/83 (8,4)	0/105 (0,0)	17,80 [1,02; 309,07] p=0,0481
≥ 7 Punkte	4/274 (1,5)	3/259 (1,2)	1,26 [0,28; 5,58] p=0,7604
Quelle: Table 14.3.1.6.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; NRS: <i>Numerical Rating Scale</i> ; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus			

Es zeigte sich für den Endpunkt spezifische UE Follikulitis ein Beleg für Effektmodifikation durch das Merkmal „Peak Pruritus NRS zu Baseline“. In der Subgruppe „4-6 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zuungunsten von Abrocitinib statistisch signifikant. In der Subgruppe „≥ 7 Punkte“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Tabelle 4-113: Detaillierte Ergebnisse der Subgruppenanalyse mit statistisch signifikantem Interaktionsterm ($p < 0,05$); Endpunkt: Spezifische UE Konjunktivitis

Subgruppe	Abrocitinib 200 mg	Dupilumab 300 mg	Abrocitinib 200 mg vs. Dupilumab 300 mg
	n/N ^a (%)	n/N ^a (%)	RR [95 %-KI] p-Wert ^b
Abstammung, p=0,0408			
Kaukasisch	7/269 (2,6)	33/248 (13,3)	0,20 [0,09; 0,43] p<0,0001
Schwarz oder afro- amerikanisch	0/25 (0,0)	0/26 (0,0)	NA
Asiatisch	3/62 (4,8)	5/83 (6,0)	0,80 [0,20; 3,23] p=0,7579
Andere	0/6	1/8 (12,5)	0,62 [0,02; 15,61] p=0,7685
Krankheitsschwere zu Baseline, p=0,0225			
IGA = 3	3/216 (1,4)	28/220 (12,7)	0,11 [0,03; 0,35] p=0,0002
IGA = 4	7/146 (4,8)	11/145 (7,6)	0,63 [0,25; 1,58] p=0,3280
Krankheitsdauer, p=0,0177			
< 26 Jahre	4/222 (1,8)	13/220 (5,9)	0,30 [0,10; 0,92] p=0,0352
≥ 26 Jahre	6/140 (4,3)	26/145 (17,9)	0,24 [0,10; 0,56] p=0,0011
Quelle: Table 14.3.1.6.4 (57)			
a: Anzahl der Patienten im SAS.			
b: RR inklusive 95 %-KI und p-Wert berechnet mittels der unstratifizierten Wald-Methode.			
IGA: <i>Investigator's Global Assessment</i> ; KI: Konfidenzintervall; mg: Milligramm; N: Anzahl der in der Analyse eingeschlossenen Patienten; n: Anzahl der Patienten mit Ereignis; NA: Nicht anwendbar;			
PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SAS: <i>Safety-Analysis-Set</i> ; vs.: Versus			

Es zeigte sich für den Endpunkt spezifische UE Konjunktivitis ein Beleg für Effektmodifikation durch das Merkmal „Abstammung“. In der Subgruppe „kaukasisch“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zugunsten von Abrocitinib statistisch signifikant. In den Subgruppen „asiatisch“ und „andere“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant. In der

Subgruppe „schwarz oder afroamerikanisch“ sind in keiner der beiden Behandlungsgruppen entsprechende Ereignisse aufgetreten.

Es zeigte sich für den Endpunkt spezifische UE Konjunktivitis ein Beleg für Effektmodifikation durch das Merkmal „Krankheitsschwere zu *Baseline*“. In der Subgruppe „IGA = 3“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zugunsten von Abrocitinib statistisch signifikant. In der Subgruppe „IGA = 4“ ist der Effekt gleichgerichtet zur Hauptanalyse und nicht statistisch signifikant.

Es zeigte sich für den Endpunkt spezifische UE Konjunktivitis ein Beleg für Effektmodifikation durch das Merkmal „Krankheitsdauer“. In den Subgruppen „< 26 Jahre“ und „≥ 26 Jahre“ ist der Effekt gleichgerichtet zur Hauptanalyse und der Unterschied zwischen den Behandlungsarmen zugunsten von Abrocitinib statistisch signifikant.

Insgesamt zeigen sich in den Subgruppenanalysen der Studie JADE DARE zu Woche 26 nur wenige signifikante Effektmodifikationen. Die beobachteten Effektmodifikationen traten in diversen Merkmalen auf, wodurch kein konstanter Effektmodifikator über mehrere Endpunkte hinweg identifiziert werden konnte. Insbesondere wiesen die Interaktionstest für die Endpunkte SCORAD-90 und POEM-Gesamtscore = 0-2 keine signifikanten Interaktionen auf, sodass die Ergebnisse der Hauptanalysen als besonders valide anzusehen sind.

4.3.1.4 Liste der eingeschlossenen Studien - RCT

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Tabelle 4-114: Liste der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Verfügbare Quellen		
	Studienregistereinträge (ja/nein [Zitat])	Studienberichte (ja/nein [Zitat])	Publikationen (ja/nein [Zitat])
JADE DARE	ja (53–55)	ja (56)	nein

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).** Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie

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

Nicht zutreffend.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

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

Nicht zutreffend.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-115: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in drei Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

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

Tabelle 4-116: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

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

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

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

Tabelle 4-117: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-118: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

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

Tabelle 4-119: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Nicht zutreffend.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse:* Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.
- *Ergebnisse zu den Effekten:* Stellen Sie die gepoolten Ergebnisse dar.
- *Konsistenzprüfung:* Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere inkonsistente Ergebnisse.

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

Nicht zutreffend.

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

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

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

Nicht zutreffend.

4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*

- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.2 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 4-120: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus nicht randomisierten vergleichenden Studien beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-121: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

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

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-122: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-123: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.**

Nicht zutreffend.

4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie***

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

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

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

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

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

Nicht zutreffend.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus weiteren Untersuchungen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-124: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

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

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 4-125: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

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

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Nicht zutreffend.

4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Für die Darstellung des medizinischen Nutzens und Zusatznutzens wurden die Ergebnisse der multizentrischen, randomisierten, doppelblinden, *double-dummy*, aktivkontrollierten Phase-III-Studie JADE DARE herangezogen, welche nach § 5 des 5. Kapitels der Verfo des G-BA der Evidenzstufe Ib (randomisierte klinische Studie) zuzuordnen ist. Die Studie JADE DARE war nicht Teil der initialen Einreichung bei der EMA und wurde erst nachträglich in das Studienprogramm für Abrocitinib aufgenommen, um den Anforderungen des G-BA an die Dauer von Studien zur Nutzenbewertung von Arzneimitteln zur Behandlung chronischer Erkrankungen zu entsprechen. Die Studie schloss erwachsene Patienten mit mittelschwerer bis schwerer AD ein, die gemäß Leitlinienkriterien für eine systemische Therapie infrage kamen und beinhaltet einen direkten Vergleich mit der vom G-BA festgelegten zVT Dupilumab (Beratungsanforderung 2021-B-183-z (3)). In beiden Behandlungsarmen der Studie konnten die Patienten eine Hintergrundtherapie mit einer nicht-medikamentösen AD-Therapie (z. B. Emollentien) sowie mit einer medikamentösen AD-Therapie (schwach bis mittelstark wirksame TCS/ TCI) verwenden. Mit einer für die Nutzenbewertung bei chronischen Erkrankungen angemessenen Behandlungsdauer von 26 Wochen erlaubt die Studie eine Bewertung patientenrelevanter Endpunkte im vorliegenden Anwendungsgebiet.

Das Verzerrungspotenzial der Studie JADE DARE wurde als niedrig eingestuft (siehe Abschnitt 4.3.1.2.2). Auch für die patientenrelevanten Endpunkte, die für die Darstellung des Zusatznutzens herangezogen wurden, wurde aufgrund der doppelten Verblindung, der annähernd gleichlangen Nachbeobachtungszeiten in den Behandlungsarmen, der adäquaten Umsetzung des ITT-Prinzips und einer ergebnisunabhängigen Berichterstattung das

Verzerrungspotenzial jeweils als niedrig eingestuft. Die Rücklaufquoten für Bewertungsbögen und Fragebögen zur Erfassung patientenberichteter Endpunkte der Studie JADE DARE lagen zu jedem Zeitpunkt über 82 % und unterschieden sich zwischen den Behandlungsarmen zu keinem Zeitpunkt um mehr als 5,7 %. Insgesamt konnte die Robustheit aller Ergebnisse durch Sensitivitätsanalysen mit verschiedenen Ersetzungsstrategien für fehlende Werte bestätigt werden (siehe Anhang 4-G).

Die Studienpopulation umfasste Patienten mit mittelschwerer bis schwerer AD, die für eine systemische Therapie geeignet waren. Zur Überprüfung der Eignung für eine systemische Therapie wurden die *Baseline*-Charakteristika der Patienten der Studie JADE DARE mit den Kriterien der Checkliste der Aktualisierung „Systemtherapie bei Neurodermitis“ der S2k-Leitlinie (6) abgeglichen. Die Patienten entsprachen der Zielpopulation von Abrocitinib gemäß zugelassenem Anwendungsgebiet und waren sowohl hinsichtlich der allgemeinen Patientencharakteristika als auch hinsichtlich der krankheitsspezifischen Kriterien auf den deutschen Versorgungskontext übertragbar. Die Hintergrundtherapie in beiden Behandlungsarmen der Studie JADE DARE entsprach dabei den Empfehlungen deutscher Leitlinien für die Behandlung von AD (siehe Abschnitt 4.3.1.2.1).

Alle in der vorliegenden Nutzenbewertung zur Quantifizierung des Zusatznutzens herangezogenen Endpunkte sind gemäß der Begründung in Abschnitt 4.2.5.2 patientenrelevant. Sie wurden mit validen Instrumenten erhoben und unter Verwendung adäquater Methoden analysiert.

Die Aussagesicherheit der für die Nutzenbewertung vorliegenden Studie JADE DARE ist sowohl auf Studien- als auch auf Endpunktebene durchweg als hoch einzuschätzen, sodass Aussagen zum medizinischen Zusatznutzen als **Hinweis** zu bewerten sind.

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.

Das Anwendungsgebiet des zu bewertenden Arzneimittels Abrocitinib umfasst die Behandlung der mittelschweren bis schweren AD bei erwachsenen Patienten, die für eine systemische Therapie infrage kommen.

Die AD ist eine chronische, entzündliche Hauterkrankung, die durch entzündete, juckende und schmerzende Hautläsionen, die entweder dauerhaft oder in Schüben auftreten, gekennzeichnet ist. Die Erkrankung geht oft mit einer Beeinträchtigung der Lebensqualität (Schlafstörungen, reduzierte Arbeitsproduktivität, negative Auswirkungen auf das psychische Wohlbefinden etc.) einher und wird von Patienten, insbesondere durch den Juckreiz und die Schmerzen, als sehr belastend empfunden (siehe Modul 3 A, Abschnitt 3.2.1).

Es besteht ein hoher Bedarf an langfristig gut wirksamen und verträglichen Therapien, welche dauerhaft bei Patienten mit mittelschwerer bis schwerer AD gegeben werden können und die eine hohe Therapietreue und einfache Handhabbarkeit aufweisen. Mit Abrocitinib steht nun eine weitere für die AD hochwirksame und oral einzunehmende Therapie zur Verfügung. Aufgrund seiner Wirkweise zeichnet sich Abrocitinib durch einen schnellen Wirkeintritt, eine rasche und anhaltende Linderung des Juckreizes als Leitsymptom der AD, eine Verringerung der Anzahl an Exazerbationen bzw. der betroffenen BSA und eine Verbesserung des Schweregrades der Erkrankung, der patientenberichteten Symptomatik, Ängsten und Depressionen, Schlafstörungen, des Hautschmerzes sowie der krankheitsspezifischen Lebensqualität bei insgesamt vorhersehbarem und beherrschbarem Verträglichkeitsprofil aus.

In der folgenden Tabelle 4-126 wird die Wahrscheinlichkeit und das Ausmaß des Zusatznutzens von Abrocitinib gegenüber der zVT Dupilumab bei erwachsenen Patienten mit mittelschwerer bis schwerer AD zu Woche 26 für die betrachteten Endpunkte der Studie JADE DARE aufgezeigt.

Tabelle 4-126: Wahrscheinlichkeit und Ausmaß des Zusatznutzens

Endpunkt	Abrocitinib 200 mg vs. Dupilumab 300 mg Ereignisanteile % vs. % Effektschätzer [95 %-KI]; p-Wert ^a	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
Morbidität		
EASI		Hinweis auf einen geringen Zusatznutzen
Verbesserung um ≥ 75 %	70,2 % vs. 71,5 % RR = 0,98 [0,89; 1,08]; p = 0,6980	
Verbesserung um ≥ 90 %	52,5 % vs. 47,1 % RR = 1,11 [0,96; 1,29]; p = 0,1472	
SCORAD		
Verbesserung um ≥ 75 %	42,0 % vs. 36,4 % RR = 1,15 [0,96; 1,38]; p = 0,1282	
Verbesserung um ≥ 90 %	22,1 % vs. 14,3 % RR = 1,55 [1,13; 2,13]; p = 0,0067	
POEM		
Gesamtscore = 0-2	29,6 % vs. 19,0 % RR = 1,56 [1,19; 2,03]; p = 0,0011	
Verbesserung um ≥ 5 Punkte	76,1 % vs. 79,6 % RR = 0,96 [0,88; 1,03]; p = 0,2519	
Peak Pruritus NRS		
Verbesserung um ≥ 4 Punkte	67,5 % vs. 62,9 % RR = 1,07 [0,96; 1,19]; p = 0,1978	
SCORAD VAS für Juckreiz		
Verbesserung um ≥ 2 Punkte	75,4 % vs. 80,5 % RR = 0,94 [0,87; 1,01]; p = 0,0926	
MOS Sleep Scale Score Index-1		
Verbesserung um ≥ 15 Punkte	36,2 % vs. 32,2 % RR = 1,12 [0,92; 1,37]; p = 0,2642	
MOS Sleep Scale Score Index-2		
Verbesserung um ≥ 15 Punkte	38,4 % vs. 38,5 % RR = 1,00 [0,83; 1,20]; p = 0,9724	
SCORAD VAS für Schlafstörung		
Verbesserung um ≥ 2 Punkte	76,9 % vs. 79,6 % RR = 0,97 [0,89; 1,05]; p = 0,3960	
Skin Pain NRS		
Verbesserung um ≥ 4 Punkte	64,9 % vs. 62,2 % RR = 1,04 [0,93; 1,17]; p = 0,4745	
EQ-5D VAS		
Verbesserung um ≥ 15 Punkte	3,6 % vs. 1,7 % RR = 2,20 [0,85; 5,71]; p = 0,1040	

Endpunkt	Abrocitinib 200 mg vs. Dupilumab 300 mg Ereignisanteile % vs. % Effektschätzer [95 %-KI]; p-Wert ^a	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
Lebensqualität		
DLQI		Zusatznutzen nicht belegt
Gesamtscore = 0 oder 1	38,3 % vs. 31,6 % RR = 1,21 [0,99; 1,48]; p = 0,0603	
Verbesserung um ≥ 5 Punkte	72,3 % vs. 78,3 % RR = 0,92 [0,85; 1,01]; p = 0,0707	
Verträglichkeit		
Gesamtraten		Zusatznutzen nicht belegt
UE jeglichen Schweregrades (ergänzend)	74,0 % vs. 65,5 % RR = 1,13 [1,03; 1,24]; p = 0,0124	
Schwere UE	3,0 % vs. 2,2 % RR = 1,39 [0,56; 3,41]; p = 0,4763	
SUE	1,7 % vs. 1,6 % RR = 1,01 [0,33; 3,10]; p = 0,9885	
UE, die zum Therapieabbruch führten	2,5 % vs. 2,5 % RR = 1,01 [0,40; 2,51]; p = 0,9859	
UE jeglichen Schweregrades nach SOC und PT		
SOC: Infektionen und parasitäre Erkrankungen	18,8 % vs. 20,8 % RR = 0,90 [0,67; 1,21]; p = 0,4911	
PT: COVID-19	4,1 % vs. 3,3 % RR = 1,26 [0,60; 2,65]; p = 0,5427	
PT: Konjunktivitis	2,2 % vs. 9,6 % RR = 0,23 [0,11; 0,49]; p = 0,0001	
PT: Follikulitis	3,3 % vs. 0,8 % RR = 4,03 [1,15; 14,17]; p = 0,0296	
PT: Herpes simplex	3,3 % vs. 1,4 % RR = 2,42 [0,86; 6,80]; p = 0,0936	
PT: Nasopharyngitis	3,9 % vs. 3,3 % RR = 1,18 [0,55; 2,51]; p = 0,6742	
PT: Oraler Herpes	2,5 % vs. 4,1 % RR = 0,60 [0,27; 1,36]; p = 0,2260	
PT: Infektion der oberen Atemwege	2,8 % vs. 2,5 % RR = 1,12 [0,46; 2,72]; p = 0,8022	
SOC: Erkrankungen des Gastrointestinaltrakts	19,3 % vs. 3,0 % RR = 6,42 [3,46; 11,91]; p < 0,0001	
PT: Übelkeit	19,3 % vs. 2,2 % RR = 8,82 [4,31; 18,07]; p < 0,0001	
PT: Erbrechen	3,0 % vs. 1,6 % RR = 1,85 [0,69; 4,95]; p = 0,2211	

Endpunkt	Abrocitinib 200 mg vs. Dupilumab 300 mg Ereignisanteile % vs. % Effektschätzer [95 %-KI]; p-Wert ^a	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort PT: Erschöpfung	2,8 % vs. 1,4 % RR = 2,02 [0,70; 5,84]; p = 0,1962 2,8 % vs. 1,4 % RR = 2,02 [0,70; 5,84]; p = 0,1962	
SOC: Untersuchungen PT: Kreatinphosphokinase im Blut erhöht PT: Abnahme der natürlichen Killerzellen PT: SARS-CoV-2 positiv	10,5 % vs. 7,1 % RR = 1,47 [0,91; 2,37]; p = 0,1112 3,9 % vs. 3,6 % RR = 1,09 [0,52; 2,28]; p = 0,8275 2,8 % vs. 0,0 % RR = 20,19 [1,18; 344,43]; p = 0,0378 4,1 % vs. 3,6 % RR = 1,16 [0,56; 2,41]; p = 0,6838	
SOC: Erkrankungen des Nervensystems PT: Schwindelgefühl PT: Kopfschmerzen	15,2 % vs. 7,4 % RR = 2,05 [1,33; 3,18]; p = 0,0012 2,8 % vs. 1,1 % RR = 2,52 [0,80; 7,96]; p = 0,1152 13,0 % vs. 6,6 % RR = 1,97 [1,23; 3,16]; p = 0,0045	
SOC: Erkrankungen der Haut und des Unterhautgewebes PT: Akne PT: Dermatitis atopisch	16,9 % vs. 6,3 % RR = 2,67 [1,69; 4,22]; p < 0,0001 12,7 % vs. 2,7 % RR = 4,64 [2,38; 9,05]; p < 0,0001 4,7 % vs. 3,8 % RR = 1,22 [0,61; 2,45]; p = 0,5666	
Spezifische UE jeglichen Schweregrades		
Superinfektionen	8,6 % vs. 6,0 % RR = 1,42 [0,84; 2,41]; p = 0,1912	
Herpes zoster	2,5 % vs. 0,6 % RR = 4,54 [0,99; 20,85]; p = 0,0520	
Konjunktivitis	2,8 % vs. 10,7 % RR = 0,26 [0,13; 0,51]; p < 0,0001	
Akne	13,3 % vs. 3,0 % RR = 4,40 [2,32; 8,33]; p < 0,0001	
Follikulitis	3,0 % vs. 0,8 % RR = 3,70 [1,04; 13,14]; p = 0,0433	
Analysen mit signifikantem Gruppenunterschied wurden fett markiert. a: In den Analysen zur Morbidität und Lebensqualität wurde der Effektschätzer inklusive 95 %-KI und p-Wert berechnet mittels Cochran-Mantel-Haenszel-Test stratifiziert nach Krankheitsschwere zu <i>Baseline</i> (IGA = 3 vs. IGA = 4). In den Analysen zur Verträglichkeit wurde der Effektschätzer inklusive 95 %-KI und p-Wert berechnet anhand der unstratifizierten Approximation der Normalverteilung nach der Wald-Methode.		

Endpunkt	Abrocitinib 200 mg vs. Dupilumab 300 mg Ereignisanteile % vs. % Effektschätzer [95 %-KI]; p-Wert ^a	Wahrscheinlichkeit und Ausmaß des Zusatznutzens
CoV: Coronavirus; DLQI: <i>Dermatology Life Quality Index</i> ; EASI: <i>Eczema Area and Severity Index</i> ; EQ-5D: <i>European Quality of Life Group 5 Dimensions</i> ; IGA: <i>Investigator's Global Assessment</i> ; KI: Konfidenzintervall; mg: Milligramm; MOS: <i>Medical Outcomes Study</i> ; NRS: <i>Numerical Rating Scale</i> ; POEM: <i>Patient-Oriented Eczema Measures</i> ; PT: Bevorzugte Bezeichnung (<i>Preferred Term</i>); RR: Relatives Risiko; SARS: <i>Severe Acute Respiratory Syndrome</i> ; SCORAD: <i>Scoring Atopic Dermatitis</i> ; SOC: Systemorganklasse (<i>System Organ Class</i>); UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala; vs.: Versus		

Morbidität

Die Kategorie Morbidität umfasst AD-relevante Endpunkte, wie EASI, SCORAD, patientenberichtete Symptomatik, Juckreiz, Schlafstörungen, Hautschmerz, Ängste und Depressionen sowie den Gesundheitszustand.

Bei EASI und SCORAD handelt es sich um auch in der Praxis anerkannte und etablierte Instrumente, die den Schweregrad der AD und die Veränderungen der Haut, wie z. B. Erythem, Ödem, Abschürfung und Lichenifikation erfassen. Die für AD wesentlichen Symptome und deren Häufigkeit sowie der Gesundheitszustand wurden vom Patienten selbst berichtet und sind demnach als patientenrelevant anzusehen. Juckreiz, Schlafstörungen und Hautschmerz gelten aufgrund der stark belastenden Auswirkungen auf die Lebensqualität der Patienten als weitere bedeutsame Endpunkte in der AD. Zudem weisen Patienten mit AD eine erhöhte Vulnerabilität für psychische Erkrankungen auf, wonach Ängste und Depressionen als AD-relevante Endpunkte betrachtet werden (siehe Abschnitt 4.2.5.2). Eine klinisch relevante Verbesserung der Krankheits schwere der AD und deren Symptome wie Juckreiz und Hautschmerz wurde meist schon nach 2 Wochen erreicht.

Eine deutliche Verbesserung des therapierelevanten Nutzens gegenüber der zVT zeigte sich in der Studie JADE DARE zu Woche 26 in einem statistisch signifikant höheren Anteil an Patienten mit einer 90%igen Verbesserung des Schweregrades der AD gemessen anhand des SCORAD (SCORAD-90) und anhand eines statistisch signifikant höheren Anteils an Patienten, die zu Woche 26 nahezu beschwerdefrei waren, gemessen mit dem patientenberichteten Symptomfragebogen POEM (POEM-Gesamtscore von 0-2). Dabei zeigten Patienten im Abrocitinib-Arm, die eine 90%ige Reduzierung des SCORAD erreichten, statistisch signifikant mehr Tage, an denen sie dieses Ansprechen ohne eine medikamentöse Hintergrundtherapie aufrechterhalten konnten als Patienten im Dupilumab-Arm (*Medicated Topical Background Therapy-Free Days*: 15,2 Tage unter Abrocitinib, 6,2 Tage unter Dupilumab). Zudem zeigten Patienten unter Abrocitinib statistisch signifikant mehr Tage ein SCORAD-90-Ansprechen als Patienten unter Dupilumab (Abrocitinib: 17,6 Tage versus Dupilumab: 7,1 Tage) (siehe Anhang 4-G). Ein ähnliches Bild zeigte sich in den *Medicated Topical Background Therapy-Free Days* und der Tage mit Ansprechen für Patienten, die einen POEM-Gesamtscore von 0-2 und ein EASI-90-Ansprechen erreicht haben (siehe Anhang 4-G). Die dargestellten Sensitivitätsanalysen zum SCORAD-90 und POEM-Gesamtscore von 0-2 unterstreichen die Robustheit der Ergebnisse der Hauptanalyse und zeigten ebenfalls statistisch signifikante

Unterschiede zugunsten von Abrocitinib gegenüber Dupilumab. Zudem zeigten sich in den Sensitivitätsanalysen zum EASI-90 und SCORAD-75 statistisch signifikante Unterschiede zugunsten von Abrocitinib (siehe Anhang 4-G). Die Reduktion des Schweregrades der AD unter Abrocitinib wurde weiterhin durch den IGA als primären Studienendpunkt in der Studie JADE DARE bestätigt (siehe Anhang 4-G). Weiterhin wiesen zu Woche 26 die Patienten in der Abrocitinib-Gruppe statistisch signifikant häufiger eine 90%ige Verbesserung im oSCORAD auf als in der Dupilumab-Gruppe (siehe Anhang 4-G). Auch wurde eine vergleichbare Verbesserung des Gesundheitszustandes, sowie eine vergleichbare Verminderung des Juckreizes, der Schlafstörungen, des Hautschmerzes und der Reduzierung von Ängsten und Depressionen zwischen den Behandlungsarmen festgestellt.

Schneller Wirkeintritt

Auch wenn die langfristige Krankheitskontrolle im Mittelpunkt der AD-Therapie steht, sind auch kurzfristige Effekte geeignet einen patientenrelevanten Zusatznutzen bei Patienten mit moderater bis schwerer AD zu zeigen. Erste Behandlungseffekte, wie z. B. die Reduktion des Juckreizes, und die Verbesserung der Hautscores (EASI und SCORAD) traten mit Abrocitinib früher auf als mit Dupilumab. Dies wird durch statistisch signifikante und klinisch relevante Gruppenunterschiede bei der Veränderung der Hautscores zugunsten von Abrocitinib zwischen Woche 2 und Woche 16 belegt, die sich auch in einer Verbesserung der Lebensqualität widerspiegeln (siehe Abschnitt 4.3.1.3.1, Abbildung 4-3, Abbildung 4-4, Abbildung 4-6, Abbildung 4-7, Abbildung 4-8 und Abbildung 4-16).

Bereits 2 Wochen nach Behandlungsbeginn war der Anteil der Patienten mit einer deutlichen Verbesserung der Hautsymptomatik (EASI-90 und SCORAD-90) im Abrocitinib-Arm größer als im Dupilumab-Arm (siehe Abbildung 4-1 und Abbildung 4-2). Dieser Vorteil blieb während der gesamten Behandlungsdauer bestehen und war größtenteils statistisch signifikant (siehe Anhang 4-G). Ebenso war der Anteil der symptomfreien oder nahezu symptomfreien Patienten (POEM-Gesamtscore = 0-2) ab dem ersten Nachbeobachtungszeitpunkt (Woche 12) und zu allen weiteren Beobachtungszeitpunkten in der Abrocitinib-Gruppe statistisch signifikant größer (siehe Abbildung 4-3 und Anhang 4-G). Der im Vergleich zu Dupilumab frühere Wirkeintritt unter Abrocitinib ging dabei mit einer höheren Anzahl an Tagen mit SCORAD-90-Ansprechen sowie einem POEM-Gesamtscore = 0-2 und dementsprechend mehr symptomfreien oder nahezu symptomfreien Tagen unter Abrocitinib einher (siehe Anhang 4-G).

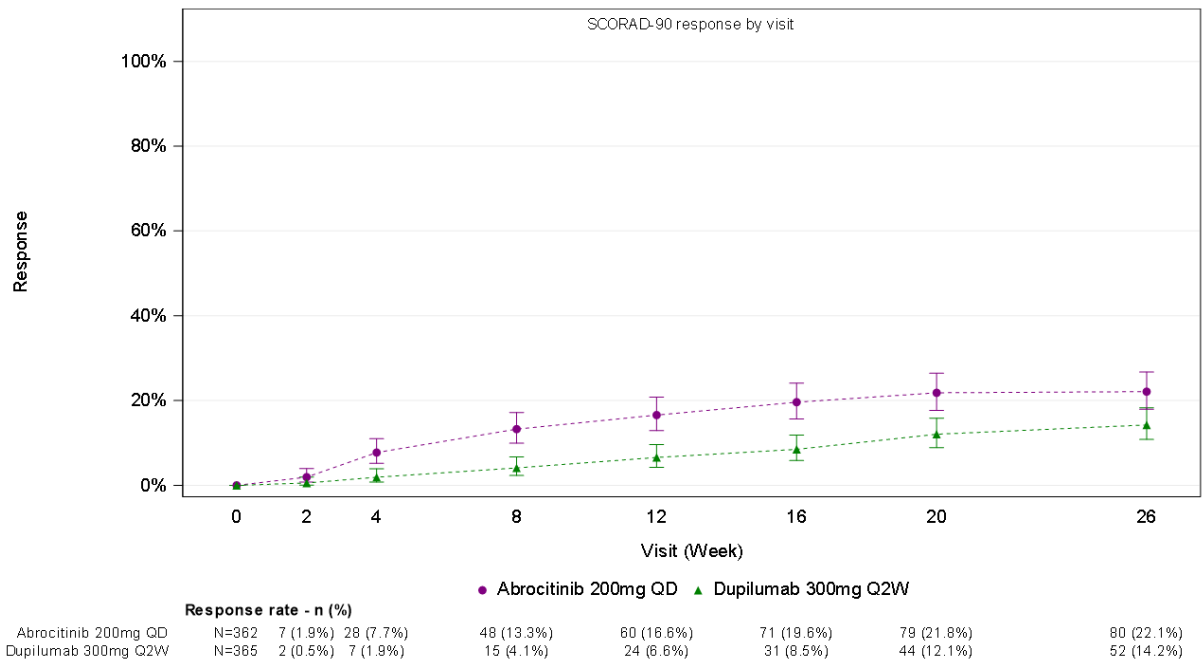


Abbildung 4-17: Verlaufskurve für den Anteil der Patienten mit SCORAD-90 Ansprechen von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

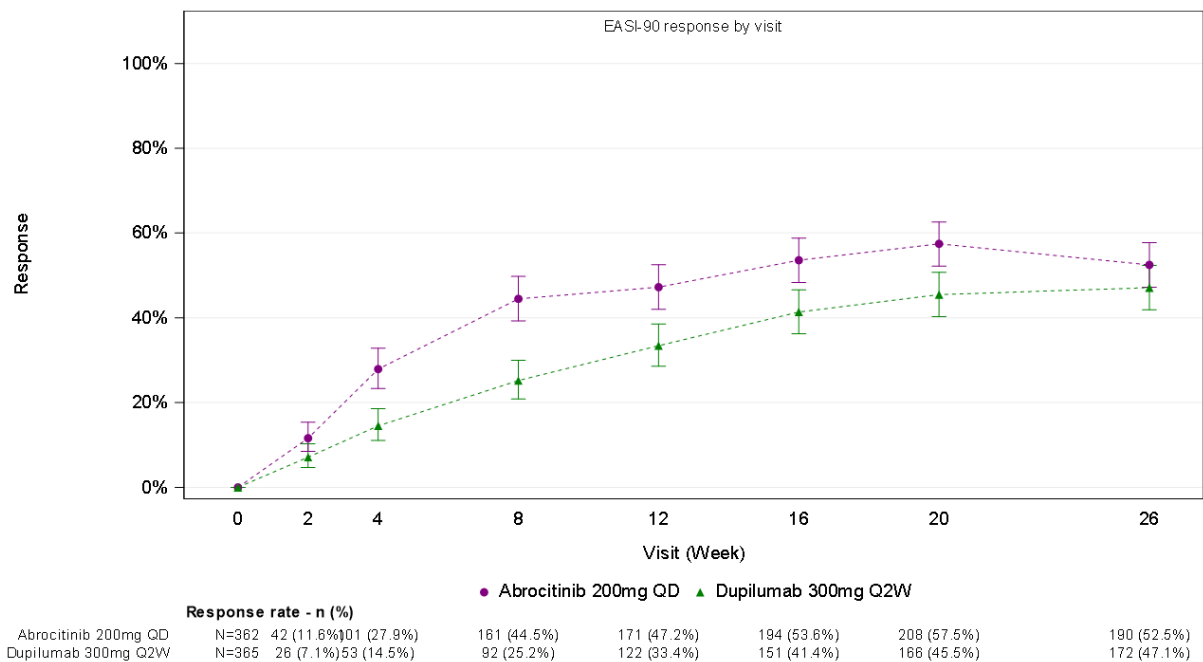


Abbildung 4-18: Verlaufskurve für den Anteil der Patienten mit EASI-90 Ansprechen von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

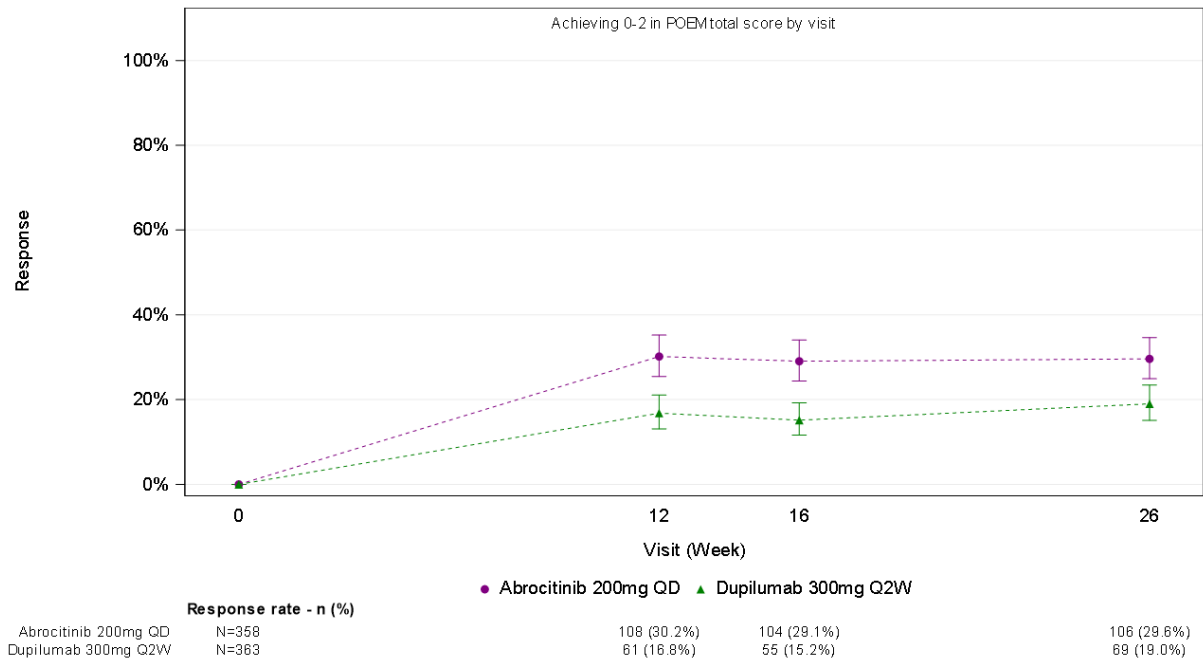


Abbildung 4-19: Verlaufskurve für den Anteil der Patienten mit einem POEM-Gesamtscore von 0-2 von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Die unter Abrocitinib beobachteten Effekte zeigten sich als dauerhaft, wie die Langzeit-*Extension*-Studie JADE EXTEND zeigte, die die Monotherapie mit Abrocitinib über einen Behandlungszeitraum von 48 Wochen untersuchte (58).

Insgesamt ergibt sich für die Behandlung mit Abrocitinib gegenüber Dupilumab aufgrund des statistisch signifikant größeren Anteils an Patienten mit einer deutlichen Reduktion der Krankheitsschwere und des signifikant höheren Anteils an Patienten, die eine Symptombefreiheit oder annähernde Symptombefreiheit erreichten, in der Kategorie Morbidität ein **Hinweis auf einen geringen Zusatznutzen**.

Gesundheitsbezogene Lebensqualität

Abrocitinib zeigte im DLQI zu Woche 26 eine vergleichbare Verbesserung in der krankheitsspezifischen gesundheitsbezogenen Lebensqualität gegenüber der hoch wirksamen zVT Dupilumab. Im Studienverlauf zeigte sich bereits ab Woche 2 ein höherer Anteil der Patienten mit keiner bzw. nahezu keiner Beeinträchtigung der Lebensqualität (DLQI-Gesamtscore = 0 oder 1) in der Abrocitinib-Gruppe gegenüber der Dupilumab-Gruppe (siehe Abbildung 4-4). Die Gruppenunterschiede zu den Beobachtungszeitpunkten waren dabei größtenteils statistisch signifikant (siehe Anhang 4-G). Bei vorrangiger Berücksichtigung der Ergebnisse zu Woche 26 ergibt sich in der Kategorie gesundheitsbezogene Lebensqualität **kein Beleg für einen Zusatznutzen** von Abrocitinib bei erwachsenen Patienten, die für eine systemische Therapie infrage kommen, gegenüber Dupilumab.

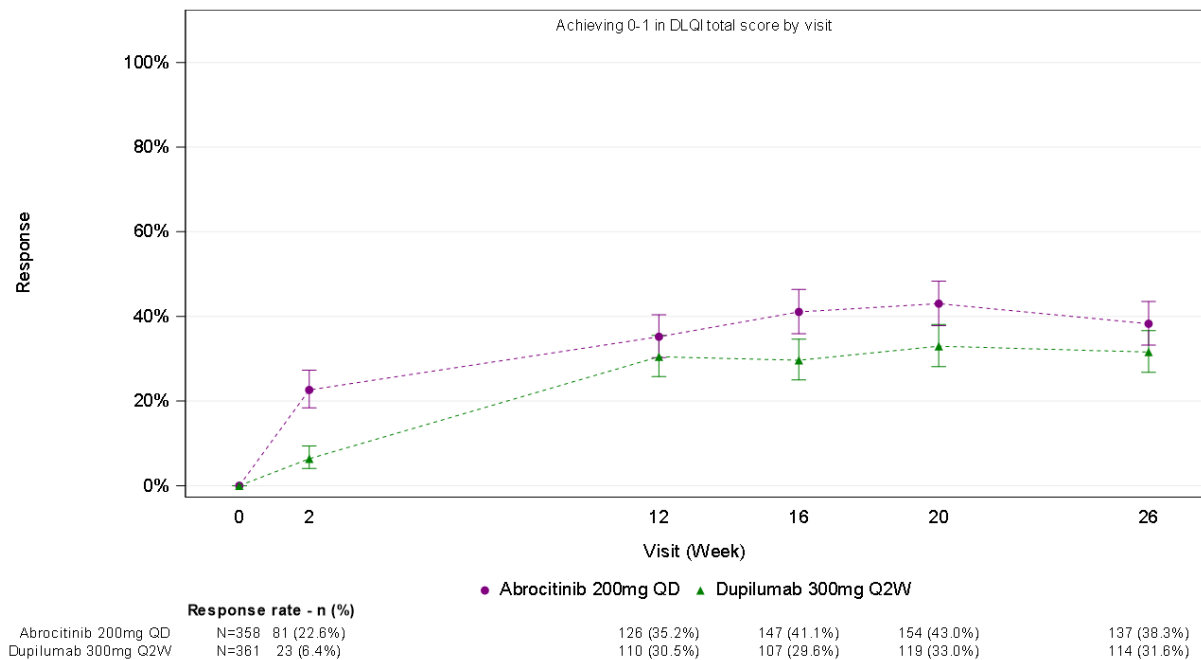


Abbildung 4-20: Verlaufskurve für den Anteil der Patienten mit einem DLQI von 0 oder 1 von *Baseline* bis Woche 26 aus RCT (JADE DARE) mit dem zu bewertenden Arzneimittel

Verträglichkeit

Endpunkte zur Verträglichkeit, die UE erfassen, sind aufgrund einer möglichen zusätzlichen physischen und psychischen Belastung als patientenrelevant anzusehen. In der Studie JADE DARE traten schwere UE oder SUE in beiden Behandlungsarmen selten auf; die entsprechenden Raten waren zwischen den Behandlungsgruppen vergleichbar. Ebenso niedrig und zwischen den Behandlungsgruppen vergleichbar war der Anteil der Patienten, die die Therapie aufgrund von UE abbrachen. Die Gesamtrate der UE jeglichen Schweregrades wurde zur Bewertung des Zusatznutzens nicht herangezogen, sondern lediglich ergänzend dargestellt. Bei den beobachteten Unterschieden wird von einem nicht mehr als geringfügigen Effekt ausgegangen.

Bei den UE nach PT zeigte sich ein statistisch signifikanter Unterschied zuungunsten von Abrocitinib bei den Endpunkten Übelkeit, Kopfschmerzen, Akne, Follikulitis und Abnahme der natürlichen Killerzellen sowie bei den entsprechenden SOC Erkrankungen des Gastrointestinaltrakts, Erkrankungen des Nervensystems und Erkrankungen der Haut und des Unterhautgewebes. Bei den UE handelte es sich um nicht schwere, nicht schwerwiegende Ereignisse, die mehrheitlich nicht zum Abbruch der Therapie führten, gut beherrschbar waren und somit keine Auswirkungen auf die Lebensqualität der Patienten hatten. Für den Endpunkt Konjunktivitis zeigte sich ein statistisch signifikanter Unterschied zugunsten von Abrocitinib.

In der Kategorie Verträglichkeit lässt sich somit aufgrund vergleichbarer, in beiden Behandlungsarmen niedriger Gesamtraten schwerer UE, SUE und der geringen und

vergleichbaren Zahl an Therapieabbrüchen wegen UE, **kein Beleg für einen Zusatznutzen** von Abrocitinib gegenüber Dupilumab ableiten.

Gesamtschau des medizinischen Zusatznutzens

Mit Abrocitinib, als einem oral einzunehmenden Wirkstoff, steht eine wirksame und gut verträgliche Therapiemöglichkeit zur Behandlung von mittelschwerer bis schwerer AD bei erwachsenen Patienten, die für eine systemische Therapie infrage kommen, zur Verfügung.

Gemäß AM-NutzenV konnte in den Wirksamkeitsergebnissen durch die Behandlung mit Abrocitinib eine bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens festgestellt werden. Dieser besteht vor allem in einer Reduzierung des Schweregrades der AD und der damit einhergehenden spürbaren Linderung der Erkrankung im Sinne der patientenberichteten Symptomatik bei einer allgemein guten Verträglichkeit. Es wird daher ein **Hinweis auf einen geringen Zusatznutzen** abgeleitet.

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 4-127: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene mit mittelschwerer bis schwerer atopischer Dermatitis, die für eine systemische Therapie infrage kommen.	Hinweis auf einen geringen Zusatznutzen

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht zutreffend.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Nicht zutreffend.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden

patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹⁷, Molenberghs 2010¹⁸). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrundeliegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁹) 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 2006²⁰) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Nicht zutreffend.

¹⁷ Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.

¹⁸ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.

¹⁹ Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.

²⁰ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

4.6 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. Pfizer Europe MA EEIG. Fachinformation Cibinqo® 50 mg/100 mg/200 mg Filmtabletten. Stand: Dezember 2021. 2021. Verfügbar unter: <https://www.fachinfo.de/>. [Zugriff am: 03.01.2022].
2. Sanofi Genzyme. Fachinformation Dupixent® 300 mg Injektionslösung in einer Fertigspritze / Dupixent® 300 mg Injektionslösung im Fertigpen. Stand: November 2021. 2021. Verfügbar unter: <https://www.fachinfo.de/>. [Zugriff am: 03.01.2022].
3. Gemeinsamer Bundesausschuss (G-BA). Information über eine Änderung der zweckmäßigen Vergleichstherapie. Beratungsanforderung 2021-B-183-z (2020-B-164; 2020-B086-z; 2019-B-159). Abrocitinib zur Behandlung der atopischen Dermatitis. 2021.
4. Gemeinsamer Bundesausschuss (G-BA). Verfahrensordnung des Gemeinsamen Bundesausschusses: in der Fassung vom 18. Dezember 2008 veröffentlicht im Bundesanzeiger Nr. 84a (Beilage) vom 10. Juni 2009 in Kraft getreten am 1. April 2009 zuletzt geändert durch den Beschluss vom 4. November 2021 veröffentlicht im Bundesanzeiger BAnz AT 08.12.2021 B3 in Kraft getreten am 9. Dezember 2021. 2021. Verfügbar unter: https://www.g-ba.de/downloads/62-492-2667/VerfO_2021-11-04_iK-2021-12-09.pdf. [Zugriff am: 03.01.2022].
5. Gemeinsamer Bundesausschuss (G-BA). Niederschrift (vorläufige Fassung) zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV. Beratungsanforderung 2019-B-159. Abrocitinib angezeigt für die Behandlung von Patienten ≥ 12 Jahren mit mittelschwerer bis schwerer atopischer Dermatitis. 2019.
6. Werfel T, Heratizadeh A, Aberer W, Ahrens F, Augustin M, Biedermann T et al. Aktualisierung „Systemtherapie bei Neurodermitis“ zur S2k-Leitlinie Neurodermitis. J Dtsch Dermatol Ges 2021; 19(1):151–69. doi: 10.1111/ddg.14371_g.
7. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden: Version 6.0 vom 05.11.2020. 2020. Verfügbar unter: https://www.iqwig.de/methoden/allgemeine-methoden_version-6-0.pdf. [Zugriff am: 07.05.2021].
8. Wong SS-L, 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.
9. Hanifin JM, Thurston M, Omoto M, Cherill R, Tofte SJ, Graeber M. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. EASI Evaluator Group. Exp Dermatol 2001; 10(1):11–8. doi: 10.1034/j.1600-0625.2001.100102.x.

10. Gemeinsamer Bundesausschuss (G-BA). Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Dupilumab. 2018. Verfügbar unter: https://www.g-ba.de/downloads/40-268-4986/2018-05-17_AM-RL-XII_Dupilumab_D-328_TrG.pdf. [Zugriff am: 22.09.2021].
11. Gemeinsamer Bundesausschuss (G-BA). Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Dupilumab (neues Anwendungsgebiet: atopische Dermatitis, Jugendliche ab 12 bis < 18 Jahre): Vom 20. Februar 2020. 2020. Verfügbar unter: https://www.g-ba.de/downloads/40-268-6363/2020-02-20_AM-RL-XII_Dupilumab_nAWG_D-483_TrG.pdf. [Zugriff am: 22.09.2021].
12. Gemeinsamer Bundesausschuss (G-BA). Mündliche Anhörung gemäß 5. Kapitel § 19 Abs. 2 Verfahrensordnung des Gemeinsamen Bundesausschusses hier: Wirkstoff Dupilumab: Sitzung im Hause des Gemeinsamen Bundesausschusses in Berlin am 9. April 2018 von 14.08 Uhr bis 15.15 Uhr. 2018. Verfügbar unter: https://www.g-ba.de/downloads/91-1031-333/2017_12_01_Wortprotokoll_Dupilumab_D-328.pdf. [Zugriff am: 10.05.2021].
13. Schmitt J, Langan S, Deckert S, Svensson A, Kobyletzki L von, Thomas K et al. Assessment of clinical signs of atopic dermatitis: a systematic review and recommendation. *J Allergy Clin Immunol* 2013; 132(6):1337–47. doi: 10.1016/j.jaci.2013.07.008.
14. Rehal B, Armstrong AW, Armstrong A. Health outcome measures in atopic dermatitis: a systematic review of trends in disease severity and quality-of-life instruments 1985-2010. *PLoS One* 2011; 6(4):e17520. doi: 10.1371/journal.pone.0017520.
15. ETFAD. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. *Dermatology* 1993; 186(1):23–31. doi: 10.1159/000247298.
16. Charman C, Williams H. Outcome measures of disease severity in atopic eczema. *Arch Dermatol* 2000; 136(6):763–9. doi: 10.1001/archderm.136.6.763.
17. Gerbens LAA, Prinsen CAC, Chalmers JR, Drucker AM, Kobyletzki LB von, Limpens J et al. Evaluation of the measurement properties of symptom measurement instruments for atopic eczema: a systematic review. *Allergy* 2017; 72(1):146–63. doi: 10.1111/all.12959.
18. Silverberg JI, Margolis DJ, Boguniewicz M, Fonacier L, Grayson MH, Ong PY et al. Validation of five patient-reported outcomes for atopic dermatitis severity in adults. *Br J Dermatol* 2020; 182(1):104–11. doi: 10.1111/bjd.18002.
19. Hawro T, Przybyłowicz K, Spindler M, Hawro M, Steć M, Altrichter S et al. The characteristics and impact of pruritus in adult dermatology patients: A prospective, cross-

- sectional study. *J Am Acad Dermatol* 2021; 84(3):691–700. doi: 10.1016/j.jaad.2020.08.035.
20. Newton L, DeLozier AM, Griffiths PC, Hill JN, Hudgens S, Symonds T et al. Exploring content and psychometric validity of newly developed assessment tools for itch and skin pain in atopic dermatitis. *J Patient Rep Outcomes* 2019; 3(1):42. doi: 10.1186/s41687-019-0128-z.
 21. Yosipovitch G, Reaney M, Mastey V, Eckert L, Abbé A, Nelson L et al. Peak Pruritus Numerical Rating Scale: psychometric validation and responder definition for assessing itch in moderate-to-severe atopic dermatitis. *Br J Dermatol* 2019; 181(4):761–9. doi: 10.1111/bjd.17744.
 22. Spindler M, Przybyłowicz K, Hawro M, Weller K, Reidel U, Metz M et al. Sleep disturbance in adult dermatologic patients: A cross-sectional study on prevalence, burden, and associated factors. *J Am Acad Dermatol* 2021; 85(4):910–22. doi: 10.1016/j.jaad.2021.04.015.
 23. Kage P, Simon J-C, Treudler R. Atopische Dermatitis und psychosoziale Komorbidität. *JDDG: Journal der Deutschen Dermatologischen Gesellschaft* 2020; 18(2):93–102. doi: 10.1111/ddg.14029_g.
 24. Bawany F, Northcott CA, Beck LA, Pigeon WR. Sleep Disturbances and Atopic Dermatitis: Relationships, Methods for Assessment, and Therapies. *J Allergy Clin Immunol Pract* 2021; 9(4):1488–500. doi: 10.1016/j.jaip.2020.12.007.
 25. Fishbein AB, Vitaterna O, Haugh IM, Bavishi AA, Zee PC, Turek FW et al. Nocturnal eczema: Review of sleep and circadian rhythms in children with atopic dermatitis and future research directions. *J Allergy Clin Immunol* 2015; 136(5):1170–7. doi: 10.1016/j.jaci.2015.08.028.
 26. Ramirez FD, Chen S, Langan SM, Prather AA, McCulloch CE, Kidd SA et al. Association of Atopic Dermatitis With Sleep Quality in Children. *JAMA Pediatr* 2019; 173(5):e190025. doi: 10.1001/jamapediatrics.2019.0025.
 27. Kahn D, Iturriaga C, Bertran K, Fernandez I, Perez-Mateluna G, Borzutzky A et al. Sleep quality in children with atopic dermatitis during flares and after treatment. *Sleep Sci* 2020; 13(2):172–5. doi: 10.5935/1984-0063.20190139.
 28. Riemann D, Baum E, Cohrs S, Crönlein T, Hajak G, Hertenstein E et al. S3-Leitlinie Nicht erholsamer Schlaf/Schlafstörungen. *Somnologie* 2017; 21(1):2–44. doi: 10.1007/s11818-016-0097-x.
 29. Allen RP, Kosinski M, Hill-Zabala CE, Calloway MO. Psychometric evaluation and tests of validity of the Medical Outcomes Study 12-item Sleep Scale (MOS sleep). *Sleep Med* 2009; 10(5):531–9. doi: 10.1016/j.sleep.2008.06.003.
 30. Rejas J, Ribera MV, Ruiz M, Masramón X. Psychometric properties of the MOS (Medical Outcomes Study) Sleep Scale in patients with neuropathic pain. *Eur J Pain* 2007; 11(3):329–40. doi: 10.1016/j.ejpain.2006.05.002.

31. Hays RD, Martin SA, Sesti AM, Spritzer KL. Psychometric properties of the Medical Outcomes Study Sleep measure. *Sleep Med* 2005; 6(1):41–4. doi: 10.1016/j.sleep.2004.07.006.
32. Vakharia PP, Chopra R, Sacotte R, Patel KR, Singam V, Patel N et al. Burden of skin pain in atopic dermatitis. *Ann Allergy Asthma Immunol* 2017; 119(6):548-552.e3. doi: 10.1016/j.anai.2017.09.076.
33. Silverberg JI, Gelfand JM, Margolis DJ, Boguniewicz M, Fonacier L, Grayson MH et al. Pain Is a Common and Burdensome Symptom of Atopic Dermatitis in United States Adults. *J Allergy Clin Immunol Pract* 2019; 7(8):2699-2706.e7. doi: 10.1016/j.jaip.2019.05.055.
34. Silverberg JI. Validity and reliability of a novel numeric rating scale to measure skin-pain in adults with atopic dermatitis. *Arch Dermatol Res* 2021; 313(10):855–61. doi: 10.1007/s00403-021-02185-3.
35. Barrett A, Hahn-Pedersen J, Kragh N, Evans E, Gnanasakthy A. Patient-Reported Outcome Measures in Atopic Dermatitis and Chronic Hand Eczema in Adults. *Patient* 2019; 12(5):445–59. doi: 10.1007/s40271-019-00373-y.
36. Patel KR, Immaneni S, Singam V, Rastogi S, Silverberg JI. Association between atopic dermatitis, depression, and suicidal ideation: A systematic review and meta-analysis. *J Am Acad Dermatol* 2019; 80(2):402–10. doi: 10.1016/j.jaad.2018.08.063.
37. Silverberg JI, Gelfand JM, Margolis DJ, Boguniewicz M, Fonacier L, Grayson MH et al. Symptoms and diagnosis of anxiety and depression in atopic dermatitis in U.S. adults. *Br J Dermatol* 2019; 181(3):554–65. doi: 10.1111/bjd.17683.
38. Silverberg JI, Gelfand JM, Margolis DJ, Boguniewicz M, Fonacier L, Grayson MH et al. Measurement Properties of the Hospital Anxiety and Depression Scale Used in Atopic Dermatitis in Adults. *Journal of Investigative Dermatology* 2019; 139(6):1388–91. doi: 10.1016/j.jid.2018.11.029.
39. Snaith RP. The Hospital Anxiety And Depression Scale. *Health Qual Life Outcomes* 2003; 1(1):29. doi: 10.1186/1477-7525-1-29.
40. Snaith RP, Zigmond AS. The hospital anxiety and depression scale. *Br Med J (Clin Res Ed)* 1986; 292(6516):344. doi: 10.1136/bmj.292.6516.344.
41. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002; 52(2):69–77. doi: 10.1016/s0022-3999(01)00296-3.
42. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983; 67(6):361–70. doi: 10.1111/j.1600-0447.1983.tb09716.x.
43. Yang Y, Brazier J, Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. *Eur J Health Econ* 2015; 16(9):927–39. doi: 10.1007/s10198-014-0638-9.
44. McCombs K, Chen SC. Patient preference quality of life measures in dermatology. *Dermatol Ther* 2007; 20(2):102–9. doi: 10.1111/j.1529-8019.2007.00118.x.

45. Silverberg JI, Lei D, Yousaf M, Janmohamed SR, Vakharia PP, Chopra R et al. Measurement properties of the product of investigator's global assessment and body surface area in children and adults with atopic dermatitis. *J Eur Acad Dermatol Venereol* 2021; 35(1):180–7. doi: 10.1111/jdv.16846.
46. Patel KR, Singam V, Vakharia PP, Chopra R, Sacotte R, Patel N et al. Measurement properties of three assessments of burden used in atopic dermatitis in adults. *Br J Dermatol* 2019; 180(5):1083–9. doi: 10.1111/bjd.17243.
47. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994; 19(3):210–6. doi: 10.1111/j.1365-2230.1994.tb01167.x.
48. Lewis-Jones MS, Finlay AY. The Children's Dermatology Life Quality Index (CDLQI): initial validation and practical use. *Br J Dermatol* 1995; 132(6):942–9. doi: 10.1111/j.1365-2133.1995.tb16953.x.
49. Heintz D, Chalmers J, Nankervis H, Apfelbacher CJ. Eczema Trials: Quality of Life Instruments Used and Their Relation to Patient-reported Outcomes. A Systematic Review. *Acta Derm Venereol* 2016; 96(5):596–601. doi: 10.2340/00015555-2322.
50. Basra MKA, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. *Br J Dermatol* 2008; 159(5):997–1035. doi: 10.1111/j.1365-2133.2008.08832.x.
51. European Medicines Agency (EMA). Guideline for good clinical practice E6(R2): Step 5. 2016. Verfügbar unter: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-6-r2-guideline-good-clinical-practice-step-5_en.pdf. [Zugriff am: 10.05.2021].
52. Pfizer Inc. Final Study Protocol B7451050: A PHASE 3B RANDOMIZED, DOUBLE-BLIND, DOUBLE-DUMMY, ACTIVE CONTROLLED MULTI-CENTER STUDY ASSESSING THE EFFICACY AND SAFETY OF ABROCITINIB COMPARED WITH DUPILUMAB IN ADULT PARTICIPANTS ON BACKGROUND TOPICAL THERAPY WITH MODERATE TO SEVERE ATOPIC DERMATITIS. 2020.
53. ClinicalTrialsGOV. Study of Abrocitinib Compared With Dupilumab in Adults With Moderate to Severe Atopic Dermatitis on Background Topical Therapy: NCT04345367 / 2019-004013-13 [B7451050]. 2021. Verfügbar unter: <https://ClinicalTrials.gov/show/NCT04345367>. [Zugriff am: 16.11.2021].
54. EU-Clinical Trials Register. A PHASE 3B RANDOMIZED, DOUBLE-BLIND, DOUBLE-DUMMY, ACTIVE CONTROLLED MULTI-CENTER STUDY ASSESSING THE EFFICACY AND SAFETY OF ABROCITINIB COMPARED WITH DUPILUMAB IN ADULT PARTICIPANTS ON BACKGROUND: 2019-004013-13 [B7451050]. 0000. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2019-004013-13. [Zugriff am: 16.11.2021].
55. WHO ICTRP. Study of Abrocitinib Compared With Dupilumab in Adults With Moderate to Severe Atopic Dermatitis on Background Topical Therapy: NCT04345367 / 2019-004013-13 [B7451050]. 2021. Verfügbar unter:

<http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04345367>. [Zugriff am: 16.11.2021].

56. Pfizer Inc. Full Clinical Study Report B7451050: A Phase 3b Randomized, Double-Blind, Double-Dummy, Active-Controlled Multi-Center Study Assessing the Efficacy and Safety of Abrocitinib Compared with Dupilumab in Adult Participants on Background Topical Therapy With Moderate to Severe Atopic Dermatitis. 2021.
57. Pfizer Deutschland GmbH. Zusatzanalysen zur Studie B7451050 (JADE DARE). 2021.
58. Reich K, Bruin-Weller M de, Deleuran M, Beck LA, Papp K, Werfel T et al. Abrocitinib Efficacy and Safety as Monotherapy Over 48 Weeks: Results From a Long-Term Extension Study. Conference: European Academy of Dermatology and Venereology Virtual Congress, October 29-31, 2020 (Abstract No. 3137). 2020.
59. Pfizer Inc. Final Statistical Analysis Plan: A PHASE 3B RANDOMIZED, DOUBLE-BLIND, DOUBLE DUMMY, ACTIVE CONTROLLED MULTI-CENTER STUDY ASSESSING THE EFFICACY AND SAFETY OF ABROCITINIB COMPARED WITH DUPILUMAB IN ADULT PARTICIPANTS ON BACKGROUND TOPICAL THERAPY WITH MODERATE TO SEVERE ATOPIC DERMATITIS. 2021.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	07.11.2016	
Zeitsegment	1974 to 2016 November 04	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ²¹] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Diabetes Mellitus/	552986
2	Non Insulin Dependent Diabetes Mellitus/	195234
3	(diabet* or niddm or t2dm).ab,ti.	714228
4	or/1-3	847068
5	linagliptin*.mp.	1562
6	(random* or double-blind*).tw.	1193849
7	placebo*.mp.	388057
8	or/6-7	1382838
9	and/4,5,8	633

²¹ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	16.11.2021	
Zeitsegment	1974 to 2021 November 15	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 ^a	
#	Suchbegriffe	Ergebnis
1	exp abrocitinib/	162
2	abrocitinib.mp.	170
3	cibinqo.mp.	0
4	(PF-04965842 OR PF04965842 OR "PF 04965842").mp.	49
5	1 OR 2 OR 3 OR 4	194
6	exp dermatitis/	166686
7	dermatit*.mp.	140691
8	neurodermatit*.mp.	2678
9	atopic dermatit*.mp.	53028
10	eczema.mp.	43814
11	(prurig* besnier* OR besnier* prurig*).mp.	22
12	6 OR 7 OR 8 OR 9 OR 10 OR 11	190065
13	random*.tw.	1724644
14	placebo*.mp.	484181
15	double-blind*.tw.	224903
16	13 OR 14 OR 15	1988904
17	5 AND 12 AND 16	93
18	limit 17 to (english or german)	93
19	remove duplicates from 18	90
a: 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.		

Datenbankname	MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions(R)	
Suchoberfläche	Ovid	
Datum der Suche	16.11.2021	
Zeitsegment	1946 to November 15, 2021	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 ^a	
#	Suchbegriffe	Ergebnis
1	abrocitinib.mp.	55
2	cibinqo.mp.	0
3	(PF-04965842 OR PF04965842 OR "PF 04965842").mp.	7
4	1 OR 2 OR 3	57
5	exp dermatitis/	109038
6	dermatit*.mp.	98535
7	neurodermatit*.mp.	1770
8	atopic dermatit*.mp.	23575
9	eczema.mp.	23024
10	(prurig* besnier* OR besnier* prurig*).mp.	68
11	5 OR 6 OR 7 OR 8 OR 9 OR 10	136518
12	randomi#ed controlled trial.pt.	550333
13	randomi#ed.mp.	969445
14	placebo*.mp.	244865
15	12 OR 13 OR 14	1042110
16	4 AND 11 AND 15	24
17	limit 16 to (english or german)	24
18	remove duplicates from 17	24
a: 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.		

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	16.11.2021	
Zeitsegment	1991 – October 2021	
Suchfilter	Keine	
#	Suchbegriffe	Ergebnis
1	abrocitinib.mp.	47
2	cibinqo.mp.	0
3	(PF-04965842 OR PF04965842 OR "PF 04965842").mp.	42
4	1 OR 2 OR 3	87
5	exp dermatitis/	4296
6	dermatit*.mp.	9480
7	neurodermatit*.mp.	149
8	atopic dermatit*.mp.	4799
9	eczema.mp.	4228
10	(prurig* besnier* OR besnier* prurig*).mp.	8
11	5 OR 6 OR 7 OR 8 OR 9 OR 10	11963
12	4 AND 11	69
13	remove duplicates from 12	69

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. clinicaltrials.gov), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07.11.2016
Eingabeoberfläche	Standard Search
Suchstrategie	linagliptin OR BI 1356
Treffer	169

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Studienregister/ Studienergebnisdatenbank	clinicaltrials.gov
Internetadresse	https://clinicaltrials.gov/ct2/home
Datum der Suche	16.11.2021
Eingabeoberfläche	Other terms
Suchstrategie	Abrocitinib OR Cibirinco OR PF-04965842 OR PF04965842 OR "PF 04965842"
Treffer	32

Studienregister/ Studienergebnisdatenbank	EU Clinical Trials Register
Internetadresse	https://www.clinicaltrialsregister.eu/ctr-search/search
Datum der Suche	16.11.2021
Eingabeoberfläche	Standard Search
Suchstrategie	Abrocitinib OR Cibirinco OR PF-04965842 OR PF04965842 OR "PF 04965842"
Treffer	9

Studienregister/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	https://trialsearch.who.int/Default.aspx
Datum der Suche	16.11.2021
Eingabeoberfläche	Standard Search
Suchstrategie	Abrocitinib OR Cibirno OR PF-04965842 OR PF04965842 OR "PF 04965842"
Treffer	29

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der /den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

1. Nusbaum KB, Fleischer S, Fleischer AB. Efficacy of biologics and oral small molecules for atopic dermatitis: a systematic review and meta-analysis. *J Dermatolog Treat.* 2021):1–11. doi:10.1080/09546634.2021.1986204.
Ausschlussgrund: A6
2. Zhou S, Qi F, Gong Y, Zhang J, Zhu B. Biological Therapies for Atopic Dermatitis: A Systematic Review. *Dermatology.* 2021:1–11. doi:10.1159/000514535.
Ausschlussgrund: A3

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken)

Listen Sie nachfolgend die durch die Studienregistersuche(n)/ Studienergebnisdatenbanksuche(n) identifizierten, aber ausgeschlossenen Registereinträgen auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

1. ClinicalTrialsGOV 2014. *A Study to Evaluate the Bioavailability of PF 04965842 in Healthy Subjects: NCT02163161*. Studiennummer des Sponsors: B7451004. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT02163161>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
2. ClinicalTrialsGOV 2014. *First-in-Human Study to Evaluate Safety and Tolerability of Single and Multiple Ascending Doses of Janus Kinase-1 Inhibitor PF-04965842 in Healthy Western and Japanese Subjects: NCT01835197*. Studiennummer des Sponsors: B7451001. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT01835197>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
3. ClinicalTrialsGOV 2016. *Study to Evaluate PF-04965842 in Patients With Moderate to Severe Psoriasis: NCT02201524*. Studiennummer des Sponsors: B7451005 JAK-1 FOR PSORIASIS. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT02201524>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
4. ClinicalTrialsGOV 2017. *Absorption, Metabolism, Excretion and Absolute Bioavailability: NCT03250039 / 2017-000461-73*. Studiennummer des Sponsors: B7451008. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03250039>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
5. ClinicalTrialsGOV 2019. *A Study to Assess Pharmacokinetics of PF-04965842 and Its Metabolites and Effect of Probenecid in Healthy Participants: NCT03937258*. Studiennummer des Sponsors: B7451043. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03937258>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
6. ClinicalTrialsGOV 2019. *A Study To Determine The Effect Of Single-dose PF-04965842 On QTc Interval In Healthy Volunteers Compared With Placebo and Moxifloxacin: NCT03386279 / 2017-004472-61*. Studiennummer des Sponsors: B7451027. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03386279>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
7. ClinicalTrialsGOV 2019. *Drug Drug Interaction Study Evaluating the Effect of Fluvoxamine or Fluconazole on PK and Safety of PF-04965842: NCT03634345 / 2018-*

- 001943-29. Studiennummer des Sponsors: B7451017. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03634345>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
8. ClinicalTrialsGOV 2019. *Study of PF-04965842 Effect on Rosuvastatin Pharmacokinetics in Healthy Participants: NCT03806101 / 2018-003425-28*. Studiennummer des Sponsors: B7451033 DDI. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03806101>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
9. ClinicalTrialsGOV 2019. *Study to Evaluate Efficacy and Safety of PF-04965842 in Subjects Aged 12 Years And Older With Moderate to Severe Atopic Dermatitis: NCT03349060 / 2017-003651-29*. Studiennummer des Sponsors: B7451012 MONO-1 JADE MONO-1. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03349060>, abgerufen am: 16.11.2021.
Ausschlussgrund: A6.
10. ClinicalTrialsGOV 2019. *Study To Evaluate Pf-04965842 In Subjects With Moderate To Severe Atopic Dermatitis: NCT02780167 / 2015-005513-72*. Studiennummer des Sponsors: B7451006. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT02780167>, abgerufen am: 16.11.2021.
Ausschlussgrund: A6.
11. ClinicalTrialsGOV 2020. *A Hepatic Impairment Study for PF-04965842: NCT03626415*. Studiennummer des Sponsors: B7451020. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03626415>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
12. ClinicalTrialsGOV 2020. *A Renal Impairment Study for PF-04965842: NCT03660241 / 2018-002865-20*. Studiennummer des Sponsors: B7451021. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03660241>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
13. ClinicalTrialsGOV 2020. *A Study to Determine the Effects of PF-04965842 on Midazolam PK in Healthy Volunteers: NCT03647670 / 2018-001198-26*. Studiennummer des Sponsors: B7451022. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03647670>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
14. ClinicalTrialsGOV 2020. *A Study to Determine the Effects of PF-04965842 on the Pharmacokinetics of Oral Contraceptive Steroids in Healthy Female Subjects: NCT03662516*. Studiennummer des Sponsors: B7451016. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03662516>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
15. ClinicalTrialsGOV 2020. *Drug Interaction Study Evaluating the Effect of Rifampin on PK and Safety of PF 04965842: NCT03637790 / 2018-002167-25*. Studiennummer des Sponsors: B7451019. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03637790>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
16. ClinicalTrialsGOV 2020. *Study Evaluating Efficacy and Safety of PF-04965842 in Subjects Aged 12 Years And Older With Moderate to Severe Atopic Dermatitis: NCT03575871 / 2018-001136-21*. Studiennummer des Sponsors: B7451013 MONO-2. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03575871>, abgerufen am: 16.11.2021.
Ausschlussgrund: A3.

17. ClinicalTrialsGOV 2020. *Study of Commercial and Phase 3 of PF-04965842 Formulations, Estimation of Effect of Food on Commercial Formulation: NCT04065633*. Studiennummer des Sponsors: B7451032. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT04065633>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
18. ClinicalTrialsGOV 2020. *Study of PF 04965842 Effect on MATE1/2K Activity in Healthy Participants: NCT03796182 / 2018-003683-31*. Studiennummer des Sponsors: B7451034 DDI. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03796182>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
19. ClinicalTrialsGOV 2020. *Study of PF-04965842 Effect on Dabigatran Pharmacokinetics in Healthy Participants: NCT03742336 / 2018-003579-36*. Studiennummer des Sponsors: B7451026. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03742336>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
20. ClinicalTrialsGOV 2020. *Study to Assess Pharmacokinetics, Safety and Tolerability of PF-04965842 in Chinese Healthy Participants: NCT04099563*. Studiennummer des Sponsors: B7451028. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT04099563>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
21. ClinicalTrialsGOV 2021. *A Study Evaluating Relative Bioavailability of an Oral Suspension of Abrocitinib and Effect of an Acid Reducing Agent on the Bioavailability of Abrocitinib and Assessing the Taste of Abrocitinib Oral Formulations: NCT04903093*. Studiennummer des Sponsors: B7451061. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT04903093>, abgerufen am: 16.11.2021.
Ausschlussgrund: A2.
22. ClinicalTrialsGOV 2021. *A Study to Estimate the Effect of Multiple Dose Abrocitinib on Caffeine, Efavirenz, and Omeprazole in Healthy Participants: NCT05067439*. Studiennummer des Sponsors: B7451092. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT05067439>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
23. ClinicalTrialsGOV 2021. *Abrocitinib Expanded Access Protocol in Adolescents and Adults With Moderate to Severe Atopic Dermatitis: NCT04564755 / 2020-003610-12*. Studiennummer des Sponsors: B7451064 JADE REAL. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT04564755>, abgerufen am: 16.11.2021.
Ausschlussgrund: A5.
24. ClinicalTrialsGOV 2021. *Efficacy of Abrocitinib for Reducing Pruritus in Adults With Prurigo Nodularis and Chronic Pruritus of Unknown Origin: NCT05038982*. Studiennummer des Sponsors: IRB00262268. Johns Hopkins University (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT05038982>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
25. ClinicalTrialsGOV 2021. *JAK Inhibition in Food Allergy: NCT05069831*. Studiennummer des Sponsors: GCO 21-0781. Icahn School of Medicine at Mount Sinai (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT05069831>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
26. ClinicalTrialsGOV 2021. *JAK1 Inhibitor With Medicated Topical Therapy in Adolescents With Atopic Dermatitis: NCT03796676 / 2018-003804-37*. Studiennummer des Sponsors:

- B7451036 JADE TEEN*. Pfizer (Hrsg.). Verfügbar unter:
<https://ClinicalTrials.gov/show/NCT03796676>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
27. ClinicalTrialsGOV 2021. *Molecular Signatures in Inflammatory Skin Disease: NCT03358693*. Studiennummer des Sponsors: A100/12. Prof. Dr. Stephan Weidinger (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03358693>, abgerufen am: 16.11.2021.
Ausschlussgrund: A5.
28. ClinicalTrialsGOV 2021. *Study Evaluating Efficacy and Safety of PF-04965842 and Dupilumab in Adult Subjects With Moderate to Severe Atopic Dermatitis on Background Topical Therapy: NCT03720470 / 2018-002573-21*. Studiennummer des Sponsors: B7451029 COMPARE. Pfizer (Hrsg.). Verfügbar unter:
<https://ClinicalTrials.gov/show/NCT03720470>, abgerufen am: 16.11.2021.
Ausschlussgrund: A6.
29. ClinicalTrialsGOV 2021. *Study Evaluating the Mechanism of Action of PF-04965842 Monotherapy for Moderate-to-severe Atopic Dermatitis: NCT03915496*. Studiennummer des Sponsors: B7451037 JADE MOA. Pfizer (Hrsg.). Verfügbar unter:
<https://ClinicalTrials.gov/show/NCT03915496>, abgerufen am: 16.11.2021.
Ausschlussgrund: A3.
30. ClinicalTrialsGOV 2021. *Study to Evaluate Efficacy and Safety of PF-04965842 With or Without Topical Medications in Subjects Aged 12 Years and Older With Moderate to Severe Atopic Dermatitis: NCT03422822 / 2017-004851-22*. Studiennummer des Sponsors: B7451015 JADE EXTEND B745105 JADE EXTEND. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03422822>, abgerufen am: 16.11.2021.
Ausschlussgrund: A3.
31. ClinicalTrialsGOV 2021. *Study to Investigate Efficacy and Safety of PF-04965842 in Subjects Aged 12 Years and Over With Moderate to Severe Atopic Dermatitis With the Option of Rescue Treatment in Flaring Subjects: NCT03627767 / 2018-000501-23*. Studiennummer des Sponsors: B7451014 JADE REGIMEN REGIMEN. Pfizer (Hrsg.). Verfügbar unter: <https://ClinicalTrials.gov/show/NCT03627767>, abgerufen am: 16.11.2021.
Ausschlussgrund: A3.
32. EU-Clinical Trials Register 0000. *A PHASE 2B RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL, MULTICENTER, DOSE-RANGING, STUDY TO EVALUATE THE EFFICACY AND SAFETY PROFILE OF PF-04965842 IN SUBJECTS WITH MODERATE TO SEVERE AT: 2015-005513-72*. Studiennummer des Sponsors: B7451006. Pfizer Inc, 235 East 42nd Street, New York, NY 10017 (Hrsg.). Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-005513-72, abgerufen am: 16.11.2021.
Ausschlussgrund: A6.
33. EU-Clinical Trials Register 0000. *A phase 3 multi-center, long-term extension study investigating the efficacy and safety of Abrocitinib, with or without topical medications, administered to subjects aged 12 years and older with m: 2017-004851-22*. Studiennummer des Sponsors: B7451015. Pfizer Inc., 235 East 42nd Street, New York, NY 10017 (Hrsg.). Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-004851-22, abgerufen am: 16.11.2021.
Ausschlussgrund: A3.
34. EU-Clinical Trials Register 0000. *A phase 3 randomized, double-blind, placebo-*

- controlled, parallel group, multi-center study to evaluate the efficacy and safety of PF-04965842 monotherapy in subjects aged 12 years and older, with m: 2017-003651-29.* Studiennummer des Sponsors: B7451012. Pfizer Inc., 235 East 42nd Street, New York, NY 10017 (Hrsg.). Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-003651-29, abgerufen am: 16.11.2021.
Ausschlussgrund: A6.
35. EU-Clinical Trials Register 0000. *A Phase 3 Randomized, Double-Blind, Placebo-Controlled, Parallel Group, Multi-Center Study to Evaluate the Efficacy and Safety of PF-04965842 Monotherapy in Subjects Aged 12 Years and Older, with M: 2018-001136-21.* Studiennummer des Sponsors: B7451013. Pfizer Inc., 235 East 42nd Street, New York, NY 10017 (Hrsg.). Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-001136-21, abgerufen am: 16.11.2021.
Ausschlussgrund: A3.
36. EU-Clinical Trials Register 0000. *A PHASE 3 RANDOMIZED, DOUBLE-BLIND, DOUBLE-DUMMY, PLACEBO-CONTROLLED, PARALLEL GROUP, MULTI-CENTER STUDY INVESTIGATING THE EFFICACY AND SAFETY OF PF-04965842 AND DUPILUMAB IN COMPARISON WITH PLACEB: 2018-002573-21.* Studiennummer des Sponsors: B7451029. Pfizer Inc. (Hrsg.). Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-002573-21, abgerufen am: 16.11.2021.
Ausschlussgrund: A6.
37. EU-Clinical Trials Register 0000. *A phase 3 randomized withdrawal, double-blind, placebo-controlled, multi-center study investigating the efficacy and safety of PF-04965842 in subjects aged 12 years and over, with moderate to severe: 2018-000501-23.* Studiennummer des Sponsors: B7451014. Pfizer Inc., 235 East 42nd Street, New York, NY 10017 (Hrsg.). Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-000501-23, abgerufen am: 16.11.2021.
Ausschlussgrund: A3.
38. EU-Clinical Trials Register 0000. *A PHASE 3, RANDOMIZED, DOUBLE BLIND, PLACEBO CONTROLLED, MULTI CENTER STUDY INVESTIGATING THE EFFICACY AND SAFETY OF PF 04965842 CO ADMINISTERED WITH BACKGROUND MEDICATED TOPICAL THERAPY IN ADOLESC: 2018-003804-37.* Studiennummer des Sponsors: B7451036. Pfizer Inc. (Hrsg.). Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-003804-37, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
39. EU-Clinical Trials Register 0000. *ABROCITINIB EXPANDED ACCESS PROTOCOL FOR THE TREATMENT OF ADOLESCENTS AND ADULTS WITH MODERATE TO SEVERE ATOPIC DERMATITIS: 2020-003610-12.* Studiennummer des Sponsors: B7451064. Pfizer Inc. (Hrsg.). Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2020-003610-12, abgerufen am: 16.11.2021.
Ausschlussgrund: A5.
40. WHO ICTRP 2015. *A Study to Evaluate the Bioavailability of PF 04965842 in Healthy Subjects: NCT02163161.* Studiennummer des Sponsors: B7451004. Pfizer (Hrsg.). Verfügbar unter: <https://trialsearch.who.int/Trial2.aspx?TrialID=NCT02163161>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
41. WHO ICTRP 2015. *First-in-Human Study to Evaluate Safety and Tolerability of Single*

- and Multiple Ascending Doses of Janus Kinase-1 Inhibitor PF-04965842 in Healthy Western and Japanese Subjects: NCT01835197.* Studiennummer des Sponsors: B7451001. Pfizer (Hrsg.). Verfügbar unter:
<https://trialssearch.who.int/Trial2.aspx?TrialID=NCT01835197>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
42. WHO ICTRP 2017. *Study to Evaluate PF-04965842 in Patients With Moderate to Severe Psoriasis: NCT02201524.* Studiennummer des Sponsors: B7451005 JAK-1 FOR PSORIASIS. Pfizer (Hrsg.). Verfügbar unter:
<https://trialssearch.who.int/Trial2.aspx?TrialID=NCT02201524>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
43. WHO ICTRP 2018. *Absorption, Metabolism, Excretion and Absolute Bioavailability: NCT03250039 / 2017-000461-73.* Studiennummer des Sponsors: B7451008. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03250039>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
44. WHO ICTRP 2020. *A Hepatic Impairment Study for PF-04965842: NCT03626415.* Studiennummer des Sponsors: B7451020. Pfizer (Hrsg.). Verfügbar unter:
<https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03626415>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
45. WHO ICTRP 2020. *A Renal Impairment Study for PF-04965842: NCT03660241 / 2018-002865-20.* Studiennummer des Sponsors: B7451021. Pfizer (Hrsg.). Verfügbar unter:
<https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03660241>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
46. WHO ICTRP 2020. *A Study to Assess Pharmacokinetics of PF-04965842 and Its Metabolites and Effect of Probenecid in Healthy Participants: NCT03937258.* Studiennummer des Sponsors: B7451043. Pfizer (Hrsg.). Verfügbar unter:
<https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03937258>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
47. WHO ICTRP 2020. *A Study to Determine the Effects of PF-04965842 on Midazolam PK in Healthy Volunteers: NCT03647670 / 2018-001198-26.* Studiennummer des Sponsors: B7451022. Pfizer (Hrsg.). Verfügbar unter:
<https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03647670>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
48. WHO ICTRP 2020. *A Study to Determine the Effects of PF-04965842 on the Pharmacokinetics of Oral Contraceptive Steroids in Healthy Female Subjects: NCT03662516.* Studiennummer des Sponsors: B7451016. Pfizer (Hrsg.). Verfügbar unter:
<https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03662516>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
49. WHO ICTRP 2020. *A Study To Determine The Effect Of Single-dose PF-04965842 On QTc Interval In Healthy Volunteers Compared With Placebo and Moxifloxacin: NCT03386279 / 2017-004472-61.* Studiennummer des Sponsors: B7451027. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03386279>, abgerufen am: 16.11.2021.
Ausschlussgrund: A1.
50. WHO ICTRP 2020. *Drug Drug Interaction Study Evaluating the Effect of Fluvoxamine or Fluconazole on PK and Safety of PF-04965842: NCT03634345 / 2018-001943-29.* Studiennummer des Sponsors: B7451017. Pfizer (Hrsg.). Verfügbar unter:
<https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03634345>, abgerufen am: 16.11.2021.

- Ausschlussgrund: A1.
51. WHO ICTRP 2020. *Drug Interaction Study Evaluating the Effect of Rifampin on PK and Safety of PF 04965842: NCT03637790 / 2018-002167-25*. Studiennummer des Sponsors: B7451019. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03637790>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.
52. WHO ICTRP 2020. *JAK1 Inhibitor With Medicated Topical Therapy in Adolescents With Atopic Dermatitis: NCT03796676 / 2018-003804-37*. Studiennummer des Sponsors: B7451036 JADE TEEN. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03796676>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.
53. WHO ICTRP 2020. *Study Evaluating Efficacy and Safety of PF-04965842 in Subjects Aged 12 Years And Older With Moderate to Severe Atopic Dermatitis: NCT03575871 / 2018-001136-21*. Studiennummer des Sponsors: B7451013 MONO-2. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03575871>, abgerufen am: 16.11.2021. Ausschlussgrund: A3.
54. WHO ICTRP 2020. *Study Evaluating Efficacy and Safety of PF-04965842 and Dupilumab in Adult Subjects With Moderate to Severe Atopic Dermatitis on Background Topical Therapy: NCT03720470 / 2018-002573-21*. Studiennummer des Sponsors: B7451029 COMPARE. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03720470>, abgerufen am: 16.11.2021. Ausschlussgrund: A6.
55. WHO ICTRP 2020. *Study of Commercial and Phase 3 of PF-04965842 Formulations, Estimation of Effect of Food on Commercial Formulation: NCT04065633*. Studiennummer des Sponsors: B7451032. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT04065633>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.
56. WHO ICTRP 2020. *Study of PF 04965842 Effect on MATE1/2K Activity in Healthy Participants: NCT03796182 / 2018-003683-31*. Studiennummer des Sponsors: B7451034 DDI. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03796182>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.
57. WHO ICTRP 2020. *Study of PF-04965842 Effect on Dabigatran Pharmacokinetics in Healthy Participants: NCT03742336 / 2018-003579-36*. Studiennummer des Sponsors: B7451026. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03742336>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.
58. WHO ICTRP 2020. *Study of PF-04965842 Effect on Rosuvastatin Pharmacokinetics in Healthy Participants: NCT03806101 / 2018-003425-28*. Studiennummer des Sponsors: B7451033 DDI. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03806101>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.
59. WHO ICTRP 2020. *Study to Assess Pharmacokinetics, Safety and Tolerability of PF-04965842 in Chinese Healthy Participants: NCT04099563*. Studiennummer des Sponsors: B7451028. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT04099563>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.

60. WHO ICTRP 2020. *Study to Evaluate Efficacy and Safety of PF-04965842 in Subjects Aged 12 Years And Older With Moderate to Severe Atopic Dermatitis: NCT03349060 / 2017-003651-29*. Studiennummer des Sponsors: B7451012 JADE MONO-1 MONO-1. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03349060>, abgerufen am: 16.11.2021. Ausschlussgrund: A6.
61. WHO ICTRP 2020. *Study To Evaluate Pf-04965842 In Subjects With Moderate To Severe Atopic Dermatitis: NCT02780167 / 2015-005513-72*. Studiennummer des Sponsors: B7451006. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT02780167>, abgerufen am: 16.11.2021. Ausschlussgrund: A6.
62. WHO ICTRP 2021. *A Study Evaluating Relative Bioavailability of an Oral Suspension of Abrocitinib and Effect of an Acid Reducing Agent on the Bioavailability of Abrocitinib and Assessing the Taste of Abrocitinib Oral Formulations: NCT04903093*. Studiennummer des Sponsors: B7451061. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT04903093>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.
63. WHO ICTRP 2021. *Abrocitinib Expanded Access Protocol in Adolescents and Adults With Moderate to Severe Atopic Dermatitis: NCT04564755 / 2020-003610-12*. Studiennummer des Sponsors: B7451064 JADE REAL. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT04564755>, abgerufen am: 16.11.2021. Ausschlussgrund: A5.
64. WHO ICTRP 2021. *Efficacy of Abrocitinib for Reducing Pruritus in Adults With Prurigo Nodularis and Chronic Pruritus of Unknown Origin: NCT05038982*. Studiennummer des Sponsors: IRB00262268. Johns Hopkins University (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT05038982>, abgerufen am: 16.11.2021. Ausschlussgrund: A1.
65. WHO ICTRP 2021. *Molecular Signatures in Inflammatory Skin Disease: NCT03358693*. Studiennummer des Sponsors: A100/12. Prof. Dr. Stephan Weidinger (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03358693>, abgerufen am: 16.11.2021. Ausschlussgrund: A5.
66. WHO ICTRP 2021. *Study Evaluating the Mechanism of Action of PF-04965842 Monotherapy for Moderate-to-severe Atopic Dermatitis: NCT03915496*. Studiennummer des Sponsors: B7451037 JADE MOA. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03915496>, abgerufen am: 16.11.2021. Ausschlussgrund: A3.
67. WHO ICTRP 2021. *Study to Investigate Efficacy and Safety of PF-04965842 in Subjects Aged 12 Years and Over With Moderate to Severe Atopic Dermatitis With the Option of Rescue Treatment in Flaring Subjects: NCT03627767 / 2018-000501-23*. Studiennummer des Sponsors: B7451014 JADE REGIMEN REGIMEN. Pfizer (Hrsg.). Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03627767>, abgerufen am: 16.11.2021. Ausschlussgrund: A3.

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-128 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 4-128 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-128 (Anhang): Studiendesign und -methodik für Studie JADE DARE

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Primäres Ziel</p> <p>Vergleich der Wirksamkeit von Abrocitinib (PF-04965842) 200 mg 1-mal täglich mit Dupilumab (gemäß Fachinformation) bei erwachsenen Studienteilnehmern unter topischer Hintergrundtherapie mit mittelschwerer bis schwerer atopischer Dermatitis (AD).</p> <p>Sekundäre Ziele</p> <ul style="list-style-type: none"> • Vergleich der Wirksamkeit von Abrocitinib 200 mg 1-mal täglich mit Dupilumab in Bezug auf zusätzliche Wirksamkeitsendpunkte bei erwachsenen Studienteilnehmern unter topischer Hintergrundtherapie mit mittelschwerer bis schwerer AD. • Vergleich der Sicherheit und Verträglichkeit von Abrocitinib 200 mg 1-mal täglich mit Dupilumab in Bezug auf zusätzliche Wirksamkeitsendpunkte bei erwachsenen Studienteilnehmern unter topischer Hintergrundtherapie mit mittelschwerer bis schwerer AD. <p>Weitere explorative Ziele</p> <ul style="list-style-type: none"> • Vergleich der Wirksamkeit von Abrocitinib 200 mg 1-mal täglich mit Dupilumab in Bezug auf die Inanspruchnahme von Gesundheitsressourcen (<i>Health Care Resource Utilization, HCRU</i>) bei erwachsenen Studienteilnehmern unter topischer Hintergrundtherapie mit mittelschwerer bis schwerer AD. • Untersuchung der Wirkung von Abrocitinib 200 mg 1-mal täglich auf die Häufigkeit und Dauer des nächtlichen Kratzens und die Schlafdauer mittels Accelerometern bei erwachsenen Studienteilnehmern unter topischer Hintergrundtherapie mit mittelschwerer bis schwerer AD mit Hilfe. • Untersuchung der Wirksamkeit von Abrocitinib 200 mg 1-mal täglich bei erwachsenen Studienteilnehmern unter topischer Hintergrundtherapie mit Handekzemen. • Ermöglichung explorativer Analysen durch Entnahme von gelagerten Bioproben, es sei denn, dies war aufgrund lokaler Vorschriften oder einer Entscheidung der Ethikkommission verboten.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>Zuteilung: Randomisiert (Zuteilungsverhältnis 1:1)</p> <p>Randomisierungseinheit: Studienteilnehmer</p> <p>Verblindung: doppelblindes Studiendesign (Prüfarzt, Sponsor und Studienteilnehmer)</p> <p>Studienhorizont: Die Gesamtdauer der Studienteilnahme beträgt 26 Wochen. Studienteilnehmer, die die Behandlung vorzeitig abbrechen oder aus anderen Gründen nicht für die</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Verlängerungsstudie B7451015 in Frage kommen, durchlaufen eine 4-wöchige Nachbeobachtungsphase in der Studie B7451050.</p> <p>Design: randomisiert, aktivkontrolliert, parallel</p> <p>Studienorganisation: multizentrisch, international</p> <p>Phase: III</p>
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Das originale Studienprotokoll stammt vom 04.11.2019. Der originale Statistische Analyseplan (SAP) stammt vom 04.11.2019. Es liegen keine Änderungen des Studienprotokolls und des SAP vor.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <ul style="list-style-type: none"> • Der Studienteilnehmer muss, einschließlich zum Zeitpunkt der Unterzeichnung der Einverständniserklärung ≥ 18 Jahre alt sein. • Klinische Diagnose von chronischer AD (auch bekannt als atopisches Ekzem) seit mindestens 1 Jahr vor Tag 1 und bestätigte AD bei der <i>Screening</i>- und <i>Baseline</i>-Visite gemäß den Kriterien nach Hanafin und Rajka für AD. • Dokumentierte Anamnese (innerhalb von 6 Monaten vor der <i>Screening</i>-Visite) mit unzureichendem Ansprechen auf die Behandlung mit einer medikamentösen topischen Therapie für AD für mindestens 4 Wochen oder eine notwendige systemische Therapie innerhalb des letzten Jahres zur Kontrolle der Erkrankung. Eine medikamentöse topische Therapie ist definiert als ein topisches Produkt, das einen pharmazeutischen Wirkstoff enthält, der für die Behandlung von AD indiziert ist (unabhängig davon, ob es sich um ein rezeptfreies [<i>Over-the-counter</i>, OTC] oder verschreibungspflichtiges Produkt handelt). • Mittelschwere bis schwere AD: von AD betroffene Körperoberfläche (<i>Body Surface Area</i>, BSA) ≥ 10 %, Investigator's Global Assessment (IGA) ≥ 3, <i>Eczema Area and Severity Index</i> (EASI) ≥ 16 und <i>Peak Pruritis Numerical Rating Scale</i> (PP-NRS) ≥ 4 zu <i>Baseline</i> • Die Verwendung von Verhütungsmitteln für Männer oder Frauen sollte mit den lokalen Vorschriften bezüglich der Verhütungsmethoden für Studienteilnehmer in klinischen Studien übereinstimmen. <ul style="list-style-type: none"> ○ Für männliche Studienteilnehmer waren keine empfängnisverhütenden Maßnahmen erforderlich. ○ Weibliche Studienteilnehmerinnen waren zur Studienteilnahme berechtigt, wenn sie nicht schwanger waren oder stillen und mindestens eine der folgenden Bedingungen zutraf: <ul style="list-style-type: none"> ▪ Studienteilnehmerinnen, die nicht im gebärfähigen Alter waren.

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ▪ Für Studienteilnehmerinnen im gebärfähigen Alter (alle weiblichen Studienteilnehmerinnen, unabhängig von dem Status der Menarche, es sei denn, sie waren dauerhaft steril oder nachweislich unfruchtbar) gilt: Bei sexueller Aktivität, muss während des Interventionszeitraums und für mindestens 28 Tage nach der letzten Dosis der Studienintervention eine hochwirksame Methode der Schwangerschaftsverhütung mit einer Versagensrate von < 1 % angewendet werden. Der Prüfarzt sollte die Wirksamkeit der Verhütungsmethode vor der ersten Dosis der Studienintervention bewerten. ▪ Studienteilnehmerinnen im gebärfähigen Alter mussten bei der <i>Screening</i>-Visite einen negativen hochsensitiven Serum-Schwangerschaftstest aufweisen. Ein Urin-Schwangerschaftstest mit einer Sensitivität von ≥ 25 mIU/ml wurde vor der ersten Dosis der Studienintervention und bei jeder Visite vor Ort, einschließlich bis Behandlungsende (<i>End of treatment</i>, EOT) und der Nachuntersuchungen, durchgeführt, um zu bestätigen, dass die Testperson nicht schwanger geworden ist. Wenn ein Urintest nicht als negativ bestätigt werden konnte (z. B. ein mehrdeutiges Ergebnis), war ein Serum-Schwangerschaftstest erforderlich. In solchen Fällen musste die Studienteilnehmerin von der Studienteilnahme ausgeschlossen werden, wenn das Ergebnis des Serum-Schwangerschaftstests positiv war. ▪ Der Prüfarzt war für die Überprüfung der medizinischen Anamnese, der Menstruationsanamnese und der jüngsten sexuellen Aktivität verantwortlich, um das Risiko für den Einschluss von Studienteilnehmerinnen mit einer frühen unerkannten Schwangerschaft zu verringern <ul style="list-style-type: none"> • Der Studienteilnehmer musste in der Lage sein, eine unterzeichnete Einwilligungserklärung abzugeben, die die Einhaltung der in der Einwilligungserklärung und in diesem Protokoll aufgeführten Anforderungen und Einschränkungen beinhaltet. • Zur Behandlung der AD konnte der Studienteilnehmer eine leichte oder mittelstark wirksame medikamentöse und nicht-

Item ^a	Charakteristikum	Studieninformation
		<p>medikamentöse topische Therapie anwenden, wobei das Ansprechen auf die Behandlung zu Studienbeginn unzureichend sein muss. Der Studienteilnehmer musste außerdem bereit und in der Lage sein, während der gesamten Dauer der Studie eine standardisierte topische Hintergrundtherapie einzuhalten.</p> <ul style="list-style-type: none"> • Der Studienteilnehmer musste bereit und in der Lage sein, die geplanten Visiten, den Behandlungsplan, die Labortests, die Hinweise zum Lebensstil und andere Studienverfahren einzuhalten. • Der Studienteilnehmer musste sich bereit erklären, längere Sonnenexposition zu vermeiden und während der Studie keine Solarien, Sonnenlampen oder andere ultraviolette Lichtquellen zu verwenden. • Wenn die Studienteilnehmer Begleitmedikamente aus einem anderen Grund als AD erhielten, mussten sie ein stabiles Therapieschema einhalten, d. h. sie durften innerhalb von 7 Tagen oder 5 Halbwertszeiten (je nachdem, was länger ist) vor Tag 1 und während der gesamten Studiendauer kein neues Medikament einnehmen oder die Dosierung ändern. <p>Ausschlusskriterien</p> <ul style="list-style-type: none"> • Andere akute oder chronische medizinische oder Laboranomalien, die das Risiko im Zusammenhang mit der Studienteilnahme oder der Verabreichung der Studienintervention erhöhen oder die Interpretation der Studienergebnisse beeinträchtigen können und nach dem Urteil des Prüfarztes den Studienteilnehmer für die Studienteilnahme ungeeignet machen würden. • Der Studienteilnehmer sollte eine Risikobewertung durch eine qualifizierte psychosoziale Fachkraft (<i>Mental health professional</i>, MHP) durchführen lassen, um zu beurteilen, ob die Studienteilnahme sicher ist, wenn die Antworten des Studienteilnehmers auf eines der <i>Screening</i>-Instrumente oder andere <i>Screening</i>-Informationen hinweisen: <ul style="list-style-type: none"> ○ Selbstmordgedanken in Verbindung mit konkreter Absicht und einer Methode oder einem Plan im vergangenen Jahr: "Ja" als Antwort auf die Items 4 oder 5 der <i>Columbia Suicide Severity Rating Scale</i> (C-SSRS) ○ Vorgeschichte von suizidalem Verhalten in den letzten 5 Jahren: "Ja" als Antwort (für Ereignisse, die in den letzten 5 Jahren aufgetreten waren) auf eines der Items zu suizidalem Verhalten des C-SSRS ○ Jegliche Anamnese von schwerem oder wiederkehrendem suizidalem Verhalten (nicht-suizidales selbstverletzendes Verhalten ist kein Auslöser für eine Risikobewertung, es sei denn, es ist nach Einschätzung des Prüfarztes indiziert)

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ○ Klinisch signifikante Depression: <i>Patient Health Questionnaire 8-Items (PHQ-8)</i> Gesamtwert von ≥ 15 ○ Das Vorhandensein einer aktuellen schweren psychischen Störung, die in den Ein-/ Ausschlusskriterien nicht explizit aufgeführt ist ○ Nach dem Urteil des Prüfarztes ist eine Risikobewertung oder ein Ausschluss erforderlich • Eine aktuelle oder frühere Krankheitsanamnese von Thrombozytopenie, Koagulopathie oder Thrombozytenfunktionsstörung. • Einnahme von Antikoagulanzen oder Medikamenten, die bekanntermaßen eine Thrombozytopenie verursachen (es sei denn, es wurde als sicher erachtet, diese abzusetzen und für die Dauer der Studie auszuschleichen). • Derzeit aktive Formen anderer entzündlicher Hauterkrankungen, d. h. keine AD, oder Anzeichen von Hauterkrankungen (z. B. Psoriasis, seborrhoische Dermatitis, Lupus) zum Zeitpunkt von Tag 1, die die Bewertung der AD oder das Ansprechen auf die Behandlung beeinträchtigen würden. • Anamnese einer lymphoproliferativen Erkrankung wie Epstein-Barr-Virus (EBV)-bezogene lymphoproliferative Erkrankung, Anamnese eines Lymphoms, einer Leukämie oder Anzeichen oder Symptome, die auf eine aktuelle lymphatische oder lymphoide Erkrankung hindeuteten. • Anamnese einer systemischen Infektion, die einen Krankenhausaufenthalt oder eine parenterale antimikrobielle Therapie erforderte oder anderweitig vom Prüfarzt als klinisch bedeutsam beurteilt wurde, innerhalb von 6 Monaten vor Tag 1. • Eine bekannte Helminthen-Infektion. • Aktive chronische oder akute Hautinfektion, die eine Behandlung mit systemischen Antimikrobika innerhalb von 2 Wochen vor Tag 1 erforderte oder oberflächliche Hautinfektionen innerhalb von einer Woche vor Tag 1. • Bekannte Infektion mit dem humanen Immundefizienz-Virus (HIV), Hepatitis B oder Hepatitis C. <ul style="list-style-type: none"> ○ Studienteilnehmer, die beim <i>Screening</i> Hepatitis-B-Oberflächenantigen-negativ (HBsAg-), Hepatitis-B-Kernantikörper-positiv (HBcAb+) und Hepatitis-B-Oberflächenantikörper-positiv (HBsAb+) waren, wurden auf Hepatitis-B-Virus-Desoxyribonukleinsäure (HBV-DNA) getestet. Studienteilnehmer, die HBV-DNA oberhalb der unteren Bestimmungsgrenze (<i>Lower limit of quantification, LLQ</i>) aufwiesen, wurden ausgeschlossen. Studienteilnehmer, die HBV-DNA negativ oder unterhalb der LLQ waren, konnten randomisiert werden. Studienteilnehmer, die HBV-DNA oberhalb der LLQ aufwiesen, wurden

Item ^a	Charakteristikum	Studieninformation
		<p>ausgeschlossen. Studienteilnehmer, die HBV-DNA-negativ oder unterhalb der LLQ waren, konnten randomisiert werden, mussten aber zu Woche 16 und Woche 26 zum EOT oder zum vorzeitigen Studienabbruch (<i>Early discontinuation</i>, ED) erneut auf HBV-DNA getestet werden, je nachdem, welcher Zeitpunkt früher liegt.</p> <ul style="list-style-type: none"> • Anamnese (einmalige Episode) von disseminiertem Herpes zoster oder disseminiertem Herpes simplex oder ein rezidivierender (mehr als eine Episode) lokalisierter, dermatomaler Herpes zoster. • Bekannte Immundefizienz oder ein Verwandter ersten Grades mit einer erblichen Immundefizienz. • Vorgeschichte von Alkohol- oder Substanzmissbrauch innerhalb von 6 Monaten vor Tag 1, die nach Meinung des Prüfarztes eine Studienteilnahme an der Studie ausschließt. • Bösartige Erkrankungen oder eine Vorgeschichte von bösartigen Erkrankungen mit Ausnahme von adäquat behandeltem oder exzidiertem nicht-metastasiertem Basalzell- oder Plattenepithelkarzinom der Haut oder Zervixkarzinom <i>in situ</i>. • Behandlung mit verbotener(n) Begleitmedikation(en) oder eine verbotene Begleitmedikation innerhalb des angegebenen Zeitrahmens vor der ersten Dosis der Studienmedikation. • Hinweise auf eine aktive oder latente oder unzureichend behandelte Infektion mit <i>Mycobacterium Tuberculosis</i> (Tuberkulose), die durch einen der folgenden Punkte nachgewiesen wurde: <ul style="list-style-type: none"> ○ Ein positiver QuantiFERON®-Tuberkulose Gold In-Tube Test (QFT-G) oder ein positiver <i>Mantoux/ Purified Protein Derivative</i> (PPD)/ Tuberkulin-Hauttest, der an oder innerhalb der letzten 12 Wochen vor Tag 1 durchgeführt wurde, ist ein Ausschlusskriterium. Ein negativer QFT-G- oder Mantoux/ PPD/ Tuberkulin-Hauttest war erforderlich, es sei denn, der Studienteilnehmer hatte zuvor eine dokumentierte adäquate Therapie für eine latente (9 Monate Isoniazid an einem Ort, an dem die Raten der primären multiresistenten Tuberkulose < 5 % waren oder ein akzeptables alternatives Regime) oder aktive (akzeptables multiresistentes Regime) Tuberkulose-Infektion erhalten. Wenn die aktuellen Inzidenzraten von multiresistenten Tuberkulose-Infektionen in der Region nicht verfügbar waren, sollte ein adäquates Behandlungsschema als das vom Gesundheitsministerium oder Expertengremium in der Region empfohlene Schema definiert werden. ○ Es wurde empfohlen, dass Studienteilnehmer mit einer Vorgeschichte einer <i>Bacille-Calmette-Guérin</i>

Item ^a	Charakteristikum	Studieninformation
		<p>(BCG)-Impfung mit dem QFT-G-Test getestet werden, da der Mantoux/ PPD/ Tuberkulin-Hauttest aufgrund der Impfung positiv sein kann.</p> <ul style="list-style-type: none"> ○ Röntgenaufnahme des Brustkorbs (oder Computertomographie des Brustkorbs, falls verfügbar) zum Zeitpunkt des Screenings mit Veränderungen, die auf eine aktive Tuberkulose-Infektion hinweisen, die von einem qualifizierten Radiologen festgestellt wurde. Eine Röntgenaufnahme des Brustkorbs war erforderlich, sofern sie nicht innerhalb von 12 Wochen vor dem ersten Studientag durchgeführt und dokumentiert wurde. ○ Anamnese einer unbehandelten oder unzureichend behandelten latenten oder aktiven Tuberkulose-Infektion. ○ Derzeitige Behandlung einer aktiven Tuberkulose-Infektion. <ul style="list-style-type: none"> • Geimpft oder exponiert gegenüber einem Lebendimpfstoff oder abgeschwächten Impfstoff innerhalb der 6 Wochen vor der ersten Dosis der Studienintervention, oder es wurde erwartet, dass der Studienteilnehmer während der Behandlung oder während der 6 Wochen nach Absetzen der Studienintervention geimpft wurde oder im Haushalt Exposition gegenüber diesen Impfstoffen hatte. • Studienteilnehmer, die zuvor eine Behandlung mit einem Januskinase (JAK)-Inhibitor erhalten hatten. Eine vorherige Behandlung mit topischen JAK-Inhibitoren war kein Ausschlusskriterium. • Frühere Behandlung mit Dupilumab und/ oder eine Vorgeschichte von Überempfindlichkeit, Unverträglichkeit, UE oder allergischen Reaktionen im Zusammenhang mit einer früheren Exposition gegenüber Hilfsstoffen von Dupilumab. • Erhalt einer der folgenden Behandlungen in den unten angegebenen Zeiträumen: <ul style="list-style-type: none"> ○ Innerhalb von einem Jahr vor der ersten Dosis der Studienintervention: <ul style="list-style-type: none"> ▪ Vorherige Behandlung mit nicht B-Zell-spezifischen lymphozytendepletierenden Mitteln/ Therapien (z. B. Alemtuzumab (CAMPATH), alkylierende Mittel (z. B. Cyclophosphamid oder Chlorambucil), totale lymphoide Bestrahlung usw.). <p>Studienteilnehmer, die Rituximab oder andere selektive B-Lymphozytendepletierende Wirkstoffe (einschließlich experimenteller Wirkstoffe) erhalten haben, waren teilnahmeberechtigt, wenn sie seit mindestens einem Jahr vor Studienbeginn keine solche Therapie erhalten hatten und</p>

Item ^a	Charakteristikum	Studieninformation
		<p>normale <i>Cluster of Differentiation</i> (CD) 19/20+ Zählungen durch Fluoreszenz-aktivierte Zellsortierung (<i>Flourescence Activated Cell Sorting</i>, FACS) Analyse aufwiesen.</p> <ul style="list-style-type: none"> ○ Innerhalb von 12 Wochen nach der ersten Dosis der Studienintervention: <ul style="list-style-type: none"> ▪ Andere Biologika: innerhalb von 12 Wochen nach der ersten Dosis der Studienintervention oder 5 Halbwertszeiten (falls bekannt), je nachdem, was länger ist. ○ Innerhalb von 4 Wochen nach der ersten Dosis der Studienintervention: <ul style="list-style-type: none"> ▪ Verwendung von oralen Immunsuppressiva (z. B. Cyclosporin A [CsA], Azathioprin, Methotrexat, systemische Kortikosteroide, Mycophenolat-Mofetil, Interferon-gamma) innerhalb von 4 Wochen nach der ersten Dosis der Studienintervention oder innerhalb von 5 Halbwertszeiten (falls bekannt), je nachdem, was länger ist. Kortikosteroid-Inhalatoren, intranasale Sprays und ophthalmische Kortikosteroide waren für Studienteilnehmer, die eine stabile Dosis erhielten, zulässig. ▪ Verwendung von CYP2C9- und CYP2C19-Induktoren innerhalb von 5 Halbwertszeiten des Induktors plus 14 Tage nach der ersten Dosis der Studienintervention. Z. B. beträgt die durchschnittliche Halbwertszeit von Carbamazepin nach wiederholter Verabreichung 15 Stunden. Die Ausschleichphase berechnet sich aus der Summe von 5 Halbwertszeiten (ca. 3 Tage) und zusätzlichen 14 Tagen für insgesamt 17 Tage vor der ersten Dosis der Studienintervention. ▪ Schmalband Phototherapie-UVB (<i>Narrowband ultraviolet B light</i>, NB-UVB) oder Breitband-Phototherapie. ▪ Regelmäßige Nutzung (mehr als 2 Visiten pro Woche) einer Bräunungskabine/ eines Solariums. ▪ Pflanzliche Arzneimittel mit unbekanntem Eigenschaften oder bekannten positiven Wirkungen bei AD. ○ Innerhalb der ersten Woche nach der ersten Dosis der Studienintervention: <ul style="list-style-type: none"> ▪ Verwendung von CYP2C9- und CYP2C19-Inhibitoren innerhalb der ersten Woche vor

Item ^a	Charakteristikum	Studieninformation
		<p>der ersten Dosis der Studienintervention oder innerhalb von 5 Halbwertszeiten (falls bekannt) des Inhibitors, je nachdem, was länger ist.</p> <ul style="list-style-type: none"> ▪ Anti-Thrombozyten-Medikamente. niedrig dosierte Acetylsalicylsäure (≤ 100 mg 1-mal täglich (QD)) war zum Zweck der kardiovaskulären Prophylaxe nach Ermessen des Prüfarztes erlaubt. <ul style="list-style-type: none"> • Teilnahme an anderen Studien mit Prüfpräparat(en) innerhalb von 8 Wochen oder innerhalb von 5 Halbwertszeiten (falls bekannt), je nachdem, was länger ist, vor Studienbeginn und/ oder während der Studienteilnahme. Jede prüfungsrelevante oder experimentelle Therapie, die für AD, Psoriasis, psoriatische Arthritis oder rheumatoide Arthritis im vorangegangenen Jahr eingenommen oder durchgeführt wurde, sollte mit dem Kliniker von Pfizer (oder einem Beauftragten) besprochen werden. Die Studienteilnehmer durften zu keinem Zeitpunkt während ihrer Teilnahme an dieser Studie an Studien mit anderen prüfungsrelevanten oder experimentellen Therapien oder Verfahren teilnehmen. • Ein 12-Kanal-<i>Screening</i>-Elektrokardiogramm (EKG), das klinisch bedeutsame Anomalien zeigte, die eine Behandlung erfordern (z. B. akuter Herzinfarkt, schwere Tachy- oder Bradyarrhythmien) oder die auf eine schwerwiegende zugrunde liegende Herzerkrankung hinweisen (z. B. Kardiomyopathie, schwere angeborene Herzerkrankung, niedrige Voltzahl in allen Ableitungen, Wolff-Parkinson-White-Syndrom) oder Kriterien, die mit Anomalien des Q-Wellen-Intervalls (QT)/ Fridericia-korrigierten Q-Wellen-Intervalls (QTcF) verbunden sind, einschließlich: <ul style="list-style-type: none"> ○ Eine deutliche Verlängerung des Fridericia-korrigierten Q-Wellen-Intervalls (> 500 Millisekunden) auf dem <i>Screening</i>-EKG; • Jede der folgenden Anomalien in den klinischen Labortests vor dem <i>Screening</i>, wie vom studienspezifischen Labor beurteilt und durch eine einmalige Wiederholung bestätigt, falls dies als notwendig erachtet wurde: <ul style="list-style-type: none"> ○ Absolute Neutrophilenzahl von $< 1,2 \times 10^9/l$ ($< 1200/mm^3$); ○ Hämoglobin $< 10,0$ g/dl oder Hämatokrit < 30 %; ○ Thrombozytenzahl von $< 150 \times 10^9/l$ ($< 150.000/mm^3$); ○ Absolute Lymphozytenzahl von $< 0,50 \times 10^9/l$ ($< 500/mm^3$); ○ Geschätzte Kreatinin-Clearance < 40 ml/min basierend auf der altersgerechten Berechnung oder Serumkreatinin $> 1,5$-mal die obere Grenze der Norm (<i>Upper limit of normal</i>, ULN);

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ○ Aspartat-Aminotransferase (AST) oder Alanin-Aminotransferase (ALT) Werte > 2-fach der ULN; ○ Gesamtbilirubin \geq 1,5-fach der ULN; <p>Studienteilnehmer mit einer Vorgeschichte des Gilbert-Syndroms konnten ein direktes Bilirubin messen lassen und wären für diese Studie geeignet, sofern das direkte Bilirubin \leq ULN erreicht.</p> <ul style="list-style-type: none"> • Nach Meinung des Prüfers oder des Sponsors jede unkontrollierte klinisch signifikante Laboranomalie, die die Interpretation der Studiendaten oder die Teilnahme des Studienteilnehmers an der Studie beeinträchtigen würde. • Signifikantes Trauma oder größere Operation innerhalb des ersten Monats vor der ersten Dosis der Studienintervention. • Mitarbeiter des Prüfzentrums, die direkt an der Durchführung der Studie beteiligt waren, und deren Familienangehörige, Mitarbeiter des Prüfzentrums, die anderweitig vom Prüfer beaufsichtigt wurden oder Studienteilnehmer, die Mitarbeiter von Pfizer waren, einschließlich deren Familienangehörige, die direkt an der Durchführung der Studie beteiligt waren.
4b	Studienorganisation und Ort der Studiendurchführung	<p><u>Weltweit etwa 220 Studienzentren</u></p> <p>Zentren in den USA, Australien, Bulgarien, Kanada, Chile, Finnland, Deutschland, Ungarn, Italien, Korea, Lettland, Polen, Slowakei, Spanien, Taiwan</p>
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p><u>Interventionsgruppe</u></p> <ul style="list-style-type: none"> • 200 mg Abrocitinib (2x 100 mg Tabletten), oral, 1-mal täglich und an Dupilumab angepasstes Placebo als subkutane (s.c.) Injektion alle 2 Wochen (2 Injektionen zu Studienbeginn; zur Vortäuschung der Ladedosis) von Tag 1 bis Woche 26 (die letzte Injektion des Placebos erfolgt zu Woche 24). <p><u>Vergleichsgruppe</u></p> <ul style="list-style-type: none"> • 300 mg Dupilumab, s.c. Injektion, alle 2 Wochen (eine Ladedosis von 600 mg zu Studienbeginn) und an Abrocitinib angepasstes Placebo (2 Tabletten), oral, 1-mal täglich von Tag 1 bis Woche 26 (die letzte Injektion von Dupilumab erfolgt zu Woche 24). <p><u>Hintergrundtherapie</u></p> <ul style="list-style-type: none"> • Nicht medikamentöse topische Therapie: Emollentien (z. B. Hyaluronsäure, Urea, Ceramide oder Filaggrin Abbauprodukte), 2-mal täglich. • Medikamentöse topische Therapie: Leichte bis mittelstark wirksame topische Glukokortikoide (TCS) oder topische Calcineurin-Inhibitoren (TCI), 1-mal täglich. <p><u>Rescue-Therapie</u></p> <ul style="list-style-type: none"> • Möglich nach 4 Wochen mit intolerablen AD-Symptomen. • Stark wirksame TCS für bis zu 2 Wochen oder systemische Glukokortikoide für bis zu 10 Tage.

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	
6a	<p>Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten</p>	<p>Primäres Zielkriterium</p> <ul style="list-style-type: none"> • Ansprechen basierend auf einer Verbesserung von ≥ 4 Punkten auf der <i>Peak Pruritus</i> NRS zu Woche 2 gegenüber <i>Baseline</i>. • Ansprechen basierend auf einer Verbesserung von ≥ 75 % des EASI zu Woche 4 gegenüber <i>Baseline</i>. <p>Sekundäre Zielkriterien</p> <ul style="list-style-type: none"> • Ansprechen basierend auf einer Verbesserung von ≥ 75 % des EASI zu Woche 16 gegenüber <i>Baseline</i>. • Ansprechen basierend auf einer Verbesserung von ≥ 75 % des EASI zu allen geplanten Beobachtungszeitpunkten bis zu Woche 26. • Ansprechen basierend auf einer Verbesserung von ≥ 90 % des EASI zu allen geplanten Beobachtungszeitpunkten bis zu Woche 26. • Ansprechen basierend auf dem IGA von 0 (normale Haut) oder 1 (fast normale Haut) auf einer 5-Punkte-Skala und einer Reduktion von ≥ 2 Punkte zu allen geplanten Beobachtungszeitpunkten bis zu Woche 26. • Ansprechen basierend auf einer Verbesserung von ≥ 4 Punkten im <i>Peak Pruritus</i> NRS zu allen geplanten Beobachtungszeitpunkten außer Woche 2 gegenüber <i>Baseline</i>. • Zeit bis zur Verbesserung von ≥ 4 Punkten im Schweregrad des <i>Peak Pruritus</i> NRS gegenüber <i>Baseline</i>. • Prozentuale Veränderung der folgenden Endpunkte zu allen geplanten Beobachtungszeitpunkten gegenüber <i>Baseline</i>: <ul style="list-style-type: none"> ○ BSA ○ <i>Scoring Atopic Dermatitis</i> (SCORAD) • Veränderung der <i>Scores</i> der folgenden Endpunkte zu allen geplanten Beobachtungszeitpunkten gegenüber <i>Baseline</i>: <ul style="list-style-type: none"> ○ <i>Hospital Anxiety and Depression Scale</i> (HADS) ○ <i>Dermatology Life Quality Index</i> (DLQI) ○ <i>EuroQol Quality of Life 5-Dimension 5-Level Scale</i> (EQ-5D-5L) ○ <i>Patient-Oriented Eczema Measure</i> (POEM) ○ <i>Medical Outcomes Study – Sleep Scale</i> (MOS Sleep Scale) ○ <i>Skin Pain</i> NRS • Therapiefreie Tage in Bezug auf die topische Hintergrundtherapie. • Inzidenz von UE. • Inzidenz von schwerwiegenden unerwünschten Ereignissen (SUE). • Inzidenz von klinischen Auffälligkeiten und Veränderung der klinischen Laborwerte, EKG-Messungen und Vitalparameter gegenüber <i>Baseline</i>.

Item ^a	Charakteristikum	Studieninformation
		<p>Weitere Ziele</p> <ul style="list-style-type: none"> • Veränderung im HCRU zu allen geplanten Beobachtungszeitpunkten gegenüber <i>Baseline</i>. • Anzahl der Kratzepisoden während der Schlafperiode, die vor der Behandlung im Vergleich zu Behandlung auftraten, ermittelt aus Datenanalysen mittels tragbaren Accelerometern zu allen geplanten Beobachtungszeitpunkten. • Dauer der Kratzepisoden während der Schlafperiode, die vor der Behandlung im Vergleich zu Behandlung auftraten, ermittelt aus Datenanalysen mittels tragbaren Accelerometern zu allen geplanten Beobachtungszeitpunkten. • Anzahl der Gesamtschlafgelegenheiten, Gesamtschlafzeit, prozentuale Einschlafzeit, Aufwachen nach dem Einsetzen des Schlafes (<i>Wake after sleep onset</i>, WASO), Einschlaf-Latenz und Anzahl der Aufwachphasen, die vor der Behandlung im Vergleich zu Behandlung auftraten, ermittelt aus Datenanalysen mittels tragbaren Accelerometern zu allen geplanten Beobachtungszeitpunkten. • Ansprechen basierend auf dem IGA von 0 (normale Haut) oder 1 (fast normale Haut) und einer Reduktion von ≥ 2 Punkte zu allen geplanten Beobachtungszeitpunkten gegenüber <i>Baseline</i>. • Potenzielle Ergebnisse der explorativen Analyse der gelagerten Bioproben (diese Ergebnisse konnten, mussten aber nicht im Rahmen der vorliegenden Studie generiert werden).
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Die Studie wurde mit einer adäquaten Power geplant, um definierte Effektgrößen für die entsprechenden Endpunkte zu detektieren. Die Fallzahlplanung basierte auf einer 1:1 Randomisierung mit jeweils 300 Studienteilnehmern pro Arm und einer Gesamtstichprobe von 600 Studienteilnehmern. Der vorgesehene Stichprobenumfang bat eine ausreichende Power (~80 %) für alle Überlegenheits-hypothesen, wie folgt:</p> <ul style="list-style-type: none"> • Woche 2: Die Stichprobenzahl verleiht dem statistischen Test (Cochran-Mantel-Haenszel, CMH) eine Power von mindestens 99 %, um einen Unterschied von mindestens 25 % in der <i>Peak Pruritus</i> NRS-Ansprechrates zwischen Abrocitinib und Dupilumab nachzuweisen, unter der Annahme, dass die Ansprechrates mit Dupilumab 11 % zu Woche 2 betrug. • Woche 4: Die Stichprobenzahl verleiht dem statistischen Test (CMH) eine Power von mindestens 97 %, um einen Unterschied von mindestens 15 % in der EASI-75-Ansprechrates zwischen Abrocitinib und Dupilumab nachzuweisen, unter der Annahme, dass die Ansprechrates mit Dupilumab 30 % zu Woche 4 betrug.

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Woche 16: Die Stichprobenzahl verleiht dem statistischen Test (CMH) eine Power von mindestens 99 %, um einen Unterschied von höchstens 10 % in der EASI-75-Ansprechrage zwischen Abrocitinib und Dupilumab zugunsten von Dupilumab mit einer Nichtunterlegenheitsgrenze (<i>non-inferiority margin</i>) von 10 % nachzuweisen, unter der Annahme, dass der Unterschied zwischen Abrocitinib und Dupilumab 10 % betrug und die Ansprechrage mit Dupilumab 66 % zu Woche 16 betrug. Dies ergab eine Teststärke von etwa 80 % für den Nachweis der Überlegenheit von Abrocitinib 200 mg 1-mal täglich gegenüber Dupilumab, gemessen mit der EASI-75-Ansprechrage zu Woche 16. • Der α-Fehler wurde für das Testen der primären und wichtigen sekundären Endpunkte auf 5 % (zweiseitig) kontrolliert, wobei eine <i>Closed-Testing-Method</i> auf Grundlage eines sequenziellen, multiplen Ansatzes verwendet wurde.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Zwischenanalysen Es waren keine Zwischenanalysen geplant.</p> <p>Vorzeitiger Studienabbruch</p> <ul style="list-style-type: none"> • Es liegen keine statistischen Kriterien zum Abbruch der Studie vor. • Der Sponsor konnte zu jeder Zeit die Studie aus klinischen und administrativen Gründen bzw. die Beteiligung von Prüfarzten oder Studienzentren aufgrund niedriger Rekrutierungsraten oder fehlender <i>Compliance</i> beenden. • Die Gabe der Studienintervention konnte u. a. aus den folgenden Gründen abgebrochen werden: <ul style="list-style-type: none"> ○ Deutliche Verlängerung des Fridericia-korrigierten Q-Wellen-Intervalls auf > 500 ms oder eine Veränderung von > 60 ms des <i>Screening</i>-EKGs. ○ Eine schwerwiegende Infektion muss zu einer vorübergehenden Unterbrechung der Studienintervention führen. Die Studienintervention konnte erst dann wieder aufgenommen werden, wenn die schwere Infektion abgeklungen war und dies mit dem medizinischen Monitor besprochen und vereinbart wurde. Wenn der Studienteilnehmer die Studienintervention innerhalb von 28 Tagen nicht wieder aufnehmen konnte oder die Infektion nicht abgeklungen war oder keine Zustimmung des medizinischen Monitors vorlag, musste der Studienteilnehmer dauerhaft aus der Studie ausgeschlossen werden. ○ Jedes Blutungsereignis, das nach dem Urteil des Prüfarztes (oder, falls erforderlich/ gewünscht, nach Rücksprache mit dem Sponsor) mit einer Verringerung der Thrombozytenzahl oder des Hämoglobins in Verbindung gebracht werden konnte. ○ UE, das nach dem Urteil des Prüfarztes einen Studienabbruch erforderte (oder, falls erforderlich/ gewünscht, nach Rücksprache mit dem Sponsor).

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ○ Jedes UE oder jede Laboranomalie, die nach dem Urteil des Prüfarztes das Aussetzen der Studienintervention für > 21 Tage erforderte. ○ Jeder anfängliche Laborwert, der darunter lag, musste innerhalb von 48 Stunden erneut getestet werden. ○ Thrombozytenzahlen von < 50.000/mm³ bei 2 aufeinanderfolgenden Messungen. Wenn der Studienteilnehmer eine Thrombozytenzahl von < 25.000/mm³ hatte, sollte die Studienbehandlung vorübergehend ausgesetzt werden, bis der Bestätigungstest erneut durchgeführt wurde. ○ Neutrophilenzahlen von < 500/mm³ bei 2 aufeinanderfolgenden Messungen. ○ Lymphozytenzahlen von < 500/mm³ bei 2 aufeinanderfolgenden Messungen. ○ Hämoglobinwerte < 8,0 g/dl und/ oder eine Abnahme von > 30 % gegenüber <i>Baseline</i> bei 2 aufeinanderfolgenden Messungen. ○ Einer der folgenden Gründe: <ul style="list-style-type: none"> ▪ 2 sequenzielle AST- oder ALT-Werte > 3-fach der ULN mit mindestens einem Gesamtbilirubinwert > 2-fach der ULN. ▪ 2 sequenzielle AST- oder ALT-Werte > 3-fach der ULN mit einem abnormalen <i>International Normalized Ratio</i>. ▪ 2 sequenzielle AST- oder ALT-Werte > 3-fach der ULN mit Symptomen, die auf eine Leberschädigung hindeuten. ▪ 2 sequenzielle AST- oder ALT-Werte > 5-fach der ULN unabhängig vom Gesamtbilirubin oder begleitenden Symptomen. ○ Anstiege des Serumkreatinins, die > 50 % über dem Durchschnitt der <i>Screening</i>- und <i>Baseline</i>-Werte liegen bei 2 aufeinanderfolgenden Messungen und ein absoluter Anstieg des Serumkreatinins um ≥ 0,5 mg/dl bei 2 aufeinanderfolgenden Messungen. Sollte ein Studienteilnehmer zum Zeitpunkt des Studienabschlusses oder -abbruchs Erhöhungen des Serumkreatinins aufweisen, die ≥ 33 % über dem Durchschnitt der <i>Screening</i>- und <i>Baseline</i>-Werte lagen, wurde er alle 1 bis 2 Wochen erneut getestet, bis sich die Erhöhung des Serumkreatinins vollständig auf weniger als 10 % des Durchschnitts der <i>Screening</i>- und <i>Baseline</i>-Werte zurückgebildet oder stabilisiert hatte.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Zuteilung auf die Studienarme erfolgte mittels einer <i>Interactive Response Technology</i> (IRT) (<i>Interactive Web-based response</i> , IWR).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierungsverhältnis 1:1 (Abrocitinib 200 mg vs. Dupilumab 300 mg)

Item ^a	Charakteristikum	Studieninformation
		Stratifizierung <ul style="list-style-type: none"> Schweregrad der AD zu <i>Baseline</i> (IGA = 3 vs. IGA = 4) Alter (< 18 Jahre vs. ≥ 18 Jahre)
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Randomisierung erfolgte zentral nach einem computergenerierten Randomisierungsplan mittels IRT (IWR) nach der Zuteilung einer individuellen Patientennummer.
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?	Zuteilung mit Hilfe der IRT (IWR) gemäß computergenerierter Randomisierungsliste.
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) Verblindet b) Verblindet c) Verblindet
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Um die Verblindung zu gewährleisten, erfolgte in jeder Studiengruppe die tägliche Gabe von 2 Tabletten (Abrocitinib 100 mg oder Placebo) und alle 2 Wochen eine s.c. Injektion (Dupilumab 300 mg oder Placebo).
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Primäre Zielkriterien <u>Peak Pruritus NRS</u> <ul style="list-style-type: none"> Binäre Analyse des Ansprechens zu Woche 2: <ul style="list-style-type: none"> Analyse auf Basis der FAS-Population. Die Studienteilnehmer mussten einen <i>Peak Pruritus NRS Baseline</i>-Wert von ≥ 4 haben, um in die Analyse aufgenommen zu werden. Auswertung mittels stratifizierter Cochran-Mantel-Haenszel-Methode und Normalapproximation.

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ○ Angabe von Anzahl der analysierten Patienten, Anzahl und Anteil der Patienten mit Ansprechen inklusive 95 %-KI und Punktschätzer des Unterschieds zwischen den Behandlungsarmen mit 95 %-KI und zweiseitigem p-Wert zu jedem spezifiziertem Beobachtungszeitpunkt. ○ Imputation von fehlenden Werten als <i>Non-Responder</i>. <p><u>EASI-90</u></p> <ul style="list-style-type: none"> ● Binäre Analyse des Ansprechens zu Woche 4: <ul style="list-style-type: none"> ○ Analyse auf Basis der FAS-Population. ○ Auswertung mittels stratifizierter Cochran-Mantel-Haenszel-Methode und Normalapproximation. ○ Angabe von Anzahl der analysierten Patienten, Anzahl und Anteil der Patienten mit Ansprechen inklusive 95 %-KI und Punktschätzer des Unterschieds zwischen den Behandlungsarmen mit 95 %-KI und zweiseitigem p-Wert. ○ Imputation von fehlenden Werten als <i>Non-Responder</i>. <p>Sekundäre Zielkriterien</p> <p><u>EASI-90</u></p> <ul style="list-style-type: none"> ● Binäre Analyse des Ansprechens zu Woche 2, 8, 12, 16 und 20: <ul style="list-style-type: none"> ○ Analyse auf Basis der FAS-Population. ○ Auswertung mittels stratifizierter Cochran-Mantel-Haenszel-Methode und Normalapproximation. ○ Angabe von Anzahl der analysierten Patienten, Anzahl und Anteil der Patienten mit Ansprechen inklusive 95 %-KI und Punktschätzer des Unterschieds zwischen den Behandlungsarmen mit 95 %-KI und zweiseitigem p-Wert. ○ Imputation von fehlenden Werten als <i>Non-Responder</i>. <p><u>EASI-75 und IGA</u></p> <ul style="list-style-type: none"> ● Binäre Analyse des Ansprechens zu Woche 2, 4, 8, 12, 16, 20 und 26: <ul style="list-style-type: none"> ○ Analyse auf Basis der FAS-Population. ○ Auswertung mittels stratifizierter Cochran-Mantel-Haenszel-Methode und Normalapproximation. ○ Angabe von Anzahl der analysierten Patienten, Anzahl und Anteil der Patienten mit Ansprechen inklusive 95 %-KI und Punktschätzer des Unterschieds zwischen den Behandlungsarmen mit 95 %-KI und zweiseitigem p-Wert zu jedem spezifiziertem Beobachtungszeitpunkt. ○ Imputation von fehlenden Werten als <i>Non-Responder</i>. <p><u>Peak Pruritus NRS</u></p> <ul style="list-style-type: none"> ● Binäre Analyse des Ansprechens von Tag 2 bis 15 und zu Woche 4, 8, 12, 16, 20 und 26: <ul style="list-style-type: none"> ○ Analyse auf Basis der FAS-Population.

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ○ Die Studienteilnehmer mussten einen <i>Peak Pruritus</i> NRS <i>Baseline</i>-Wert von ≥ 4 haben, um in die Analyse aufgenommen zu werden. ○ Auswertung mittels stratifizierter Cochran-Mantel-Haenszel-Methode und Normalapproximation. ○ Angabe von Anzahl der analysierten Patienten, Anzahl und Anteil der Patienten mit Ansprechen inklusive 95 %-KI und Punktschätzer des Unterschieds zwischen den Behandlungsarmen mit 95 %-KI und zweiseitigem p-Wert zu jedem spezifiziertem Beobachtungszeitpunkt. ○ Imputation von fehlenden Werten als <i>Non-Responder</i>. ● Ereigniszeitanalyse (Zeit bis zur Verbesserung von ≥ 4 Punkten in der <i>Peak Pruritus</i> NRS gegenüber <i>Baseline</i>): <ul style="list-style-type: none"> ○ Analyse auf Basis der FAS-Population. ○ Die Studienteilnehmer mussten einen <i>Peak Pruritus</i> NRS <i>Baseline</i>-Wert von ≥ 4 haben, um in die Analyse aufgenommen zu werden. ○ Bestimmung des Medians und der Quartile inklusive 95 %-KI sowie der Wahrscheinlichkeit zu spezifischen Beobachtungszeitpunkten ereignisfrei zu sein (Kaplan-Meier-Methode und Schätzung des 95 %-KI). ○ Auswertung des Unterschieds zwischen den Behandlungsgruppen mittels des p-Wertes anhand des stratifizierten Log-Rank-Tests (stratifiziert nach Schweregrad der Erkrankung zu <i>Baseline</i>). ○ Wenn kein Ereignis aufgetreten ist, wurden die Zeit bis zum Ereignis rechtszensiert zum letzten Beobachtungszeitpunkt. <p><u>DLQI</u></p> <ul style="list-style-type: none"> ● Binäre Analyse des Ansprechens zu Woche 4, 12, 16, 20 und 26: <ul style="list-style-type: none"> ○ Analyse auf Basis der FAS-Population. ○ Die Studienteilnehmer mussten einen DLQI <i>Baseline</i>-Wert von ≥ 4 haben, um in die Analyse aufgenommen zu werden. ○ Auswertung mittels stratifizierter Cochran-Mantel-Haenszel-Methode und Normalapproximation. ○ Angabe von Anzahl der analysierten Patienten, Anzahl und Anteil der Patienten mit Ansprechen inklusive 95 %-KI und Punktschätzer des Unterschieds zwischen den Behandlungsarmen mit 95 %-KI und zweiseitigem p-Wert zu jedem spezifiziertem Beobachtungszeitpunkt. ○ Imputation von fehlenden Werten als <i>Non-Responder</i>. ● Kontinuierliche Analyse des Ansprechens zu Woche 4, 12, 16, 20 und 26: <ul style="list-style-type: none"> ○ Analyse auf Basis der FAS-Population.

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ○ Die Studienteilnehmer mussten einen DLQI <i>Baseline</i>-Wert haben, um in die Analyse aufgenommen zu werden. ○ Auswertung anhand MMRM-Analyse mit den Variablen <i>Behandlung</i>, <i>Zeit</i>, <i>Behandlung*Zeit</i> und <i>Schweregrad der Erkrankung</i> zu <i>Baseline</i> als Kovariate mit einer unstrukturierten Kovarianzmatrix. ○ Angabe der Veränderungen zu <i>Baseline</i> anhand der im MMRM-Modell geschätzten <i>Least Squares (LS)-Means</i> und Differenzen der <i>LS-Means</i> zwischen den Behandlungsgruppen sowie die korrespondierenden p-Werte und 95 %-KI zu jeder Visite. ○ Keine Imputation von fehlenden Werten. <p><u>BSA-Ansprechen, SCORAD zu Woche 2, 4, 8, 12, 16, 20 und 26, HADS, EQ-5D-5L, POEM, MOS <i>Sleep Scale</i> zu Woche 12, 16 und 26, <i>Skin Pain NRS</i> zu Woche 2, 12, 16, 20 und 26:</u></p> <ul style="list-style-type: none"> ● Kontinuierliche Analyse des Ansprechens: <ul style="list-style-type: none"> ○ Analyse auf Basis der FAS-Population. ○ Die Studienteilnehmer mussten einen <i>Baseline</i>-Wert haben, um in die Analyse aufgenommen zu werden. ○ Auswertung anhand der MMRM-Analyse mit den Variablen <i>Behandlung</i>, <i>Zeit</i>, <i>Behandlung*Zeit</i> und <i>Schweregrad der Erkrankung</i> zu <i>Baseline</i> als Kovariate mit einer unstrukturierten Kovarianzmatrix. ○ Angabe der Veränderungen zu <i>Baseline</i> anhand der im MMRM-Modell geschätzten <i>LS-Means</i> und Differenzen der <i>LS-Means</i> zwischen den Behandlungsgruppen sowie die korrespondierenden p-Werte und 95 %-KI zu jeder Visite. ○ Keine Imputation von fehlenden Werten. <p><u>Therapiefreie Tage in Bezug auf die topische Hintergrundtherapie</u></p> <ul style="list-style-type: none"> ● Analyse auf Basis der FAS-Population. ● Auswertung anhand ANCOVA. ● Keine Imputation von fehlenden Werten. <p>Studienpopulationen</p> <p><u>FAS-Population (<i>Full-Analysis-Set</i>)</u></p> <ul style="list-style-type: none"> ● Diese Population umfasst allen Studienteilnehmern, die der Studienintervention randomisiert zugeteilt wurden und mindestens eine Studiendosis erhalten haben. Die Auswertung erfolgte gemäß der randomisierten Behandlungszuteilung. <p><u>PPAS-Population (<i>Per Protocol-Analysis-Set</i>)</u></p> <ul style="list-style-type: none"> ● Diese Population umfasst eine Subgruppe des FAS, die keine wesentlichen Protokolländerungen aufwiesen. Die PPAS-Population umfasste Studienteilnehmer, die: <ul style="list-style-type: none"> ○ das Einschlusskriterium für das qualifizierende Alter erfüllten, ○ das Einschlusskriterium $EASI \geq 16$ erfüllten,

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> ○ das Einschlusskriterium für eine vorherige qualifizierende Behandlung erfüllten, ○ beobachtete Werte im EASI zu Woche 16 hatten, ○ eine Gesamt-Einhaltung von $\geq 80\%$ aber $\leq 120\%$ der randomisierten Behandlung zugewiesenen Menge zu Woche 16 hatten und ○ keine größeren Protokollabweichungen aufwiesen, die vermutlich einen wesentlichen Einfluss auf die klinischen Beobachtungen oder die vom klinischen Team ermittelten Reaktionen des Patienten hatten. <p><u>Safety-Population (Safety-Analysis-Set)</u></p> <ul style="list-style-type: none"> ● Diese Population umfasst alle Studienteilnehmer, die der Studienintervention randomisiert zugeteilt wurden und mindestens eine Studiendosis erhalten haben. Die Auswertung erfolgte gemäß der tatsächlich erhaltenen Therapie.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Subgruppenanalysen</p> <p><u>Analysen der primären und sekundären Endpunkte</u></p> <ul style="list-style-type: none"> ● Alter in Jahren (< 40 vs. ≥ 40 vs. < 65 vs. ≥ 65) ● Geschlecht (Männlich vs. Weiblich) ● Abstammung (weiß vs. schwarz oder afroamerikanisch vs. asiatisch vs. andere) ● Region der Studieneinschreibung: (USA/ Kanada/ Australien vs. Europa vs. Asien) ● Gewicht in kg (< 70 vs. 70-100 vs. > 100 kg) ● Dauer der AD in Jahren (< 26 vs. ≥ 26) ● Schweregrad der Erkrankung zu <i>Baseline</i> (IGA = 3 vs. IGA = 4) ● EASI-Score zu <i>Baseline</i> (16-25 vs. > 25) ● Betroffene BSA zu <i>Baseline</i> (10-30 % vs. > 30-50 % vs. > 50 %) ● Vorbehandlung mit Cyclosporin (ja vs. nein) ● Vorbehandlung mit AD-Medikation (nur topische Mittel vs. systemische Mittel) <p>Sensitivitätsanalysen</p> <p><u>Peak Pruritus NRS</u></p> <ul style="list-style-type: none"> ● Sensitivitätsanalyse des Ansprechens anhand der FAS-Population unter Verwendung von multipler Imputation mit den Daten zu Woche 2, 4, 8, 12, 16, 20 und 26. Die Studienteilnehmer mussten einen <i>Peak Pruritus NRS Baseline</i>-Wert von ≥ 4 haben, um in die Analyse aufgenommen zu werden. <p><u>EASI-90</u></p> <ul style="list-style-type: none"> ● Sensitivitätsanalyse des Ansprechens anhand der FAS-Population unter Verwendung von multipler Imputation mit den Daten zu Woche 2, 4, 8, 12, 16, 20 und 26.

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> Sensitivitätsanalyse des Ansprechens anhand der PPAS-Population unter Verwendung der gleichen Methodik wie bei der Hauptanalyse.
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) Abrocitinib: N = 362 Dupilumab: N = 365 b) Abrocitinib: N = 362 Dupilumab: N = 365 c) Abrocitinib: N = 362 Dupilumab: N = 365
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart (siehe Abbildung 4-21)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studienbeginn: 11.06.2020 (erste Visite des ersten Studienteilnehmers) Studienende: 13.07.2021 (letzte Nachbeobachtung des letzten Patienten)
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant beendet.
a: nach CONSORT 2010. AD: Atopische Dermatitis; ALT: Alanin-Aminotransferase; ANCOVA: <i>Analysis of Covariance</i> ; AST: Aspartat-Aminotransferase; BSA: Körperoberfläche (<i>Body Surface Area</i>); BCG: Bacille Calmette Guérin; CaPS: CDISC und Pfizer Standards; CD: <i>Cluster of differentiation</i> ; CMH: Cochran-Mantel-Haenszel; CsA: Cyclosporin A; C-SSRS: <i>Columbia Suicide Severity Rating Scale</i> ; CONSORT: <i>Consolidated Standards of Reporting Trials</i> ; d. h.: Das heißt; DLQI: <i>Dermatology Life Quality Index</i> ; DNA: Desoxyribonukleinsäure (<i>Deoxyribonucleic Acid</i>); EASI: <i>Eczema Area and Severity Index</i> ; EBV: Epstein Barr Virus; ED: Vorzeitiger Studienabbruch (<i>Early discontinuation</i>); EKG: Elektrokardiogramm; EOT: Behandlungsende (<i>End of treatment</i>); EQ-5D-5L: <i>European Quality of Life 5 Dimension 5 Level</i> ; FACS: Fluoreszenz-aktivierte Zellsortierung (<i>Flourescence Activated Cell Sorting</i>); FAS: <i>Full-Analysis-Set</i> ; FSH: Follikelstimulierendes Hormon; HADS: <i>Hospital Anxiety and Depression Scale</i> ; HBV: Hepatitis-B-Virus; HCRU: Inanspruchnahme von Gesundheitsressourcen (<i>Health Care Resource Utilization</i>); HIV: Humanes Immundefizienz-Virus; IGA: <i>Investigator's Global Assessment</i> ; INR: <i>International normalized ratio</i> ; IRT: <i>Interactive Response</i>		

Item ^a	Charakteristikum	Studieninformation
<p><i>Technology</i>; IWR: Interaktive webbasierte Antwort (<i>Interactive Web-based Response</i>); JAK: Januskinase; kg: Kilogramm; KI: Konfidenzintervall; LLQ: Untere Quantifizierungsgrenze (<i>Lower Limit of Quantification</i>); LS: <i>Least Squares</i>; mIU: Tausend Internationale Units; ml: Milliliter; mm³: Kubikmillimeter; MMRM: <i>Mixed effect Model Repeated Measurements</i>; MOS: <i>Medical Outcomes Study</i>; MRT: Magnetresonanztomographie; ms: Millisekunden; NB-UVB: Schmalband Phototherapie-UVB (<i>Narrowband ultraviolet B light</i>); OTC: Rezeptfrei, nicht verschreibungspflichtig (<i>Over-the-counter</i>); NRS: <i>Numerical Rating Scale</i>; PHQ-8: <i>Patient Health Questionnaire 8-Items</i>; POEM: <i>Patient-Oriented Eczema Measure</i>; PPAS: <i>Per-Protocol Analysis-Set</i>; PPD: <i>Purified Protein Derivative Test</i>; PRO: Patientenberichtete Endpunkte (<i>Patient Reported Outcomes</i>); QD: 1-mal täglich (<i>Once daily</i>); QFT-G: <i>QuantiFERON®-TB Gold In-Tube Test</i>; SAP: Statistischer Analyseplan; SAS: <i>Safety-Analysis-Set</i>; s.c.: Subkutan; SCORAD: <i>Scoring Atopic Dermatitis</i>; SUE: Schwerwiegendes unerwünschtes Ereignis; TCI: Topische Calcineurin-Inhibitoren; TCS: Topische Glukokortikosteroide (<i>Topical corticosteroids</i>); u. a.: Unter anderem; UE: Unerwünschtes Ereignis; ULN: Obere Grenze der Norm (<i>Upper limit of normal</i>); UVB: Ultraviolettes Licht (<i>Ultraviolet B light</i>); USA: <i>United States of America</i>; usw.: Und so weiter; vs.: Versus; WASO: Aufwachen nach dem Einsetzen des Schlafs (<i>Wake after sleep onset</i>)z. B.: Zum Beispiel</p>		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

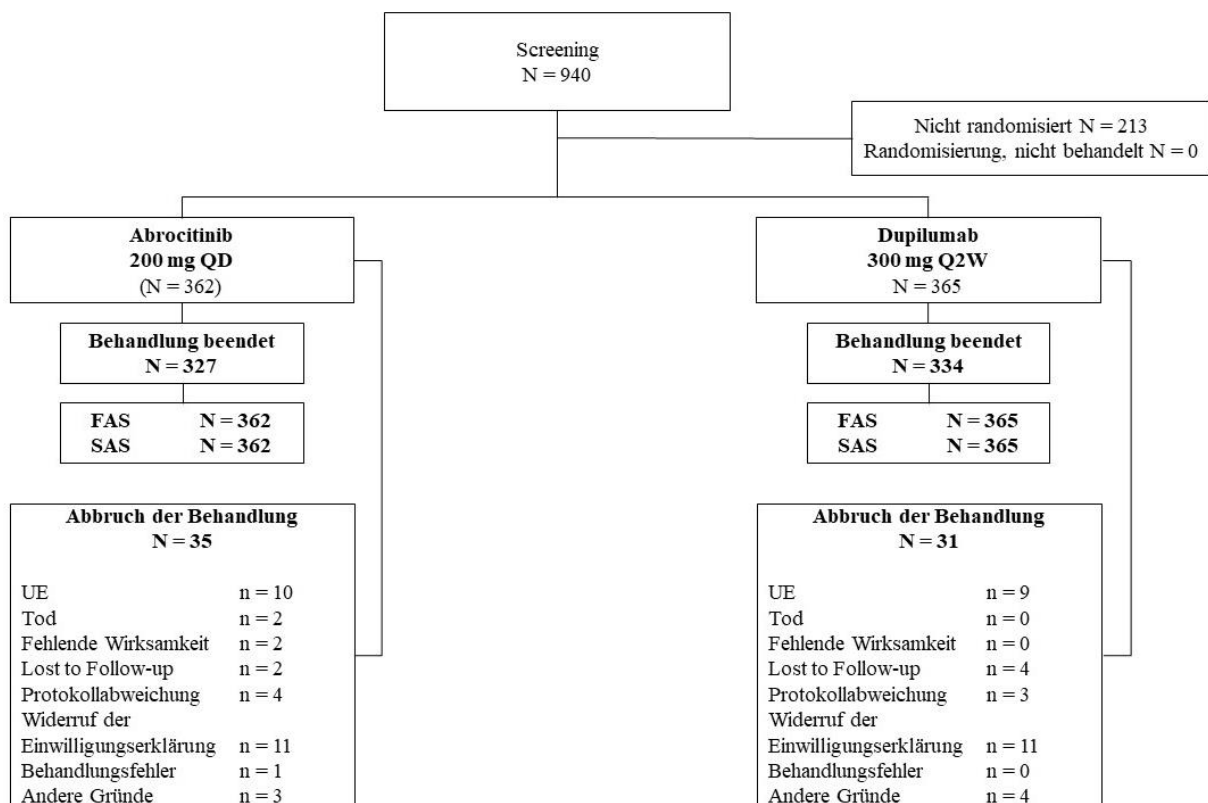


Abbildung 4-21: Flow-Chart der Studie JADE DARE

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 4-129 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie JADE DARE

Studie: JADE DARE

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienprotokoll der Studie JADE DARE vom 14.08.2020 (52)	CSP
Statistischer Analyseplan der Studie JADE DARE vom 21.07.2021 (59)	SAP
Studienbericht der Studie JADE DARE vom 25.10.2021 (56)	CSR

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Bei der Studie JADE DARE handelt es sich um eine randomisierte, aktivkontrollierte Phase-III-Studie (CSP, SAP, CSR).

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

ja **unklar** **nein**

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.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient:** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblind durchgeführte, randomisierte, aktivkontrollierte Phase-III-Studie (CSP, SAP, CSR).

behandelnde bzw. weiterbehandelnde Personen: ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblind durchgeführte, randomisierte, aktivkontrollierte Phase-III-Studie (CSP, SAP, CSR).

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine multizentrische, randomisierte, doppelblinde und aktivkontrollierte Phase-III-Studie. Die Randomisierungssequenz wurde adäquat mittels eines computerbasierten IRT-Systems erzeugt und die Gruppenzuteilung erfolgte verdeckt. Patienten und behandelnde Personen waren verblindet. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte oder Hinweise für eine ergebnisabhängige Berichterstattung identifiziert werden. Somit wird das Verzerrungspotenzial für die Studie JADE DARE als niedrig eingestuft (CSP, SAP, CSR).

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: EASI

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: SCORAD**1. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: POEM**1. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: Peak Pruritus NRS**1. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: SCORAD VAS**1. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: MOS *Sleep Scale***1. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: *Skin Pain* NRS**5. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

6. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: HADS**1. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: EQ-5D VAS**1. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: DLQI**1. Verblindung der Endpunkterheber**

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): **niedrig** **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt. Abweichungen von der im SAP präspezifizierten Methodik sind durch die methodischen Anforderungen der Nutzenbewertung begründet. Die Darstellung der Studienergebnisse erfolgte daher nicht ergebnisabhängig. Des Weiteren konnten keine das Verzerrungspotenzial beeinflussenden Aspekte identifiziert werden. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

Endpunkt: Verträglichkeit**1. Verblindung der Endpunkterheber** **ja** **unklar** **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei der Studie JADE DARE handelt es sich um eine doppelblinde Phase-III-Studie (CSP, SAP, CSR).

2. Adäquate Umsetzung des ITT-Prinzips **ja** **unklar** **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** **ja** **unklar** **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** **ja** **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Durch die Verblindung der Patienten und Behandler ist gegeben, dass die Erhebung der Endpunkte ohne Kenntnis über die tatsächlich verabreichte Studienmedikation stattfand. Für die Analysen zur Sicherheit wurde das SAS herangezogen, welches alle randomisierten Patienten beinhaltet, die mindestens eine Dosis der Studienmedikation erhalten haben. Des Weiteren wurden keine das Verzerrungspotenzial beeinflussenden Aspekte oder eine ergebnisgesteuerte Berichterstattung identifiziert. Somit wird das Verzerrungspotenzial für diesen Endpunkt als niedrig eingestuft (CSP, SAP, CSR).

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: endpunkübergreifend; B: endpunktspezifisch).

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)** **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

 unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau. **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren** **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

 unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau. **nein:** Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient:** **ja:** Die Patienten waren verblindet. **unklar:** Es finden sich keine diesbezüglichen Angaben. **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein.
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:
